Neural correlates of an illusionary sense of agency caused by virtual reality

Yiyang Cai[1,2], Huichao Yang[3], Ziyi Xiong[3], Simone Kühn[5,6], Yanchao Bi[3,4], Kunlin Wei[1,2*]

[1]School of Psychological and Cognitive Sciences and Beijing Key Laboratory of Behavior and Mental Health, Peking University, Beijing 100871, China

[2]Key Laboratory of Machine Perception (Ministry of Education), Peking University, Beijing 100871, China

[3]State Key Laboratory of Cognitive Neuroscience and Learning & IDG/McGovern Institute for Brain Research, Beijing Normal University, Beijing 100875, China,

[4]Chinese Institute for Brain Research, Beijing 102206, China

[5]Department of Psychiatry and Psychotherapy, University Medical Center Hamburg-Eppendorf, 20251 Hamburg, Germany

[6]Lise Meitner Group for Environmental Neuroscience, Max Planck Institute for Human Development, 14195 Berlin, Germany

Corresponding author (*)

Kunlin Wei (wei.kunlin@pku.edu.cn)

School of Psychological and Cognitive Sciences, Peking University, 5 Yiheyuan Road, Beijing 100871, China

Acknowledgments

This work was supported by the Ministry of Science and Technology of China (2021ZD0202601) and the National Natural Science Foundation of China (62061136001, 32071047, 31871102) awarded to KW, and the National Natural Science Foundation of China (32100837) to HY, and the German Research Foundation (TRR 169/C8) to SK. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Keywords: sense of agency; virtual reality; self-consciousness; binding; embodiment

Abstract

As a part of self-consciousness, sense of agency (SoA) refers to our sensation that we initiate actions that lead to ensuing perceptual consequences. The prospective theory of agency posits that motor prediction and its comparison with the subsequent action outcome underlies SoA. In contrast, the reconstructive theory posits that retrospective causal processing of action outcome leads to SoA without the necessity of action planning. Previous neuroimaging studies provide direct neural support for the prospective theory by showing the agency network includes brain regions for action planning and intentionality, but the equivalent neural evidence for the reconstructive theory is lacking despite behavioral evidence. Here we utilized virtual reality (VR) technology to induce an illusionary SoA in the absence of voluntary action and examined it by functional magnetic resonance imaging (fMRI). Participants' experience of controlling an avatar hand elicited a persistent increase in temporal binding, a behavioral marker of SoA when they merely passively observed the avatar hand's movement in the scanner. fMRI results showed that the increased temporal binding was associated with a cluster centered in the right angular gyrus, extending to the right inferior parietal lobule and the right precuneus. These associative areas engage in causal processing and inferential sensemaking in various tasks beyond agency. In contrast, other traditional agency-related brain areas, especially those governing action intentionality and motor planning, were not implicated. Our findings not only provide direct neural support for the reconstructive theory of agency but also suggest that the experience of controlling an avatar boosts the inferential processing of "virtual" actions that reside in posterior parietal regions, namely the angular gyrus.

Introduction

Sense of agency (SoA) is the sensation that self-initiated actions influence the external environment. We implicitly experience the feeling of the connection between our action and the resulting action consequence and attend to its disruption only when the actual action feedback conflicts with our expected consequences. As an integral part of self-consciousness, agency enables one to feel fluent control over one's surroundings (Haggard, 2017), distinct from others (Kahl & Kopp, 2018), and responsible for one's own actions (Haggard & Tsakiris, 2009). A clear understanding of the computations underlying agency is still lacking but two major competing theories are currently attested. The prospective theory emphasizes that agency is based on a predictive process in the motor system and a comparative process for comparing the predicted and actual action feedback (Frith et al., 2000; Gallagher, 2000; Haggard, 2005; Wolpert & Ghahramani, 2000). The predictive process uses the efference copy of the current motor commands to generate expectations of action consequences. A mismatch between the prediction and actual sensory feedback can disrupt the otherwise fluent SoA. The alternative reconstructive theory emphasizes that agency arises from retrospective explanations of sensory feedback after movement (Buehner & Humphreys, 2009; Wegner, 2003; Wegner & Wheatley, 1999). This inferential sensemaking process evaluates the action feedback and its contingency with prior intentions and goals, reconstructing the causal links between them. While both theories depend on the processing of sensory feedback, they differ in predictive aspects of action planning: the prospective theory necessitates the sensory prediction of action consequence, while the reconstructive theory does not rely on action planning but necessitates post-movement inferential processing of action feedback.

A strong neural support for the prospective theory would be a positive correlation between the magnitude of agency measures and the activity of action planning-specific brain regions. Recently,

Zapparoli and colleagues found that intentional binding, an implicit measure of SoA (Moore & Obhi, 2012), was associated with the activity in a brain network including the pre-supplementary motor area (pre-SMA) and dorsal parietal cortex (Zapparoli et al., 2020). The intentional binding often refers to the fact that people's timing judgment of an action (e.g., a key press) and its delayed outcome (e.g., a beep sound or flash) are biased towards each other whenever the action is voluntary as compared to involuntarily made (e.g., the finger pushed by others or triggered by transcranial magnetic stimulation (TMS); Haggard et al., 2002). Importantly, Zapparoli and colleagues found that repetitive TMS applied over the pre-SMA at the motor planning phase modulated the binding effect (see similar findings by the use of transcranial direct current stimulation and theta-burst TMS, Cavazzana et al., 2015; Moore et al., 2010). In contrast, rTMS applied over the pre-SMA during the outcome presentation or over non-planning areas (the dorsal parietal cortex and occipital cortex) did not modulate the binding effect. Given that the pre-SMA is a key structure for preparing and initiating voluntary actions (Cunnington et al., 2003), these findings offer strong support for the prospective theory of SoA.

However, a null effect of stimulating brain areas outside the pre-SMA, namely the dorsal parietal cortex and occipital cortex (Zapparoli et al., 2020), does not constitute direct evidence against the reconstructive theory of agency. In fact, behavioral studies offered strong support that action planning is not necessary for agency, given that it can be elicited without voluntary action (Borhani et al., 2017; Buehner, 2012; Kong et al., 2017; Poonian et al., 2015; Suzuki et al., 2019). The aforementioned intentional binding effect can be generated by merely observing another human's or even a machine's causal action (i.e., a key press) while observing a non-causal event (i.e., an LED flash) could not (Buehner, 2012). These observation-elicited binding effects thus support the reconstructive theory, which conceptualizes agency as a consequence of post hoc inference after

movements (Buehner & Humphreys, 2009; Desantis et al., 2011; Wegner, 2003; Wegner & Wheatley, 1999). Given this behavioral evidence, some researchers propose that both predictive and retrospective processes contribute to the manifestation of agency (Moore & Obhi, 2012). However, the neural evidence supporting the reconstructive theory of SoA is currently lacking.

Strong support would be that the generation of agency should arise in the absence of the involvement of brain regions of action planning but instead rely only on those directly related to referential sensemaking. However, previous neuroimaging studies typically required human subjects to act voluntarily, making action planning inevitable. In order to modulate the magnitude of agency, researchers either manipulated the authorship of the action, e.g., externally moving people's effector to generate passive "actions" (e.g., Balslev et al., 2006; Kuhn et al., 2013; Straube et al., 2017; Tsakiris et al., 2010; Uhlmann et al., 2020; van Kemenade et al., 2017, 2019; Zapparoli et al., 2020), or perturbed the sensory feedback of the movement or its outcome by implementing temporal and spatial discrepancies (e.g., Balslev et al., 2006; Chambon et al., 2013; David et al., 2007; de Bezenac et al., 2016; Di Plinio et al., 2020; Farrer et al., 2003, 2008; Farrer & Frith, 2002; Kikuchi et al., 2019; Kuhn et al., 2013; Leube, Knoblich, Erb, Grodd, et al., 2003; Leube, Knoblich, Erb, & Kircher, 2003; Matsuzawa et al., 2005; Nahab et al., 2011; Ohata et al., 2020; Sasaki et al., 2018; Schnell et al., 2007; Spengler et al., 2009; Uhlmann et al., 2020; van Kemenade et al., 2019; Yomogida et al., 2010; Zapparoli et al., 2020). Such contrasts between voluntary action and perturbed action revealed neural correlates of agency in extensive cortical areas such as frontal, parietal, temporal, and insula cortices and subcortical regions such as the cerebellum and striatum (Charalampaki et al., 2022; Haggard, 2017; Seghezzi et al., 2019). Among them, a frontoparietal network is highlighted, though disentangling their respective role in action planning and action feedback processing is challenging (Ohata et al., 2020).

Here we used virtual reality (VR) to modulate people's agency *without* requiring them to act voluntarily and examined whether its neural correlates were specifically tied to the inferential processing of action feedback rather than to action planning. Our recent behavioral study showed that after controlling an avatar in a first-person perspective in VR for a brief period, people increased temporal binding when passively observing an avatar's "action" (Kong et al., 2017). This "embodiment" effect was thus caused by the prior experience of controlling the avatar since the temporal binding was unchanged for people who experienced the identical VR environment but without controlling the avatar. Hence, such a VR setting would allow us to modulate the implicit agency and reveal its related neural changes when no voluntary action and motor planning were engaged. While previous studies used temporal binding as the behavioral phenotype for demonstrating the neural underpinnings of the prospective process for agency (Zapparoli et al., 2020), we would modulate the same binding effect to reveal its neural underpinnings of the reconstructive process. We hypothesized that if the change in binding involved action planning processes, we should find its neural correlates in sensorimotor regions, especially those prefrontal and frontal areas (e.g., pre-SMA and SMA proper) implicated in motor intention and planning (Seghezzi et al., 2019; Sperduti et al., 2011). Alternatively, if post-movement inferential processes contributed heavily to the binding effect, we should observe its neural correlates in the regions outside the frontal motor areas. The targeted areas included posterior parietal areas that had been attributed to causal inference and action awareness (Haggard, 2017; Renes et al., 2015; Wende et al., 2013). In particular, inferior parietal regions deserved special attention since direct stimulation of these regions induced subjective experiences of intending to move or even increased (false) reports of movements that were not objectively measured (Desmurget et al., 2009).

Another venue of our study was that our VR manipulation enabled us to examine the neural basis of agency over a virtual body. VR experience could change people's self-consciousness (Banakou & Slater, 2014; Slater et al., 2009), but previous neuroimaging studies focused on the sense of bodily ownership (Bach et al., 2012; Bekrater-Bodmann et al., 2014; Pamplona et al., 2022) and self-localization (Ionta et al., 2011; Lenggenhager et al., 2011), leaving sense of agency, the integral component of self-consciousness, largely uninvestigated. Our study thus constitutes an initial endeavor to explore the neural correlates of embodying an avatar via the control of its action.

Methods

Participants

Our study recruited 48 college students as paid volunteers. Half of the participants were randomly assigned to the experimental group and half to the control group. Both groups were exposed to a VR environment, but only the experimental group viewed an avatar hand. Four participants, inattentive in a task during fMRI scanning (rate of successful judgment < 80%, details in Procedures), were excluded from data analyses, leaving 23 participants in the experimental group (mean age = 23.69 ± 2.37 years, 12 females) and 21 participants in the control group (mean age = 21.99 ± 2.75 years, 13 females). Power analysis based on the reported effect size in our previous study with a similar design (Cohen's f = .5 for the interaction effect in a two-way mixed ANOVA; Kong et al., 2017) indicated that a sample size of n = 14 per group would lead to a power of 0.9 with an α level of 0.05 (G*Power 3.1; Faul et al., 2007). The two groups were matched in age (t_{42} = .639, p = .526) and gender (χ_1^2 = .424, p = .515). All participants were right-handed, had normal or corrected-to-normal vision, and reported no neurological diagnoses. The experiment was conducted according to the principles of the Declaration of Helsinki and was approved by the Ethical Committee of the School of Psychological and Cognitive Sciences at Peking University.

Procedures

Experimental procedure overview

Each participant went through three consecutive phases of the experiment, i.e., a pre-test, VR exposure, and a post-test. In the pre-test and post-test, participants performed the classical temporal binding task and a hand laterality judgment task (Ferri et al., 2012) in a sequel in the MRI scanner (T1 functional scan was acquired before the pre-test); each test lasted about 35 minutes. The laterality task was designed to study questions unrelated to the purpose of the current investigation.

Four participants showing inattentive performance in the laterality task (rate of successful judgment < 80%) were excluded from the current data analyses out of prudence. After the pre-test, the participants walked into the waiting room next to the scanning room to receive VR exposure for about 30 minutes. During VR exposure, participants wore a head-mounted display (HMD, HTC Vive Pro) and a motion-tracking glove (Noitom Hi5 VR Glove) to perform four gamified motor tasks by controlling an avatar hand with their right hand (see below). The VR exposure aimed to enable participants to embody the avatar hand by controlling it for goal-directed movements. The experimental group could view the avatar hand, whose motion spatially and temporally matched with that of their actual right hand. The control group was never given a chance to see the avatar hand and performed the same motor tasks. After finishing the VR exposure, the participants removed the HMD and the motion-tracking glove, and walked with their eyes closed to the scanning room under the guidance of the experimenter. They were instructed not to open their eyes until they were properly positioned in the scanner to get ready for the post-test. This procedure was employed to minimize the visual experience of the real settings and to preserve the effect of VR exposure. After the post-test, participants re-entered the waiting room and were asked to evaluate their subjective sense of embodiment (explicit rating of agency and ownership over the avatar hand) during the post-test by means of questionnaires.

Temporal binding task

The temporal binding task performed was a replicate of the same task in our previous study, though it was performed with the HMD earlier (Kong et al., 2017). In brief, participants judged the timing of an auditory stimulus with the aid of a Libet clock projected in the MRI scanner in keeping with the temporal binding task outside of the VR (Haggard et al., 2002).

The task involved temporal judgments of tones in a series of trials. Two types of trials, baseline, and operant trials, were presented in blocks. For each operant trial (top panel in Figure 1a), the Libet clock started to rotate clockwise from a random location. After a random interval of 2560 - 5120 ms, the right avatar hand pressed the white button. A tone would be presented 250 ms later (100 ms in duration). Note that the participant was required to refrain from any movement during the stimulus presentation. The clock hand kept rotating after the tone for a random duration of 1280 - 2560 ms, then reset its position to 12 o'clock. Participants were required to report the location of the clock hand when the tone was perceived. They reported by pressing the left and right keys using the left middle and index finger to move the clock hand clockwise and counterclockwise, respectively. They then confirmed their estimated position by pressing the key with the right index finger when the clock hand arrived at the desired location. The report was self-paced. The baseline trial (bottom panel in Figure 1a) was identical to the operant trial, except that the tone would be presented without the button press, i.e., the avatar's hand remained static all the time. The timing of the tone followed the same random uniform distribution as for the operant trial.

Participants familiarized themselves with the temporal binding task before the formal experiment by conducting four operant trials and four baseline trials outside the scanner. Both the pre-test and the post-test comprised of two operant runs and two baseline runs whose orders were randomized and counterbalanced between participants. Due to technical failure, two control participants underwent one operant run and three baseline runs in one of the two phases. Each run contained 20 trials.

All the visual stimuli, including the Libet clock, a pair of avatar hands, and a white button under the index finger of the right avatar hand, were back-projected onto a translucent screen located

inside the scanner (resolution: 1024×768; refresh rate: 60Hz; viewing distance about 90 cm; Figure 1a). The avatar hand was the one that the participants visually controlled during the VR exposure outside the scanner but fitted for rendering in the two-dimension display in the scanner. The clock had a 10-pixel-long hand, which rotated with a period of 2560 ms per circle. The clock face (radius = 110 pixels) was marked with conventional “1 min” intervals. The participants wore a pair of headphones to receive the auditory stimuli and to reduce noise. Before the formal data collection, participants familiarized themselves in the scanner with the temporal binding task. We adjusted the volume of the auditory stimulus on the individual basis to ensure that they could hear the stimuli clearly and comfortably in the scanner. The rendering was controlled by a customized program coded using Psychtoolbox (3.0.16) implemented in MATLAB (R2016b, MathWorks).

VR exposure

During VR exposure that was sandwiched between the pre-test and the post-test, participants wore the HMD (HTC Vive Pro) with a dual-display giving a 110° field of view and a resolution of 1080 × 1200 per eye. In the VR scene, they could see a virtual desk located at approximately the same location as the physical desk at which they sat. The motion tracking glove sampled their right hand motion at a sampling rate of 100 Hz to control a right avatar hand in the VR environment. The motion of the avatar hand was always temporally and spatially congruent with the real right hand. For participants from the experimental group, the avatar hand resembled their own right hand in size and skin tone, which were adjusted before the experiment. For control participants, no avatar hand was shown, and all the VR tasks were performed with the "invisible" hand. The games were designed to provide a sensorimotor experience of controlling the avatar's hand to interact with the virtual environment (Figure 1b), including the gesture-imitation task, the bubble-poking task, the cube-picking task, and the pad-tapping task (customized by using Unity3D 2019.3.14f1).

In the gesture-imitation task, a left-hand avatar was displayed in front of the participants at their eye height. This avatar hand would present a hand gesture randomly selected from 20 possible gestures, e.g., a closed fist, a thumb-up, and a V-sign. The participant was required to replicate the same gesture with their right avatar hand. Once the gesture was reproduced, the next target gesture would appear until all 40 trials (two repetitions for each different gesture) were successfully finished. In the bubble-poking task, a transparent bubble would appear at a random location on the desk, moving at a random speed and direction. Participants were asked to poke the bubble with the index finger of the right avatar hand before it fell off the desk. After a bubble was poked or dropped off the desk, the next bubble would appear. This task lasted until 60 target bubbles were shown and poked. In the cube-picking task, forty cubes with different sizes, colors (red or blue), and labeled numbers (1 or 2) were placed randomly on the virtual desk. Participants were required to control the avatar hand to grab specific cubes and drop them on the floor. Participants performed this task three times with different verbal instructions each time, including “pick all the red cubes and drop them on the floor," “pick all the cubes labeled as number 1 and drop them on the floor”, and “pick all the red cubes labeled as number 1 and drop them on the floor”. In the pad-tapping task, a multi-digit number would appear before disappearing after 2-4 s. Participants were required to memorize the number and then replicate the number by tapping on a virtual number pad. Forty numbers, ranging from 3-digit to 10-digit, were presented in random order for each participant. Each of the four VR tasks lasted for about 6-7 minutes, resulting in a total VR exposure time of about half an hour.

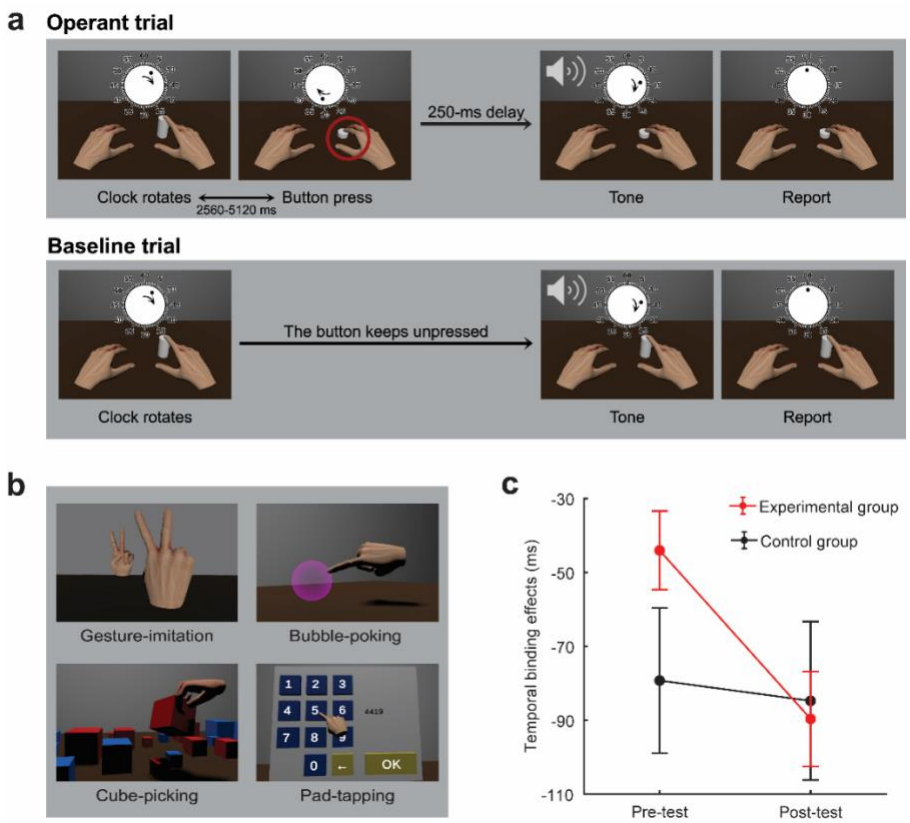

Figure 1. Illustrations of behavioral tasks and the observed temporal binding effects. (a) Graphical illustration of an operant trial (top panel) and a baseline trial (bottom panel) in the temporal binding task. (b) Scenes of motor tasks in virtual reality (VR). Participants performed four VR motor tasks: the gesture-imitation task: bending right-hand fingers to match a target gesture, shown by the distant avatar hand; the bubble-poking task: poking the bubble with the right index finger; the cube-picking task: picking target cubes that are specified by color or shape in instructions; the pad-tapping task: memorizing a multi-digit number before it disappears, and then recalling it by tapping on a keypad. The experimental group viewed the avatar hand (shown here), but the control group did not. (c) Temporal binding effects that were quantified by perceptual shifts evoked by the avatar's movement. The average binding effects in the pre-test and post-test are shown for the two groups separately.

Questionnaire

We assessed the subjective feeling of embodiment by means of questionnaires after the post-test (Table 1). The questions were identical to the ones used in our previous study on VR embodiment (Kong et al., 2017). The items were designed to assess explicit agency and ownership or to control for possible response biases by using the reversed control items. Notably, the agency items were related to the explicit sense of control over the avatar hand movements during the operant condition in the post-test, thus serving as a measure of explicit agency over the avatar hand in the post-test. As our previous behavioral study has shown that explicit agency did not differ between groups, our fMRI analysis focused on the neural correlates of temporal binding.

Table 1. Questionnaire on embodiment evaluation. All questions were rated on a Likert scale from 1 (strongly disagree) to 7 (strongly agree).

Category	**Questions: During the operant condition in the post-test,**	**Order of questions**
Explicit ownership	I felt like I was looking at my own hand.	1
	I felt like the avatar was part of my body.	6
	It seemed as if my hand was pressing when I saw the avatar's movement.	5
Explicit agency	It seemed as if the avatar moved, obeying my will.	4
	It seemed as if the avatar pressed the button instead of me.	3
	If I moved my finger, I felt that the virtual finger would move in the same way.	7
Control items	I felt as if the movement of the avatar had no relationship with me.	8
	It seemed as if the avatar had a will on its own.	2

Behavioral analysis

To examine whether VR exposure could enhance participants' implicit agency associated with an avatar hand, we compared the temporal binding effect by using a two-way mixed-design ANOVA with Phase as the within-group factor (pre- versus post-test) and Group as the between-group factor (experimental versus control group). Each trial yielded a perceptual error of temporal judgment,

quantified as the difference between the reported and the actual onset of the tone. The temporal binding effect was operationally defined as the difference in perceptual error, i.e., a perceptual shift, between the operant condition and the baseline condition. A negative perceptual shift indicated the temporal binding effect.

For the subjective measures of embodiment from questionnaires, we calculated the median scores for explicit agency, explicit ownership, and control-items separately. Between-group and within-group comparisons were performed by the Wilcoxon signed ranks test. Spearman correlation was used to test the association between the temporal binding effect and the subjective rating of embodiment.

MRI acquisition

All functional and structural MRI data were acquired on a Siemens 3T Prisma scanner with a 64-channel head-neck coil at the Center for MRI Research, Peking University. High-resolution functional images were acquired using a multi-band echo-planar sequence (62 axial slices, repetition time (TR) = 2000 ms, echo time (TE) = 30 ms, flip angle (FA) = 90°, field of view (FOV) = 224 × 224 mm^2, matrix size = 112 × 112, slice thickness = 2.0 mm, voxel size = 2 × 2 × 2 mm^3, multi-band factor = 2). High-resolution 3D T1-weighted anatomical images were acquired using the magnetization-prepared rapid-acquisition gradient-echo sequence (192 sagittal slices, TR = 2530 ms, TE = 2.98 ms, inversion time = 1100 ms, FA = 7°, FOV = 224 × 256 mm^2, matrix size = 224 × 256, interpolated to 448 × 512, slice thickness = 1.0 mm, voxel size = 0.5 × 0.5 × 1 mm^3).

MRI analysis

fMRI data were preprocessed using Statistical Parametric Mapping (SPM12) software (Wellcome Center for Human Neuroimaging, London, UK, http://www.fil.ion.ucl.ac.uk/spm12/) and analyzed using the toolbox for Data Processing & Analysis for Brain Imaging (DPABI, Version 6.2, http://rfmri.org/DPABI; Yan et al., 2016) implemented in MATLAB (R2019b, MathWorks). All of the surface brain maps were visualized with the BrainNet Viewer (Version 1.7, http://www.nitrc.org/projects/bnv/; Xia et al., 2013). Other statistical tests were conducted in R (Version 4.1.1, https://www.r-project.org/).

Data preprocessing

Functional images were preprocessed using SPM12. For each individual participant, the first five volumes of each functional run were discarded. The remaining images were corrected for slice timing and head motion and spatially normalized to Montreal Neurological Institute (MNI) space via unified segmentation (resampling into $2 \times 2 \times 2$ mm^3 voxel size). No participant revealed head motion larger than 3 mm/3°. The resulting images were spatially smoothed using a 6-mm full-width half-maximum (FWHM) Gaussian kernel for univariate contrast analyses.

Univariate contrast analyses

We carried out a whole-brain univariate analysis to test whether exposure with or without an avatar hand in VR would lead to differential brain activities for the temporal binding task. At the first level, preprocessed functional images were modeled in a general linear model (GLM) corresponding to the onset of the beep time in each trial, with four regressors for the condition (operant or baseline) and the phase (pre- or post-test) of different runs (the condition was consistent in a single run), convolved with the canonical hemodynamic response function (HRF). The GLM

also included six additional predictors of no interest representing head motion parameters in each run. The high-pass filter was set at 128 s. After model estimation, we first obtained whole-brain contrast images of operant condition versus baseline in the pre-test and post-test, respectively, for each participant. One-sample *t*-tests of these contrast images represented the general neural responses of the temporal binding effects (operant error - baseline error). The between-phase (post-test versus pre-test and pre-test versus post-test) effects of such neural responses were analyzed by paired-sample *t*-tests. Besides, the VR exposure-related neural activities were quantified by subtracting the whole-brain contrast images of the pre-test (operant versus baseline) from the post-test (operant versus baseline) for each participant. VR exposure-related neural signals were entered into the second-level analyses, testing the between-group effect by a two-sample *t*-test. All results reported survived at a voxel-wise threshold at $p < .005$ and a cluster-level FWE correction at $p < .05$.

ROI analyses

We also performed regions of interest (ROI) analysis to examine the possible involvement of the traditional agency-related network and the action-observation network in our temporal binding task. The traditional agency network contained seven ROIs as reported by a recent meta-analysis (Seghezzi et al., 2019). We used the reported peaks to form 8-mm-radium spherical ROIs (Figure 3, Figure S1a, and Table S1). For completeness, these seven ROIs covered clusters associated with both self-agency (agency attributed to self; their Table 1) and external-agency (agency attributed to others; their Table 3) as both are related to agency processing. The action observation-related network was selected to verify that the VR-related neural activities were not simply caused by action observation. We chose fourteen ROIs reported in association with the observation of general hand actions from a meta-analysis (Caspers et al., 2010), and used the reported peaks to form 8-

mm-radium spherical ROIs (Figure 3, Figure S1b and Table S2). For the same purpose, we also picked the regions particularly associated with the action observation of avatar hands since action observation in our present temporal binding task was elicited by the key press in the operant trials. This avatar hand action observation network was defined as regions showing significantly stronger activation in the operant condition relative to the baseline in the pre-test for all participants (voxel-wise uncorrected $p < .005$ and cluster size > 30). Results showed that the bilateral occipitotemporal cortices (LOTC; Figure 3, Figure S1c, and Table S2) were specifically related to the action observation of the avatar hand (details in Results).

For a preview, our whole-brain univariate analysis revealed a cluster associated with VR-induced temporal binding changes, while the ROI analyses generally returned null results. To show the effect size of our identified cluster's activation, we conducted a two-way mixed-design ANOVA, with the Brain network as the within-group factor and the Group as the between-group factor. The averaged beta values of each brain network (the traditional agency network, the action observation network, the avatar observation network, and the newly identified cluster for the temporal binding effect) were extracted as a whole mask in the two-way mixed ANOVA. The interaction effect of this ANOVA would highlight the distinguishing power of the identified cluster to account for the group effect.

Correlation analyses between fMRI and behavioral measures

To probe the link between brain activity and behavior, we correlated the behavioral measures with their corresponding neural activation. For temporal binding, the behavioral measure is the change in binding between the pre-test and the post-test and the corresponding neural signal is the difference in activation between the two scans. Since explicit reports of agency were only obtained

after the post-test, the corresponding neural signal is the activation in the post-test scan. We first used Spearman correlation to test whether the neural signals in the clusters showing significance in the contrast correlated with behavioral measures. To examine whether the correlation could be explained away by group difference, we regressed neural signals on behavioral measurements with a group label as the covariate. We then used whole-brain correlation analyses with a voxel-wise threshold at $p < .005$ and a cluster-level FWE correction at $p < .05$. For completeness, Spearman correlations between the activity in the traditional agency network and our measurements of agency, i.e., temporal binding and explicit ratings, were also conducted.

Results

Behavioral results: The temporal binding effects

Before the VR exposure (pre-test phase), both groups of participants showed a moderate level of temporal binding effect (Figure 1c). In the baseline condition, when the avatar hand remained stationary, the temporal judgment error amounted to an average of -26.45 ms (SE = 14.58 ms, same below) and -38.05 (16.25) ms for the experimental and control groups, respectively. In the operant condition with the avatar hand suddenly pressing the button before the tone, the judgment error was -70.49 (17.06) ms and -117.30 (27.44) ms for the experimental and control groups, respectively. Thus, in the pre-test, both groups showed a negative perceptual shift in timing judgment of the tone, an indicator of temporal binding.

Crucially, after the VR exposure in the post-test, the experimental group increased their binding effects by 45.59 ms (10.57) with the baseline and operant error of -7.81 (15.09) ms and -97.45 (19.66) ms, respectively. The control group's binding effects remained largely unchanged (differed from the pre-test by -5.49 ± 13.32 ms). This group difference was confirmed by a mixed ANOVA

(within-group factor: Phase; between-group factor: Group): both the interaction ($F_{1, 42}$ = 5.71, p = .021, Cohen's f = .37) and the main effect of Phase ($F_{1, 42}$ = 9.26, p = .004, Cohen's f = .47) reached significance, while the main effect of Group was not significant ($F_{1, 42}$ = .49, p = .487, Cohen's f = .11). Post-hoc comparisons indicated that the experimental group showed enhanced binding effects in the post-test (-89.64 ± 16.92 ms) compared to the pre-test (-44.04 ± 15.05 ms; t_{42} = 3.93, Bonferroni-corrected p < .001, Cohen's d = .83), while the same effect was absent in the control group (t_{42} = .45, Bonferroni-corrected p = .654, Cohen's d = .10). These behavioral results and their effect size were similar to those in our previous study (Kong et al., 2017), indicating that seeing and controlling an avatar hand increases the temporal binding effect.

Note that the two groups tended to differ in the binding effect at the pre-test, though statistically insignificant (t_{42} = 1.62, p = .113; Figure 1c), and two participants of the control group only underwent one operant run in a test phase (one in the pre-test and another in the post-test). To test the group differences robustly, we excluded these two participants and then randomly selected 18 participants from each group. We found that 99.83% of the possible pairs of subgroups were matched in the pre-test (t_{34} < 2.03, p < .05), and 70.27% of matched subsets still revealed a significant interaction of Phase and Group ($F_{1, 34}$ < 5.50, p < .05), which suggested that the enhanced binding effect found in the experimental group was stable. A pair of exemplary subgroups, which were matched in pre-test temporal binding, were selected to validate the imaging results of the entire group of participants. Main imaging results and conclusions from this subset of participants remained the same as those from all participants (details in Supplementary Material).

Behavioral results: Explicit rating of agency and ownership

Besides using temporal binding as the implicit measure of agency, we also asked participants to explicitly report their sense of embodiment for the avatar hand in the post-test. We did not find a significant group difference in explicit agency and ownership (z = -.13, p = .896 for explicit agency; z = -.10, p = .925 for explicit ownership). Across two groups, the average scores of the control items were significantly higher than those of the agency questions (z = -3.43, p = .001), indicating that participants correctly followed the instructions to fill in the questionnaire.

As expected, explicit agency and ownership were positively correlated (ρ_{42} = .60, p < .001). However, neither of them correlated with the temporal binding effects in the post-test or the binding changes between the pre- and post-test (binding effects: ρ_{42} = -.19, p = .225 and ρ_{42} = -.07, p = .676 for explicit agency and ownership in the post-test, respectively; changes in binding effects: ρ_{42} = -.03, p = .827 and ρ_{42} = .04, p = .783 for explicit agency and ownership, respectively).

Imaging results: The temporal binding effects revealed by Operant vs. Baseline condition contrast

To investigate the neural responses of the binding effects (operant error - baseline error) in the temporal binding task, we first generated whole-brain contrast images of the operant condition versus baseline, separately for different phases (pre-test and post-test) and groups (experimental group and control group). Note that the difference captured by these contrast images is that the avatar hand performed a quick keypress movement in the operant condition while it remained static in the baseline condition. All the contrasts except the experimental group in the pre-test showed significant activation in bilateral occipitotemporal cortices (LOTC; Figure 2a) with a voxel-wise threshold at p < .005 and a cluster-level FWE correction at p < .05. The experimental group in the pre-test showed a trend in this region but failed to reach the significance level.

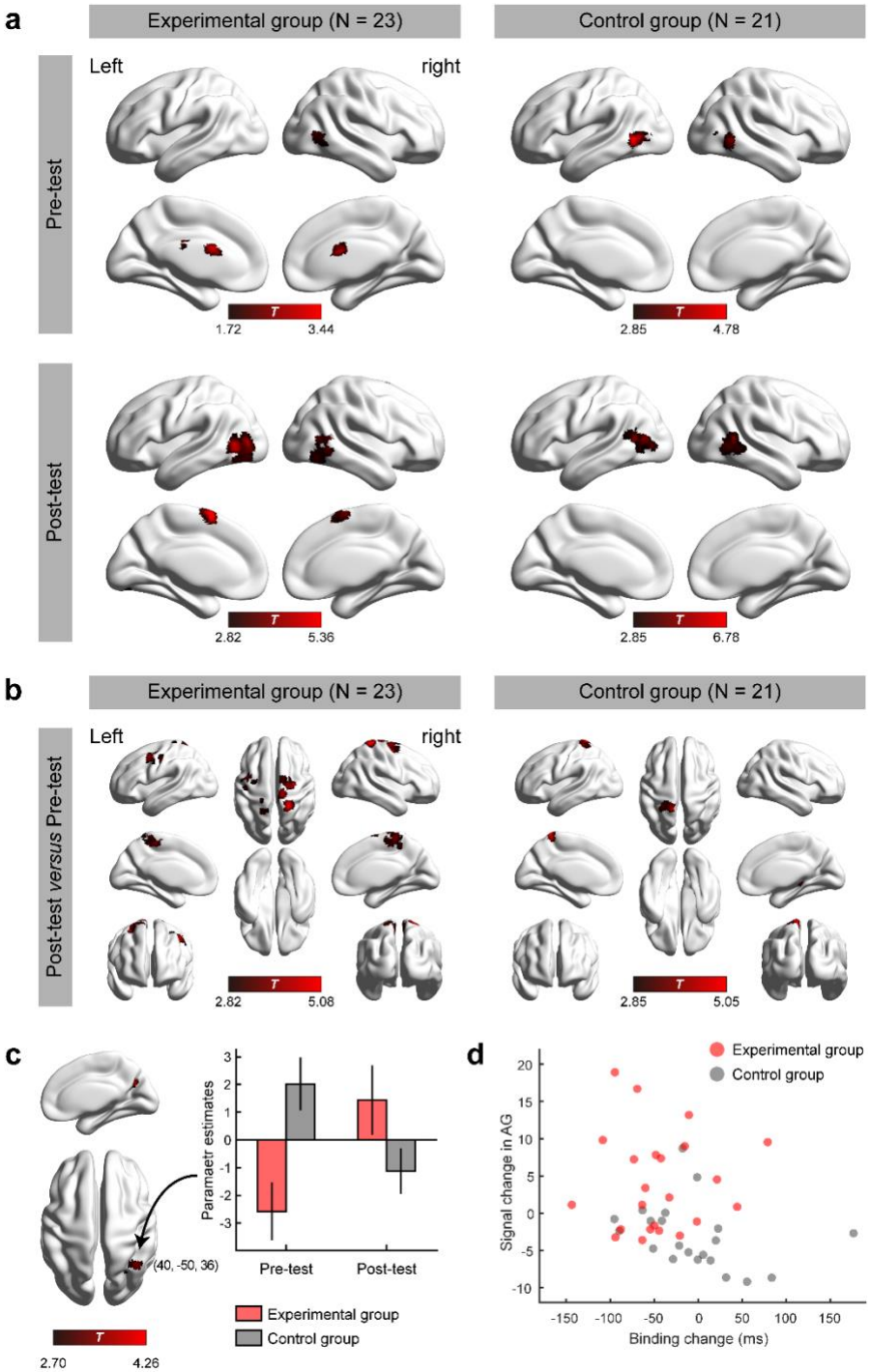

Figure 2. Results of the univariate analyses. (a) Whole-brain contrast images of operant versus baseline were plotted separately for different phases (pre-test and post-test) and different groups (experimental group and control group). A lenient threshold was applied at voxel-wise

uncorrected $p < .05$ and cluster size > 30 for the experimental group in the pre-test, while the threshold of the other three conditions was applied at voxel-wise $p < .005$ and cluster-level FWE correction at $p < .05$; (b) Whole-brain comparison for the neural response of the binding effect (operant - baseline) before and after VR exposure (post-test versus pre-test), plotted separately for the experimental group and control group. The experimental group revealed three clusters with greater activation in the post-test relative to the pre-test (voxel-wise threshold at $p < .005$, cluster-level FWE correction at $p < .05$). The control group displayed no significant activation at the same threshold, and its activation map was shown with a more lenient threshold at voxel-wise uncorrected $p < .005$ and cluster size > 30; (c) The whole-brain between-group comparison of the VR exposure effect that is depicted in (b). Greater activation with the experimental group relative to the control group was observed at the right angular gyrus (MNI coordinate: 40, -50, 36; cluster size: 290 voxels; voxel-wise threshold at $p < .005$, cluster-level FWE correction at $p < .05$). The right bar graph depicts the magnitude of neural activation of the whole cluster, as represented by the average parameter estimates across participants (operant versus baseline); (d) Association between the VR exposure-related neural signals and the temporal binding changes.

Imaging results: The VR binding effects (Post- vs. Pre-test) revealed by Experimental vs. Control group contrast

For the experimental group who behaviorally showed enhanced binding effects in the post-test, we observed increased activation in the post-test (versus pre-test) in three clusters (Table 2 and Figure 2b; voxel-wise threshold at $p < .005$ and cluster-level FWE correction at $p < .05$), located in the left precentral gyrus, right paracentral lobule and right superior frontal gyrus. No significant activation in the pre-test (versus post-test) was found in the experimental group. For the control group, no region displayed a significant effect between the two phases (either post-test versus pre-

test or pre-test versus post-test) applying the same threshold, which echoed their stable binding effects throughout the experiment.

Behaviorally, the VR exposure-induced binding changes were quantified by subtracting the binding effects (operant error - baseline error) of the pre-test from that of the post-test. Similarly, the VR exposure-related neural activities were analyzed by obtaining whole-brain contrast images between the pre-test (operant versus baseline) and the post-test (operant versus baseline). We then used these contrast images to make comparisons between the experimental and control groups who performed the same temporal binding task in the scanner and differed only in whether they had visually controlled the avatar hand during the VR exposure phase outside the scanner.

Critically, compared to the control group, the experimental group showed increased activation in a single cluster centered around the right angular gyrus extending into the right inferior parietal lobule and the right precuneus (Table 2 and Figure 2c, MNI coordinate of peak: 40, -50, 36; cluster size = 290 voxels, $t_{42} = 4.26$). Specifically, for the experimental group, the activation related to the operant versus baseline effects within this cluster increased from -2.58 (1.04) in the pre-test to 1.44 (1.25) in the post-test, while for the control group, such activation decreased from 2.02 (.95) in the pre-test to -1.12 (.83) in the post-test (Figure 2c). No significantly greater activation was detected in the control group when compared with the experimental group at the same threshold.

Table 2. Summary of the results of the univariate analyses.

Brain region		**MNI coordinates (mm)**			**Significance**		
Area	**Hemisphere**	***x***	***y***	***z***	***t***	***r***	**Cluster size (vx)**
VR binding effects (Post- *vs.* Pre-test within Exp. group)							
Precentral gyrus	Left	-42	4	46	4.33	—	256
Paracentral lobule	Right	20	-40	68	5.08	—	628
Superior frontal gyrus	Right	26	-12	56	4.97	—	237
VR binding effects (between Exp. and Control group)							
Angular gyrus	Right	40	-50	36	4.26	—	290

Note: A voxel-wise threshold at $p < .005$ and a cluster-level FWE correction at $p < .05$ was set for the univariate analyses.

Imaging results: ROI analysis for the VR binding effect

Unlike previous imaging studies using voluntary actions to study sense of agency, the temporal binding effect in our paradigm was induced by mere observation of "embodied" action. Our univariate analysis only revealed a significant between-group effect in the cluster centered around the right angular gyrus (Figure 2c). We further explored whether ROIs that have been previously linked to sense of agency induced by self-action (Seghezzi et al., 2019) and to the observation of the general hand actions (Caspers et al., 2010) could show the same between-group effect. The particular regions associated with action observation of the avatar hand (operant versus baseline) were also examined (Figure 3, details in Methods, Figure S1, Table S1, and Table S2).

Results of the between-group comparison in every single ROI are shown in Supplementary Material. Briefly, none of these ROIs produced the same between-group effect except the right AG from the meta-analysis of agency studies (Seghezzi et al., 2019), which was marginally

significant (p = 0.051, BF_{10} = 1.46). Note the peak voxel of the right AG from this meta-analysis was 13.2 mm from the peak of our cluster.

These brain networks were compared to the cluster centered around the right angular gyrus by submitting the neural changes of each brain network into a mixed ANOVA (within-group factor: Brain network; between-group factor: Group; Figure 3). Both the interaction ($F_{1.6, 65.6}$ = 9.93, p < .001, Cohen's f = .49) and the main effect of Group ($F_{1, 42}$ = 4.41, p = .042, Cohen's f = .32) reached significance, while the main effect of Brain network was not significant ($F_{1.6, 65.6}$ = .78, p = .435, Cohen's f = .14). Simple main effects indicated that the angular gyrus cluster showed more neural changes in the experimental group (4.03 ± 1.18) than in the control group (-3.14 ± 1.24; t_{42} = 4.19, Bonferroni-corrected p < .001, Cohen's d = 1.64), while the other three networks failed to show any group difference (traditional agency network: t_{42} = 1.24, Bonferroni-corrected p = .222, Cohen's d = .38; general hand action observation network: t_{42} = .43, Bonferroni-corrected p = .673, Cohen's d = .13; avatar hand action observation network: t_{42} = .99, Bonferroni-corrected p = .328, Cohen's d = .41). Thus, the enhanced binding effect was specifically tied to the activation in the right AG, not in the traditional agency network or the network associated with hand action observation.

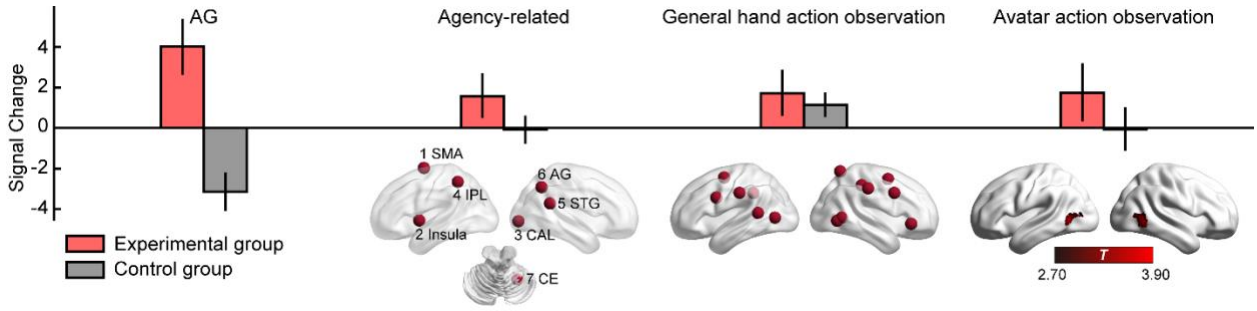

Figure 3. Group-based ROI results. Top panel: experimental (red) and control (gray) group information of neural signal change in the right angular gyrus (AG) cluster (Figure 2c) and different brain networks. Bottom panel: traditional agency-related network and the general hand action observation network defined by two recent meta-analyses (Caspers et al., 2010; Seghezzi et al., 2019), and the avatar hand action observation network defined as regions showing stronger activation in the operant condition relative to the baseline in the pre-test for all participants. Details in Table S1 (SMA = supplementary motor area, Insula = posterior insula, CAL = calcarine scissure, IPL = inferior parietal lobule, AG = angular gyrus, STG = superior temporal gyrus, CE = cerebellum).

Associations between neural activity and measurements of agency

To probe the link between neural activity and the agency measures, i.e., temporal binding and explicit ratings, we extracted neural signals from the cluster centered in the right AG, which showed a significant between-group effect (Table 2 and Figure 2c). We found the temporal binding changes positively correlated with activity changes in this cluster before and after the VR exposure (ρ_{42} = -.34, p = .023; Figure 2d). Thus, the larger the increases of the binding effects, the bigger the activation changes in the right AG. The multiple regression analysis, using the group label as an additional predictor, revealed that Group was significant (β = .51, p < .001) but the binding change was no longer significant (β = -.11, p = .437). This result thus suggested that the correlation between the VR exposure-related neural activities and the temporal binding changes was largely driven by the group difference.

Though not the focus of the current investigation, the neural correlates for the subjective rating of agency were also explored. As the rating was only collected after the post-test, we examined the

neural signal (operant versus baseline) in the post-test. Interestingly, the cluster centered on the right AG also exhibited significant activations associated with the explicit agency (ρ_{42} = .35, p = .020). Thus, the higher the explicit sense of control over the avatar hand movement in the post-test, the higher activation in the right AG. The multiple regression analysis, using the group label as an additional predictor, revealed that group was not significant (β = .23, p = .114) but the explicit agency was significant (β = .33, p = .027). This result suggests that the correlation between the VR exposure-related neural activity and the explicit agency was not driven by group differences.

However, the whole-brain correlation analyses did not return any significant results, neither for the binding changes nor the explicit ratings of agency. No region displayed significant correlation (voxel-wise threshold at p < .005 and cluster-level FWE correction at p < .05). For completeness, we also examined whether the activity in the traditional agency network was correlated with our measurements of agency. None of the seven ROIs showed a significant correlation with the binding changes (all ps > .1, Table S1). However, the left insula showed a significant correlation (ρ_{42} = .36, Bonferroni-corrected p = .018) and the left SMA showed a marginally significant correlation (ρ_{42} = .28, Bonferroni-corrected p = .070) with the explicit rating after the post-test, consistent with their role in explicit agency attribution as implicated by previous studies. The other five ROIs did not reveal a significant correlation (all ps > .1, Table S1).

Discussion

We examined the brain activation pattern supporting the changes in sense of agency that were elicited by passive observation of "embodied" avatar actions using the classical temporal binding paradigm before and after VR exposure. Behaviorally, we replicated our previous findings that controlling an avatar for a short period of time leads to increased binding upon seeing the avatar's hand action while the actual hand stays immobile (Kong et al., 2017). Extending the previous finding, here we show that this VR embodiment effect persists after exiting the VR setting and when the same avatar hand is then depicted in a two-dimensional display in the fMRI scanner. Our neuroimaging results identified increased activation in a cluster centered in the right angular gyrus, extending to the inferior parietal lobule and precuneus. This cluster is outside the regions for action planning but has been previously implicated in agency processing and, more importantly, inferential sensemaking. The change in binding is unlikely explained away by unspecific action observation as the traditional action observation network is not implicated (Figure 3 and S2, Table S2). As expected, the observed neural correlates for VR-induced SoA are distinct from neural correlates for other VR-induced self-consciousness changes, such as bodily ownership and self-localization. Our findings support the reconstructive theory of agency by showing that temporal binding, widely viewed as reflecting the prospective process of agency, is elicited without action planning, and its modulation is subserved by neural regions tied to inferential sensemaking of sensory events.

The change in SoA after VR exposure enabled us to reveal distinct neural underpinnings of agency without invoking action planning and prospective processing. Previous investigations have extensively examined the activation foci for SoA by contrasting condition pairs with presumably different levels of SoA, i.e., voluntary action versus rest, voluntary action versus stimulus-driven

action, actions with higher visuomotor congruency versus those with lower congruency (see a recent review, Seghezzi et al., 2019). Of note, all these contrasting analyses involve voluntary actions. Thus the identified extensive brain network for SoA includes sensorimotor areas, especially those related to intentionality and action planning (Seghezzi et al., 2019; Zapparoli et al., 2020). For example, the dorsolateral prefrontal cortex is associated with action selection (Khalighinejad et al., 2016), pre-SMA with intentionality (Yomogida et al., 2010; Zapparoli et al., 2018) and temporal binding in particular (Kuhn et al., 2013), and SMA proper with motor planning and initiation (Passingham & Lau, 2019). Consistent with the rationale of our task design, all these prefrontal and frontal areas, related to the generation of action before action feedback, returned a null effect in our data (Figure 3 and S1, Table S1). Interestingly, the peak of the right AG cluster we obtained with our VR manipulation had a sizable distance (13.2mm) from the peak estimated from the meta-analysis based on voluntary action-related SoA (Seghezzi et al., 2019).

Instead, the temporal binding change elicited by observing an embodied action is specifically associated with right AG (extended to IPL and precuneus), a cluster not directly tied to action planning and initiation. This finding is consistent with the reconstructive theory of agency, which emphasizes that the temporal binding is grounded by retrospective causality since right AG and IPL participate in both agency and causal processing in general. The right IPL has been one of the most frequently revealed neural correlates of sense of agency (Chambon et al., 2013, 2015; Farrer et al., 2003, 2008; Nahab et al., 2011; Schnell et al., 2007). Even anosognosia patients who often assert that they performed an action with their paralyzed, immobile limb typically have lesions in the right parietal lobule (Fotopoulou et al., 2008). More importantly, the right IPL is broadly involved in causal processing since the explicit judgment of both physical and social causality relates to neural activations in the right IPL, along with other areas (Renes et al., 2015; Wende et

al., 2013). Even seeing a causal event, such as object collision, elicits more activations in the right IPL than seeing a non-causal event, such as an object launching (Fugelsang et al., 2005).

The AG, similarly implicated by numerous agency studies, engages in diverse cognitive tasks that require inferential sensemaking. For agency, meta-analyses have shown that the TPJ, with the right AG included, is related to attributing agency to others (external agency, Sperduti et al., 2011) and to the reduction of self-agency (negative agency, Zito et al., 2020). A recent review also finds AG as a common node for encoding motor intention and SoA (Seghezzi et al., 2019). Beyond agency tasks, the AG has been reliably shown to engage in a wide range of tasks, including reasoning, semantic processing, word reading and comprehension, memory retrieval, attention and spatial cognition, default mode network, and social cognition. A well-received unified theory about the AG's function, based on the commonality of these tasks, highlights its role in sensemaking, i.e., giving meaning to external sensory information or internal thoughts (Seghier, 2013). For instance, the AG engages in the comprehension of speech and written languages (Obleser & Kotz, 2010; Xu et al., 2005), especially in solving referential ambiguity (Nieuwland et al., 2007). It also engages in inferring human intention in theory of mind tasks (Mason & Just, 2011). Given its rich anatomical connectivity to widely-distributed brain regions, the AG appears suitable for combining diverse information, linguistic and nonverbal (e.g., body movements), prior knowledge (experiences, context, and purpose) and new sensory information, to converge toward plausible accounts of the events. This sensemaking process can be implemented as an active optimization process that combines bottom-up information (i.e., sensory information) with top-down predictions (i.e., prior knowledge and purpose) to minimize surprise according to the free energy principle (Friston, 2010). Pertinent to our findings here, the AG is a central region for the inferential sensemaking process in various tasks, among which the agency-related task is an

important genre since agency sets the boundary between self and external environment (Seghier, 2013, 2022).

The involvement of AG and IPL in our observation-based temporal binding strongly suggests that retrospective processing of agency cues involves causal inference, in line with the reconstructive theory of agency (Buehner & Humphreys, 2009; Desantis et al., 2011; Tramacere, 2021; Wegner, 2003; Wegner & Wheatley, 1999). Our findings, of course, should not be taken as evidence against the importance of the prospective processing for agency during voluntary action. There exists extensive behavioral and neural evidence that both the prospective motoric process in action planning and retrospective process in outcome evaluation contribute to SoA, though their relative importance depends on available cues and task goals (Moore & Obhi, 2012; Synofzik et al., 2013). Our findings highlight that the brain can indeed invoke agency-related processing retrospectively without the involvement of action planning.

The validity of treating temporal binding as an implicit measurement of SoA has been challenged in recent years, given that our and other studies have shown that temporal binding can occur without voluntary action and intentionality (Buehner, 2012; Kong et al., 2017; Suzuki et al., 2019). However, temporal binding with voluntary actions changes according to agency manipulations, including the aforementioned experimental comparisons between active and passive movements and between congruent and incongruent action feedback. Thus, a parsimonious account of temporal binding posits that it results from top-down causal belief about the timing of sequential events, with or without voluntary action (Hoerl et al., 2020). The belief is about the causal relationship between a movement-related event, not necessarily an intentional action, and a subsequent outcome event. The causal belief is subject to influence from priming, instruction,

statistical contingency, and prior belief, which all have been shown to affect temporal binding (Aarts et al., 2005; Desantis et al., 2011; Ebert & Wegner, 2010; Moore et al., 2009; Wegner, 2003). The causal account thus views temporal binding as a general phenomenon in timing perception and casual belief, beyond a reflection of implicit agency that has been argued to embed in the motor system. This view resembles the reconstructive theory of agency with its emphasis on inferential processing for sensory events. In this light, the control experience during our VR exposure facilitates the formation of causal belief between avatar "action" and subsequent sensory event, possibly following the principles of ideomotor learning (Elsner & Hommel, 2001). The AG and IPL underlies the VR binding effects, thus might play a role in representing the causal belief. Furthermore, the fact that the activity in this cluster also positively correlates with the explicit ratings of agency over the avatar appears to suggest that it might also play a role in explicit causal attribution of avatar action. Though risking the curse of reverse inference from neural findings to cognitive processes, our findings support the causal account by showing that the neural substrate underlying our observed VR binding effect involves AG and IPL, important areas supporting the causal inference of sensory events.

Though a quantitative model of agency is currently lacking, various aspects of temporal binding have been successfully accounted for by probabilistic inference models based on Bayesian cue combination (Legaspi & Toyoizumi, 2019; Lush et al., 2019; Moore & Fletcher, 2012; Wolpe et al., 2013). The temporal shift of the action and the action outcome are modeled as resulting from optimal estimates of their specific timing when relevant sensory cues and prior expectations are integrated according to causality between cues. Specifically, the shifts occur only when the "action" is inferred as causal for the subsequent effect (Legaspi & Toyoizumi, 2019). In computational terms, the binding builds upon a prior belief of a causal relationship and the sensory evidence of related timing cues, independent of whether intentional action is involved. From the perspective of the Bayesian model, our increased binding of the outcome event can be viewed as reflecting an enhanced prior belief of the causal relationship between the avatar movement and the subsequent beep tone. Both our VR and control groups received identical sensory feedback in the temporal binding task, and the only difference is that the VR group had prior experience of visually controlling the avatar before the post-test. The embodiment of the avatar is thus akin to an enhanced prior belief that the avatar hand is responsible for the outcome (Desantis et al., 2011; Haering & Kiesel, 2012), which leads to an increased timing shift according to the Bayesian model of temporal binding (Legaspi & Toyoizumi, 2019). In fact, a similar Bayesian model based on causal inference also explains the sense of bodily ownership, another component of self-consciousness, as investigated in the classical rubber hand illusion paradigm (Chancel, Ehrsson, et al., 2022). More importantly, causal beliefs about relevant ownership cues, estimated from this paradigm, are implicated in the IPS, a region often associated with cue combination, as well as the angular gyrus (Chancel, Iriye, et al., 2022). These modeling and neuroimaging work thus suggest that classical measures of self, i.e., the rubber hand illusion in the sense of bodily ownership and

the temporal binding in the sense of agency, might be governed by the same causal inference mechanism with the involvement of AG.

As in our previous study (Kong et al., 2017), we found that the explicit agency rating over the avatar movements did not differ between groups, nor did it correlate to the binding or the change in binding. The lack of a group difference in explicit ratings might simply reflect that the brief VR exposure is insufficient to change the reflective judgment about the action attribution of the avatar hand. However, the lack of correlation between the two measures appears puzzling at first sight since many researchers view the verbal report of self-attribution of agency as an explicit measure of SoA and the temporal binding as an implicit measure of SoA (Berberian et al., 2012; Moore & Fletcher, 2012). First, we need to consider that a positive association between these two measures is not consistently found but depends on experiment settings and tasks. For example, priming disbelief in self-will can reduce temporal binding but not the explicit rating of self-agency, suggesting a dissociation of these two measures (Lynn et al., 2014). In contrast, the choice of action (free vs. instructed choice) and valence of action outcome (pleasant vs. unpleasant sound) can increase both temporal binding and explicit rating of control (Barlas et al., 2018), making them seemingly correlated. Thus, whether explicit and implicit measures of agency are correlated or simultaneously modulated depends on specific agency manipulation methods. Our study only manipulated the presence of the control experience of an avatar between groups, which appears insufficient to change the subjective rating of agency over the uncontrolled avatar hand in the delayed post-test. Second, the lack of group difference in explicit rating, along with its dissociation with temporal binding, is consistent with the proposition that explicit and implicit measures of SoA might probe agency processes at different functional and representational levels (Synofzik et al., 2008). Third, we note that explicit rating and temporal binding in our study targeted different

agency events. The former is about whether the avatar action is caused by oneself (see Table 1), while the latter is about the temporal or causal relationship between the avatar action and the ensuing beep sound. Thus, the rating of the explicit agency is more about "I am the one who is causing the avatar movement" than about the feeling of authorship that "I am the one who is causing the tone through my own action," which is relevant for temporal binding. In this line of thinking, we postulate that the difference in measured entity might explain the widely reported dissociation between explicit and implicit measures of agency (Dewey & Knoblich, 2014; Lynn et al., 2014; Saito et al., 2015; Wen et al., 2015).

Previous studies on VR embodiment have largely focused on how multisensory integration affects people's self-consciousness (Banakou & Slater, 2014; Slater et al., 2009). With a brief exposure to VR, people erroneously feel that they own a virtual body part or even a full virtual body (Blanke et al., 2015; Petkova et al., 2011), mislocate themselves (Ehrsson, 2007), or change the perception of one's identity (Banakou et al., 2013; Petkova et al., 2011). The common technique is to present a vivid visual representation of an avatar and match it spatiotemporally with sensory cues from other modalities, including tactile, auditory, and proprioceptive cues (Banakou & Slater, 2014; Slater et al., 2009). Neuroimaging studies have shown that the premotor cortex and TPJ are key areas for bodily ownership (Bekrater-Bodmann et al., 2014; Pamplona et al., 2022) and self-location (Ionta et al., 2011; Lenggenhager et al., 2011). However, the sense of agency over a virtual body has never been studied. Experimentally, its modulation in VR would demand action control of the avatar as opposed to passively receiving sensory stimuli pertinent to self. Our findings provide compelling evidence that sensorimotor control experience with an avatar can lead to agency changes over the avatar movement, whose neural correlates center at the right AG, a key area that is also associated with agency arising from actions in real settings. Given this cluster

covers high-order associative regions, we postulate that the VR embodiment effect is potentially generalizable to other tasks beyond the temporal binding. For instance, SoA arising from voluntary actions contributes to perceptual attenuation of action-induced sensory stimuli (Blakemore et al., 1998; Shergill et al., 2005) or self-other distinction (Kahl & Kopp, 2018). Whether these perceptual tasks are affected by similar avatar-control experiences in VR warrants further investigation.

Our findings raised possible problems for the era of VR or metaverse. First, despite the fact that our participants did not change their self-reported agency rating with the brief VR experience, it is still possible that people's explicit judgment of agency can be modulated by long-term VR use. Second, individuals with neurological and psychiatric disorders experience disrupted SoA and illusions in their daily life (Edwards & Bhatia, 2012; Frith et al., 2000), and even neurotypical individuals can occasionally experience faulty SoA, say, with sensory priming (Aarts et al., 2005; Wegner & Wheatley, 1999). Whether certain populations' self-consciousness is negatively affected by the experience of controlling an avatar is an important open question from the perspective of psychopathology. Third, given the observed immediate behavioral and neural effect of an embodied avatar on agency, we expect that unintended "actions" of the avatar, accidentally caused by technical glitches in the virtual worlds, might affect the avatar owner's sense of agency and even lead to psychological harm (Cheong, 2022). These previously rare scenarios might lead to potential legal issues about how to account for the responsibility of compromising someone's sense of agency in the metaverse.

In conclusion, the temporal binding elicited by passive observation of an embodied virtual body is subserved by the right AG with an extension to IPL, key regions for causal inference and inferential

sensemaking. In contrast, traditional action planning areas (e.g., pre-SMA), widely observed in studies on sense of agency arising from voluntary actions, are not implicated. These findings support the reconstructive theory of agency that emphasizes retrospective processing of agency-related cues and suggests that the experience of controlling an avatar might enhance the causal belief of avatar action and its action outcome, leading to increased temporal binding, which often serves as a surrogate of agency. Given that people's causal belief over an avatar's action can be built up, both behaviorally and neurophysiologically, by a brief period of avatar control in immersive VR, how the embodiment of a virtual body affects our self-consciousness and other psychological constructs would pose as a novel problem when our populace spends increasing time in virtual or digital worlds.

Reference

Aarts, H., Custers, R., & Wegner, D. M. (2005). On the inference of personal authorship: Enhancing experienced agency by priming effect information. *Consciousness and Cognition*, *14*(3), 439–458. https://doi.org/10.1016/j.concog.2004.11.001

Bach, F., Cakmak, H., Maass, H., Bekrater-Bodmann, R., Foell, J., Diers, M., Trojan, J., Fuchs, X., & Flor, H. (2012). Illusory Hand Ownership Induced by an MRI Compatible Immersive Virtual Reality Device. *Biomedical Engineering-Biomedizinische Technik*, *57*, 718–720. https://doi.org/10.1515/bmt-2012-4084

Balslev, D., Nielsen, F. A., Lund, T. E., Law, I., & Paulson, O. B. (2006). Similar brain networks for detecting visuo-motor and visuo-proprioceptive synchrony. *Neuroimage*, *31*(1), 308–312. https://doi.org/10.1016/j.neuroimage.2005.11.037

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

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

Barlas, Z., Hockley, W. E., & Obhi, S. S. (2018). Effects of free choice and outcome valence on the sense of agency: Evidence from measures of intentional binding and feelings of control. *Experimental Brain Research*, *236*(1), 129–139. https://doi.org/10.1007/s00221-017-5112-3

Bekrater-Bodmann, R., Foell, J., Diers, M., Kamping, S., Rance, M., Kirsch, P., Trojan, J., Fuchs, X., Bach, F., Çakmak, H. K., Maaß, H., & Flor, H. (2014). The Importance of Synchrony and Temporal Order of Visual and Tactile Input for Illusory Limb Ownership

Experiences – An fMRI Study Applying Virtual Reality. *PLoS ONE*, *9*(1), Article 1. https://doi.org/10.1371/journal.pone.0087013

Berberian, B., Sarrazin, J., Le Blaye, P., & Haggard, P. (2012). Automation Technology and Sense of Control: A Window on Human Agency. *PLoS ONE*, *7*(3). https://doi.org/10.1371/journal.pone.0034075

Blakemore, S. J., Wolpert, D. M., & Frith, C. D. (1998). Central cancellation of self-produced tickle sensation. *Nature Neuroscience*, *1*(7), 635–640. https://doi.org/10.1038/2870

Blanke, O., Slater, M., & Serino, A. (2015). Behavioral, Neural, and Computational Principles of Bodily Self-Consciousness. *Neuron*, *88*(1), 145–166. https://doi.org/10.1016/j.neuron.2015.09.029

Borhani, K., Beck, B., & Haggard, P. (2017). Choosing, Doing, and Controlling: Implicit Sense of Agency Over Somatosensory Events. *Psychological Science*, *28*(7), 882–893. https://doi.org/10.1177/0956797617697693

Buehner, M. J. (2012). Understanding the Past, Predicting the Future: Causation, Not Intentional Action, Is the Root of Temporal Binding. *Psychological Science*, *23*(12), 1490–1497. https://doi.org/10.1177/0956797612444612

Buehner, M. J., & Humphreys, G. R. (2009). Causal Binding of Actions to Their Effects. *Psychological Science*, *20*(10), 1221–1228. https://doi.org/10.1111/j.1467-9280.2009.02435.x

Caspers, S., Zilles, K., Laird, A. R., & Eickhoff, S. B. (2010). ALE meta-analysis of action observation and imitation in the human brain. *NeuroImage*, *50*(3), 1148–1167. https://doi.org/10.1016/j.neuroimage.2009.12.112

Cavazzana, A., Penolazzi, B., Begliomini, C., & Bisiacchi, P. S. (2015). Neural underpinnings of the "agent brain': New evidence from transcranial direct current stimulation. *European Journal of Neuroscience*, *42*(3), Article 3. https://doi.org/10.1111/ejn.12937

Chambon, V., Moore, J. W., & Haggard, P. (2015). TMS stimulation over the inferior parietal cortex disrupts prospective sense of agency. *Brain Structure and Function*, *220*(6), 3627–3639. https://doi.org/10.1007/s00429-014-0878-6

Chambon, V., Wenke, D., Fleming, S. M., Prinz, W., & Haggard, P. (2013). An Online Neural Substrate for Sense of Agency. *Cerebral Cortex*, *23*(5), 1031–1037. https://doi.org/10.1093/cercor/bhs059

Chancel, M., Ehrsson, H. H., & Ma, W. J. (2022). Uncertainty-based inference of a common cause for body ownership. *ELife*, *11*. https://doi.org/10.7554/elife.77221

Chancel, M., Iriye, H., & Ehrsson, H. (2022). Causal Inference of Body Ownership in the Posterior Parietal Cortex. *Journal of Neuroscience*, *42*(37), 7131–7143. https://doi.org/10.1523/JNEUROSCI.0656-22.2022

Charalampaki, A., Karabanov, A. N., Ritterband-Rosenbaum, A., Nielsen, J. B., Siebner, H. R., & Christensen, M. S. (2022). Sense of agency as synecdoche: Multiple neurobiological mechanisms may underlie the phenomenon summarized as sense of agency. *Consciousness and Cognition*, *101*, 103307. https://doi.org/10.1016/j.concog.2022.103307

Cheong, B. C. (2022). Avatars in the metaverse: Potential legal issues and remedies. *International Cybersecurity Law Review, Journal Article*. https://doi.org/10.1365/s43439-022-00056-9

Cunnington, R., Windischberger, C., Deecke, L., & Moser, E. (2003). The preparation and readiness for voluntary movement: A high-field event-related fMRI study of the

Bereitschafts-BOLD response. *NeuroImage*, *20*(1), 404–412. https://doi.org/10.1016/S1053-8119(03)00291-X

David, N., Cohen, M. X., Newen, A., Bewernick, B. H., Shah, N. J., Fink, G. R., & Vogeley, K. (2007). The extrastriate cortex distinguishes between the consequences of one's own and others' behavior. *Neuroimage*, *36*(3), 1004–1014. https://doi.org/10.1016/j.neuroimage.2007.03.030

de Bezenac, C. E., Sluming, V., Gouws, A., & Corcoran, R. (2016). Neural response to modulating the probability that actions of self or other result in auditory tones: A parametric fMRI study into causal ambiguity. *Biological Psychology*, *119*, 64–78. https://doi.org/10.1016/j.biopsycho.2016.07.003

Desantis, A., Roussel, C., & Waszak, F. (2011). On the influence of causal beliefs on the feeling of agency. *Consciousness and Cognition*, *20*(4), Article 4. https://doi.org/10.1016/j.concog.2011.02.012

Desmurget, M., Reilly, K., Richard, N., Szathmari, A., Mottolese, C., & Sirigu, A. (2009). Movement Intention After Parietal Cortex Stimulation in Humans. *Science*, *324*(5928), 811–813. https://doi.org/10.1126/science.1169896

Dewey, J. A., & Knoblich, G. (2014). Do Implicit and Explicit Measures of the Sense of Agency Measure the Same Thing? *PLoS ONE*, *9*(10), e110118. https://doi.org/10.1371/journal.pone.0110118

Di Plinio, S., Perrucci, M. G., Aleman, A., & Ebisch, S. J. H. (2020). I am Me: Brain systems integrate and segregate to establish a multidimensional sense of self. *Neuroimage*, *205*, 116284. https://doi.org/10.1016/j.neuroimage.2019.116284

Ebert, J. P., & Wegner, D. M. (2010). Time warp: Authorship shapes the perceived timing of actions and events. *Consciousness and Cognition*, *19*(1), Article 1. https://doi.org/10.1016/j.concog.2009.10.002

Edwards, M. J., & Bhatia, K. P. (2012). Functional (psychogenic) movement disorders: Merging mind and brain. *Lancet Neurology*, *11*(3), 250–260. https://doi.org/10.1016/S1474-4422(11)70310-6

Ehrsson, H. H. (2007). The Experimental Induction of Out-of-Body Experiences. *Science*, *317*(5841), 1048–1048. https://doi.org/10.1126/science.1142175

Elsner, B., & Hommel, B. (2001). Effect anticipation and action control. *Journal of Experimental Psychology-Human Perception and Performance*, *27*(1), 229–240. https://doi.org/10.1037/0096-1523.27.1.229

Farrer, C., Franck, N., Georgieff, N., Frith, C. D., Decety, J., & Jeannerod, A. (2003). Modulating the experience of agency: A positron emission tomography study. *Neuroimage*, *18*(2), 324–333. https://doi.org/10.1016/S1053-8119(02)00041-1

Farrer, C., Frey, S. H., Van Horn, J. D., Tunik, E., Turk, D., Inati, S., & Grafton, S. T. (2008). The Angular Gyrus Computes Action Awareness Representations. *Cerebral Cortex*, *18*(2), 254–261. https://doi.org/10.1093/cercor/bhm050

Farrer, C., & Frith, C. D. (2002). Experiencing oneself vs another person as being the cause of an action: The neural correlates of the experience of agency. *Neuroimage*, *15*(3), 596–603. https://doi.org/10.1006/nimg.2001.1009

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175–191. https://doi.org/10.3758/bf03193146

Ferri, F., Frassinetti, F., Ardizzi, M., Costantini, M., & Gallese, V. (2012). A Sensorimotor Network for the Bodily Self. *Journal of Cognitive Neuroscience*, *24*(7), Article 7. https://doi.org/10.1162/jocn_a_00230

Fotopoulou, A., Tsakiris, M., Haggard, P., Vagopoulou, A., Rudd, A., & Kopelman, M. (2008). The role of motor intention in motor awareness: An experimental study on anosognosia for hemiplegia. *Brain*, *131*, 3432–3442. https://doi.org/10.1093/brain/awn225

Friston, K. (2010). The free-energy principle: A unified brain theory? *Nature Reviews Neuroscience*, *11*(2), 127–138. https://doi.org/10.1038/nrn2787

Frith, C. D., Blakemore, S. J., & Wolpert, D. M. (2000). Explaining the symptoms of schizophrenia: Abnormalities in the awareness of action. *Brain Research Reviews*, *31*(2–3), 357–363. https://doi.org/10.1016/S0165-0173(99)00052-1

Fugelsang, J. A., Roser, M. E., Corballis, P. M., Gazzaniga, M. S., & Dunbar, K. N. (2005). Brain mechanisms underlying perceptual causality. *Cognitive Brain Research*, *24*(1), 41–47. https://doi.org/10.1016/j.cogbrainres.2004.12.001

Gallagher, S. (2000). Philosophical conceptions of the self: Implications for cognitive science. *Trends in Cognitive Sciences*, *4*(1), Article 1. https://doi.org/10.1016/S1364-6613(99)01417-5

Haering, C., & Kiesel, A. (2012). Mine is earlier than yours: Causal beliefs influence the perceived time of action effects. *Frontiers in Psychology*, *3*. https://doi.org/10.3389/fpsyg.2012.00393

Haggard, P. (2005). Conscious intention and motor cognition. *Trends in Cognitive Sciences*, *9*(6), 290–295. https://doi.org/10.1016/j.tics.2005.04.012

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

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

Haggard, P., & Tsakiris, M. (2009). The Experience of Agency: Feelings, Judgments, and Responsibility. *Current Directions in Psychological Science*, *18*(4), 242–246. https://doi.org/10.1111/j.1467-8721.2009.01644.x

Hoerl, C., Lorimer, S., McCormack, T., Lagnado, D., Blakey, E., Tecwyn, E., & Buehner, M. (2020). Temporal Binding, Causation, and Agency: Developing a New Theoretical Framework. *Cognitive Science*, *44*(5). https://doi.org/10.1111/cogs.12843

Ionta, S., Heydrich, L., Lenggenhager, B., Mouthon, M., Fornari, E., Chapuis, D., Gassert, R., & Blanke, O. (2011). Multisensory Mechanisms in Temporo-Parietal Cortex Support Self-Location and First-Person Perspective. *Neuron*, *70*(2), 363–374. https://doi.org/10.1016/j.neuron.2011.03.009

Kahl, S., & Kopp, S. (2018). A Predictive Processing Model of Perception and Action for Self-Other Distinction. *Frontiers in Psychology*, *9*, 2421. https://doi.org/10.3389/fpsyg.2018.02421

Khalighinejad, N., Di Costa, S., & Haggard, P. (2016). Endogenous action selection processes in dorsolateral prefrontal cortex contribute to sense of agency: A meta-analysis of tDCS studies of "intentional binding." *Brain Stimulation*, *9*(3), Article 3. https://doi.org/10.1016/j.brs.2016.01.005

Kikuchi, T., Sugiura, M., Yamamoto, Y., Sasaki, Y., Hanawa, S., Sakuma, A., Matsumoto, K., Matsuoka, H., & Kawashima, R. (2019). Neural responses to action contingency error in different cortical areas are attributable to forward prediction or sensory processing. *Scientific Reports*, *9*(1). https://doi.org/10.1038/s41598-019-46350-1

Kong, G., He, K., & Wei, K. (2017). Sensorimotor experience in virtual reality enhances sense of agency associated with an avatar. *Consciousness and Cognition*, *52*, 115–124. https://doi.org/10.1016/j.concog.2017.04.018

Kuhn, S., Brass, M., & Haggard, P. (2013). Feeling in control: Neural correlates of experience of agency. *Cortex*, *49*(7), Article 7. https://doi.org/10.1016/j.cortex.2012.09.002

Legaspi, R., & Toyoizumi, T. (2019). A Bayesian psychophysics model of sense of agency. *Nature Communications*, *10*(1), Article 1. https://doi.org/10.1038/s41467-019-12170-0

Lenggenhager, B., Halje, P., & Blanke, O. (2011). Alpha band oscillations correlate with illusory self-location induced by virtual reality. *European Journal of Neuroscience*, *33*(10), 1935–1943. https://doi.org/10.1111/j.1460-9568.2011.07647.x

Leube, D. T., Knoblich, G., Erb, M., Grodd, W., Bartels, M., & Kircher, T. T. J. (2003). The neural correlates of perceiving one's own movements. *Neuroimage*, *20*(4), 2084–2090. https://doi.org/10.1016/j.neuroimage.2003.07.033

Leube, D. T., Knoblich, G., Erb, M., & Kircher, T. T. J. (2003). Observing one's hand become anarchic: An fMRI study of action identification. *Consciousness and Cognition*, *12*(4), 597–608. https://doi.org/10.1016/S1053-8100(03)00079-5

Lush, P., Roseboom, W., Cleeremans, A., Scott, R., Seth, A., & Dienes, Z. (2019). Intentional Binding as Bayesian Cue Combination: Testing Predictions With Trait Individual Differences. *Journal of Experimental Psychology-Human Perception and Performance*, *45*(9), 1206–1217. https://doi.org/10.1037/xhp0000661

Lynn, M., Muhle-Karbe, P., Aarts, H., & Brass, M. (2014). Priming determinist beliefs diminishes implicit (but not explicit) components of self-agency. *Frontiers in Psychology*, *5*. https://doi.org/10.3389/fpsyg.2014.01483

Mason, R., & Just, M. (2011). Differentiable Cortical Networks for Inferences Concerning People's Intentions Versus Physical Causality. *Human Brain Mapping*, *32*(2), 313–329. https://doi.org/10.1002/hbm.21021

Matsuzawa, M., Matsuo, K., Sugio, T., Kato, C., & Nakai, T. (2005). Temporal Relationship between Action and Visual Outcome Modulates Brain Activation: An fMRI Study. *Magnetic Resonance in Medical Sciences*, *4*(3), 115–121. https://doi.org/10.2463/mrms.4.115

Moore, J. W., & Fletcher, P. C. (2012). Sense of agency in health and disease: A review of cue integration approaches. *Consciousness and Cognition*, *21*(1), 59–68. https://doi.org/10.1016/j.concog.2011.08.010

Moore, J. W., Lagnado, D., Deal, D. C., & Haggard, P. (2009). Feelings of control: Contingency determines experience of action. *Cognition*, *110*(2), 279–283. https://doi.org/10.1016/j.cognition.2008.11.006

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

Moore, J. W., Ruge, D., Wenke, D., Rothwell, J., & Haggard, P. (2010). Disrupting the experience of control in the human brain: Pre-supplementary motor area contributes to the sense of agency. *Proceedings of the Royal Society B-Biological Sciences*, *277*(1693), Article 1693. https://doi.org/10.1098/rspb.2010.0404

Nahab, F. B., Kundu, P., Gallea, C., Kakareka, J., Pursley, R., Pohida, T., Miletta, N., Friedman, J., & Hallett, M. (2011). The Neural Processes Underlying Self-Agency. *Cerebral Cortex*, *21*(1), 48–55. https://doi.org/10.1093/cercor/bhq059

Nieuwland, M., Petersson, K., & Van Berkum, J. (2007). On sense and reference: Examining the functional neuroanatomy of referential processing. *NeuroImage*, *37*(3), 993–1004. https://doi.org/10.1016/j.neuroimage.2007.05.048

Obleser, J., & Kotz, S. (2010). Expectancy Constraints in Degraded Speech Modulate the Language Comprehension Network. *Cerebral Cortex*, *20*(3), 633–640. https://doi.org/10.1093/cercor/bhp128

Ohata, R., Asai, T., Kadota, H., Shigemasu, H., Ogawa, K., & Imamizu, H. (2020). Sense of Agency Beyond Sensorimotor Process: Decoding Self-Other Action Attribution in the Human Brain. *Cerebral Cortex*, *30*(7), 4076–4091. https://doi.org/10.1093/cercor/bhaa028

Pamplona, G. S. P., Salgado, J. A. D., Staempfli, P., Seifritz, E., Gassert, R., & Ionta, S. (2022). Illusory Body Ownership Affects the Cortical Response to Vicarious Somatosensation. *Cerebral Cortex*, *32*(2), 312–328. https://doi.org/10.1093/cercor/bhab210

Passingham, R. E., & Lau, H. C. (2019). Acting, seeing, and conscious awareness. *Neuropsychologia*, *128*, 241–248. https://doi.org/10.1016/j.neuropsychologia.2017.06.012

Petkova, V. I., Bjornsdotter, M., Gentile, G., Jonsson, T., Li, T.-Q., & Ehrsson, H. H. (2011). From Part- to Whole-Body Ownership in the Multisensory Brain. *Current Biology*, *21*(13), 1118–1122. https://doi.org/10.1016/j.cub.2011.05.022

Poonian, S., McFadyen, J., Ogden, J., & Cunnington, R. (2015). Implicit Agency in Observed Actions: Evidence for N1 Suppression of Tones Caused by Self-made and Observed Actions. *Journal of Cognitive Neuroscience*, *27*(4), 752–764. https://doi.org/10.1162/jocn_a_00745

Renes, R. A., Van Haren, N. E. M., Aarts, H., & Vink, M. (2015). An exploratory fMRI study into inferences of self-agency. *Social Cognitive and Affective Neuroscience*, *10*(5), 708–712. https://doi.org/10.1093/scan/nsu106

Saito, N., Takahata, K., Murai, T., & Takahashi, H. (2015). Discrepancy between explicit judgement of agency and implicit feeling of agency: Implications for sense of agency and its disorders. *Consciousness and Cognition*, *37*, 1–7. https://doi.org/10.1016/j.concog.2015.07.011

Sasaki, A. T., Okamoto, Y., Kochiyama, T., Kitada, R., & Sadato, N. (2018). Distinct sensitivities of the lateral prefrontal cortex and extrastriate body area to contingency between executed and observed actions. *Cortex*, *108*, 234–251. https://doi.org/10.1016/j.cortex.2018.08.003

Schnell, K., Heekeren, K., Schnitker, R., Daumann, J., Weber, J., Hesselmann, V., Moeller-Hartmann, W., Thron, A., & Gouzoulis-Mayfrank, E. (2007). An fMRI approach to particularize the frontoparietal network for visuomotor action monitoring: Detection of incongruence between test subjects' actions and resulting perceptions. *Neuroimage*, *34*(1), 332–341. https://doi.org/10.1016/j.neuroimage.2006.08.027

Seghezzi, S., Zirone, E., Paulesu, E., & Zapparoli, L. (2019). The Brain in (Willed) Action: A Meta-Analytical Comparison of Imaging Studies on Motor Intentionality and Sense of Agency. *Frontiers in Psychology*, *10*, 804. https://doi.org/10.3389/fpsyg.2019.00804

Seghier, M. (2013). The Angular Gyrus: Multiple Functions and Multiple Subdivisions. *Neuroscientist*, *19*(1), 43–61. https://doi.org/10.1177/1073858412440596

Seghier, M. (2022). Multiple functions of the angular gyrus at high temporal resolution. *Brain Structure and Function*. https://doi.org/10.1007/s00429-022-02512-y

Shergill, S. S., Samson, G., Bays, P. M., Frith, C. D., & Wolpert, D. M. (2005). Evidence for sensory prediction deficits in schizophrenia. *American Journal of Psychiatry*, *162*(12), 2384–2386. https://doi.org/10.1176/appi.ajp.162.12.2384

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

Spengler, S., von Cramon, D. Y., & Brass, M. (2009). Was it me or was it you? How the sense of agency originates from ideomotor learning revealed by fMRI. *Neuroimage*, *46*(1), 290–298. https://doi.org/10.1016/j.neuroimage.2009.01.047

Sperduti, M., Delaveau, P., Fossati, P., & Nadel, J. (2011). Different brain structures related to self- and external-agency attribution: A brief review and meta-analysis. *Brain Structure & Function*, *216*(2), 151–157. https://doi.org/10.1007/s00429-010-0298-1

Straube, B., Van Kemenade, B. M., Arikan, B. E., Fiehler, K., Leube, D. T., Harris, L. R., & Kircher, T. (2017). Predicting the Multisensory Consequences of One's Own Action: BOLD Suppression in Auditory and Visual Cortices. *PLoS ONE*, *12*(1), e0169131. https://doi.org/10.1371/journal.pone.0169131

Suzuki, K., Lush, P., Seth, A. K., & Roseboom, W. (2019). Intentional Binding Without Intentional Action. *Psychological Science*, *30*(6), Article 6. https://doi.org/10.1177/0956797619842191

Synofzik, M., Vosgerau, G., & Newen, A. (2008). Beyond the comparator model: A multifactorial two-step account of agency. *Consciousness and Cognition*, *17*(1), Article 1. https://doi.org/10.1016/j.concog.2007.03.010

Synofzik, M., Vosgerau, G., & Voss, M. (2013). The experience of agency: An interplay between prediction and postdiction. *Frontiers in Psychology*, *4*. https://doi.org/10.3389/fpsyg.2013.00127

Tramacere, A. (2021). A causal view of the sense of agency. *Philosophical Psychology*, 1–24. https://doi.org/10.1080/09515089.2021.1985102

Tsakiris, M., Longo, M. R., & Haggard, P. (2010). Having a body versus moving your body: Neural signatures of agency and body-ownership. *Neuropsychologia*, *48*(9), 2740–2749. https://doi.org/10.1016/j.neuropsychologia.2010.05.021

Uhlmann, L., Pazen, M., Kemenade, B. M., Steinsträter, O., Harris, L. R., Kircher, T., & Straube, B. (2020). Seeing your own or someone else's hand moving in accordance with your action: The neural interaction of agency and hand identity. *Human Brain Mapping*, *41*(9), Article 9. https://doi.org/10.1002/hbm.24958

van Kemenade, B. M., Arikan, B. E., Kircher, T., & Straube, B. (2017). The angular gyrus is a supramodal comparator area in action-outcome monitoring. *Brain Structure & Function*, *222*(8), 3691–3703. https://doi.org/10.1007/s00429-017-1428-9

van Kemenade, B. M., Arikan, B. E., Podranski, K., Steinstraeter, O., Kircher, T., & Straube, B. (2019). Distinct Roles for the Cerebellum, Angular Gyrus, and Middle Temporal Gyrus in Action-Feedback Monitoring. *Cerebral Cortex*, *29*(4), 1520–1531. https://doi.org/10.1093/cercor/bhy048

Wegner, D. M. (2003). The mind's best trick: How we experience conscious will. *Trends in Cognitive Sciences*, *7*(2), 65–69. https://doi.org/10.1016/S1364-6613(03)00002-0

Wegner, D. M., & Wheatley, T. (1999). Apparent mental causation—Sources of the experience of will. *American Psychologist*, *54*(7), 480–492. https://doi.org/10.1037/0003-066X.54.7.480

Wen, W., Yamashita, A., & Asama, H. (2015). The influence of action-outcome delay and arousal on sense of agency and the intentional binding effect. *Consciousness and Cognition*, *36*, 87–95. https://doi.org/10.1016/j.concog.2015.06.004

Wende, K. C., Nagels, A., Blos, J., Stratmann, M., Chatterjee, A., Kircher, T., & Straube, B. (2013). Differences and commonalities in the judgment of causality in physical and social contexts: An fMRI study. *Neuropsychologia*, *51*(13), 2572–2580. https://doi.org/10.1016/j.neuropsychologia.2013.07.027

Wolpe, N., Haggard, P., Siebner, H. R., & Rowe, J. B. (2013). Cue integration and the perception of action in intentional binding. *Experimental Brain Research*, *229*(3), Article 3. https://doi.org/10.1007/s00221-013-3419-2

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

Xia, M., Wang, J., & He, Y. (2013). BrainNet Viewer: A Network Visualization Tool for Human Brain Connectomics. *PLoS ONE*, *8*(7), e68910. https://doi.org/10.1371/journal.pone.0068910

Xu, J., Kemeny, S., Park, G., Frattali, C., & Braun, A. (2005). Language in context: Emergent features of word, sentence, and narrative comprehension. *NeuroImage*, *25*(3), 1002–1015. https://doi.org/10.1016/j.neuroimage.2004.12.013

Yan, C.-G., Wang, X.-D., Zuo, X.-N., & Zang, Y.-F. (2016). DPABI: Data Processing & Analysis for (Resting-State) Brain Imaging. *Neuroinformatics*, *14*(3), 339–351. https://doi.org/10.1007/s12021-016-9299-4

Yomogida, Y., Sugiura, M., Sassa, Y., Wakusawa, K., Sekiguchi, A., Fukushima, A., Takeuchi, H., Horie, K., Sato, S., & Kawashima, R. (2010). The neural basis of agency: An fMRI study. *Neuroimage*, *50*(1), 198–207. https://doi.org/10.1016/j.neuroimage.2009.12.054

Zapparoli, L., Seghezzi, S., Scifo, P., Zerbi, A., Banfi, G., Tettamanti, M., & Paulesu, E. (2018). Dissecting the neurofunctional bases of intentional action. *Proceedings of the National Academy of Sciences*, *115*(28), 7440–7445. https://doi.org/10.1073/pnas.1718891115

Zapparoli, L., Seghezzi, S., Zirone, E., Guidali, G., Tettamanti, M., Banfi, G., Bolognini, N., & Paulesu, E. (2020). How the effects of actions become our own. *Science Advances*, *6*(27), Article 27. https://doi.org/10.1126/sciadv.aay8301

Zito, G. A., Wiest, R., & Aybek, S. (2020). Neural correlates of sense of agency in motor control: A neuroimaging meta-analysis. *PLoS ONE*, *15*(6), e0234321. https://doi.org/10.1371/journal.pone.0234321

Supplemental Material

ROI analysis 1: VR binding effects in the traditional agency-related network

We obtained seven ROIs associated with agency in a recent meta-analysis (Seghezzi et al., 2019). Using the reported peaks to form 8-mm-radium spherical ROIs (top panel in Figure S1a and Table S1) on the VR exposure-related contrast images (used in the between-group comparison above), we compared the ROI signals between two groups. No significant group difference was detected among these ROIs, and Bayes factors (BF_{10}, default prior scale = .707 for effect size on the alternative hypothesis H_1) indicated that for three out of the seven ROIs, the null hypothesis of no difference between groups was more than three times more likely than the alternative hypothesis (bottom panel in Figure S1a and Table S1). The ROI at the right angular gyrus was marginally significant ($t_{37.45}$ = .55, p = .051, BF_{10} = 1.46).

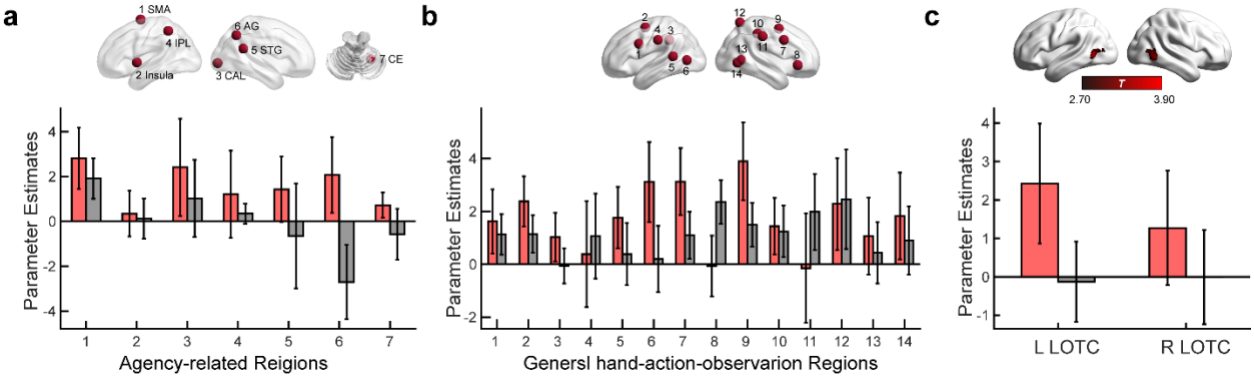

Figure S1. Group-based ROI results. (a) Agency-related ROI results. Top panel: ROIs defined by a recent meta-analysis (Seghezzi et al., 2019). Bottom panel: The experimental (red) and control (gray) information in each ROI. SMA = Supplementary motor area, Insula = Posterior insula, CAL = Calcarine scissure, IPL = Inferior parietal lobule, AG = Angular gyrus, STG = Superior temporal gyrus, CE = Cerebellum; (b) ROIs results from observation of general hand actions. Top panel: ROIs associated with the observation of hand actions. They were defined by a recent meta-analysis

(Caspers et al., 2010). Bottom panel: The experimental (red) and control (gray) information in each ROI. 1 = L IFG / PrG, 2 = L lat dPMC, 3 = L SI / IPS, 4 = L IPL, 5 = L STS / pMTG, 6 = L lat occipital, 7 = R IFG / PrG, 8 = R IFG, 9 = R lat dPMC / MFG, 10 = R SI, 11 = R IPL, 12 = R SPL, 13 = R STS / pMTG, 14 = R lat occipital; (c) ROIs results from particular observation of the avatar hand actions (operant condition) in the pre-test. Top panel: ROIs defined by the contrast of operant vs. baseline in the pre-test at the group level. Bottom panel: The experimental (red) and control (gray) information in each ROI.

Table S1. Group-based results for signals in traditional agency network (Seghezzi et al., 2019).

Location		**Coordinates (MNI, mm)**			**Correlation with binding change**		**Correlation with explicit agency**		**Between-group comparison (n_1 = 23, n_2 = 21)**				**Between-group comparison (n_1 = 18, n_2 = 18)**			
Area	**H**	***x***	***y***	***z***	ρ_{42}	***p***	ρ_{42}	***p***	***t***	***p***	**Cohen's *d***	**BF_{10}**	***t***	***p***	**Cohen's *d***	**BF_{10}**
SMA	L	-7	-4	69	-.04	.818	.28	.070	.55	.585	.16	.335	.68	.501	.23	.385
Insula	L	-41	2	1	.13	.413	.36	.018	.16	.872	.05	.301	-.32	.751	-.11	.335
CAL	R	18	-90	-1	-.11	.464	.10	.529	.49	.625	.15	.328	.38	.707	.13	.340
IPL	L	-46	-48	51	-.09	.567	.24	.110	.61	.543	.19	.347	.47	.642	.16	.351
STG	R	54	-49	22	-.11	.463	.12	.432	1.07	.292	.32	.470	1.30	.202	.43	.620
AG	R	45	-60	43	-.06	.699	.20	.196	2.01	.051	.61	1.46	2.09	.044	.70	1.68
CE	R	24	-53	-27	.11	.492	.22	.152	.55	.585	.16	.335	.11	.914	.04	.323

Note: H = Hemisphere, L = left, R = right, SMA = Supplementary motor area, Insula = Posterior insula, CAL = Calcarine scissure, IPL = Inferior parietal lobule, AG = Angular gyrus, STG = Superior temporal gyrus, CE = Cerebellum

ROI analysis 2: VR binding effects in the hand action observation network

ROIs associated with the observation of hand actions were first defined by a recent meta-analysis (Caspers et al., 2010). We used the reported peaks to form 8-mm-radium spherical ROIs (top panel in Figure S1b and Table S2) on the VR exposure-related contrast images (used in the between-group comparison) and compared the ROI signals between the two groups. We did not obtain significant group differences in these ROIs (All *p*s >.1). Bayes factors were between .298 and .912,

indicating anecdotal evidence for the null hypothesis of no difference between groups (bottom panel in Figure S1b and Table S2). For six out of the fourteen ROIs, the null hypothesis was more than three times more likely than the alternative hypothesis.

Furthermore, for the present temporal binding task, action observation was specifically elicited by the key press of the avatar hand in the operant trials. Thus, the particular ROIs associated with action observation of the avatar hand were defined as regions showing significantly stronger activation in the operant condition relative to the baseline in the pre-test for all participants (voxel-wise uncorrected $p < .005$ and cluster size > 30), and bilateral LOTC were selected (top panel in Figure S1c and Table S2; L LOTC: 144 voxels, R LOTC: 211 voxels). We did not obtain significant group differences in these two ROIs (*p*s >.1 with BF_{10} = .607 and .354, respectively; bottom panel in Figure S1c and Table S2). In sum, ROI analyses failed to support that the temporal binding is associated with neural response changes in observation-related regions.

Table S2. Group-based results for signals in the hand action observation network.

Location		**Coordinates (MNI, mm)**			**Between-group comparison (n_1 = 23, n_2 = 21)**				**Between-group comparison (n_1 = 18, n_2 = 18)**			
Area	**H**	***x***	***y***	***z***	***t***	***p***	**Cohen's *d***	**BF_{10}**	***t***	***p***	**Cohen's *d***	**BF_{10}**
Observation of general hand actions												
IFG / PrG	L	-50	6	30	.34	.733	.10	.312	.37	.716	.12	.339
lat dPMC	L	-26	-4	56	1.02	.312	.31	.454	.97	.341	.32	.463
SI / IPS	L	-36	-42	36	.94	.352	.28	.425	1.16	.254	.39	.542
IPL	L	-58	-24	36	-.26	.795	-.08	.306	-.62	.543	-.21	.373
STS / pMTG	L	-56	-48	10	.83	.411	.25	.393	.47	.640	.16	.351
lat occipital	L	-46	-70	4	1.47	.149	.44	.705	1.17	.251	.39	.546
IFG / PrG	R	52	8	36	1.29	.205	.39	.577	1.31	.198	.44	.626
IFG	R	56	30	-4	-1.68	.101	-.51	.912	-1.23	.227	-.41	.578
lat dPMC / MFG	R	36	0	54	1.43	.162	.43	.647	1.08	.289	.36	.505

Location		Coordinates (MNI, mm)			Between-group comparison (n_1 = 23, n_2 = 21)				Between-group comparison (n_1 = 18, n_2 = 18)			
Area	**H**	***x***	***y***	***z***	***t***	***p***	**Cohen's *d***	**BF_{10}**	***t***	***p***	**Cohen's *d***	**BF_{10}**
SI	R	42	-34	46	.14	.892	.04	.300	-.26	.797	-.09	.330
IPL	R	60	-26	42	-.83	.411	-.25	.393	-1.50	.144	-.05	.762
SPL	R	22	-62	64	-.07	.944	-.02	.298	.13	.898	.04	.324
STS / pMTG	R	52	-60	4	.33	.741	.10	.311	.19	.850	.06	.326
lat occipital	R	50	-66	0	.44	.664	.13	.322	.16	.871	.05	.325
Observation of avatar hand actions												
LOTC	L	-52	-66	2	1.36	.181	.41	.607	.98	.335	.33	.469
LOTC	R	50	-66	-4	.66	.515	.20	.354	.34	.740	.11	.336

Note: H = Hemisphere, L = left, R = right, IFG = inferior frontal gyrus, PrG = precentral gyrus, dPMC = dorsal premotor cortex, SI = primary somatosensory cortex, IPS = intraparietal sulcus, IPL = inferior parietal lobule, STS = superior temporal sulcus, pMTG = posterior middle temporal gyrus, MFG = middle frontal gyrus, SPL = superior parietal lobule, LOTC = lateral occipitotemporal cortex, lat = lateral.

Selection of subgroup participants

We noticed that the two groups tended to differ in the binding effect at the pre-test, though statistically insignificant (t_{42}= 1.62, p = .113; Figure 1c). Besides, two participants from the control group only underwent one operant run in a phase (one in the pre-test and another in the post-test). To test group differences in the neuroimaging data robustly, we excluded these two participants and then randomly selected 18 participants from each group to match their pre-test performance. We found that 99.83% of the possible pairs of subgroups were matched in the pre-test (t_{34} < 2.03, p < .05), and 70.27% of matched subsets still revealed a significant interaction of Phase and Group ($F_{1, 34}$ < 5.50, p < .05), which suggested that the enhanced binding effect found in the experimental group was stable. A pair of exemplary subgroups were selected to validate the imaging results of whole-group participants.

Behavioral results of subgroup participants

The temporal binding effects

The mixed ANOVA of the temporal binding effects revealed a significant Group*Phase interaction ($F_{1, 34}$ = 7.61, p = .009) and a significant main effect of Phase ($F_{1, 34}$ = 9.40, p = .004), while the main effect of Group was not significant ($F_{1, 34}$ = .65, p = .425). The Group*Phase interaction was further examined by post hoc comparisons, with participants from the experimental group showing larger temporal binding effects in the post-test (M = -112.67, $S.E.$ = 16.94) than in the pre-test (M = -59.69, $S.E.$ = 15.92): t_{34}= 4.12, Bonferroni-corrected p <.001, Cohen's d = .985. The temporal binding effects of the control group showed no significant difference between the post-test (M = -70.31, $S.E.$ =16.94) and the pre-test (M = -67.53, $S.E.$ = 15.92: t_{34}= .22, Bonferroni-corrected p =.830, Cohen's d = .052.

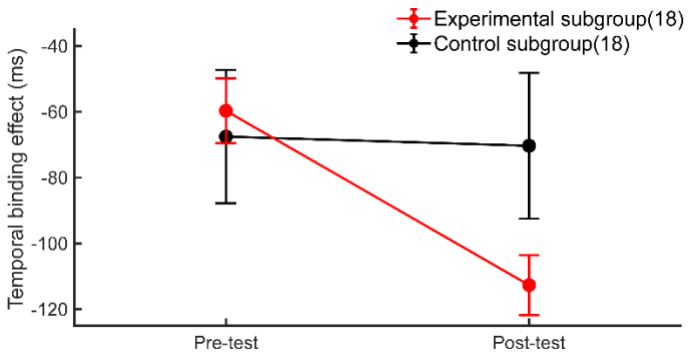

Figure S2. Temporal binding effects from two randomly selected subgroups of participants matched based on pre-test performance.

The subjective feeling of explicit agency and ownership

For all the subgroup participants, the average scores of control items were significantly higher than those of explicit agency (z = -2.76, p = .006), indicating that they correctly understood the

instructions in the questionnaire. The perceived agency and ownership did not differ between the two groups (z = -.45, p = .654 for explicit agency; z = .00, p = 1 for explicit ownership.

As expected, explicit agency and ownership were positively correlated (ρ_{34} = .62, p < .001). However, neither of them correlated with the temporal binding effect in the post-test or the binding changes between the pre- and post-test (ρ_{34} = -.24, p = .159 for explicit agency and binding effect in post-test; ρ_{34} = -.09, p = .601 for explicit ownership and binding effect in post-test; ρ_{34} = -.029, p = .865 for explicit agency and binding changes; ρ_{34} = .002, p = .989 for explicit ownership and binding changes).

fMRI results of subgroup participants

Operant binding effects (operant versus baseline) in different conditions

For the subgroup participants, we generated whole-brain contrast images of the operant condition versus the baseline, and then contrasted these images within different phases (pre- and post-test) and groups (experimental and control subgroup). All the contrasts except the experimental group in the pre-test showed significant activation in bilateral occipitotemporal cortices (LOTC; Figure S3) with a voxel-wise threshold at p < .005 and a cluster-level FWE correction at p < .05. The experimental group in the pre-test showed a trend in this region but failed to reach the significance level. These results replicated the operant binding effects in the whole group analysis.

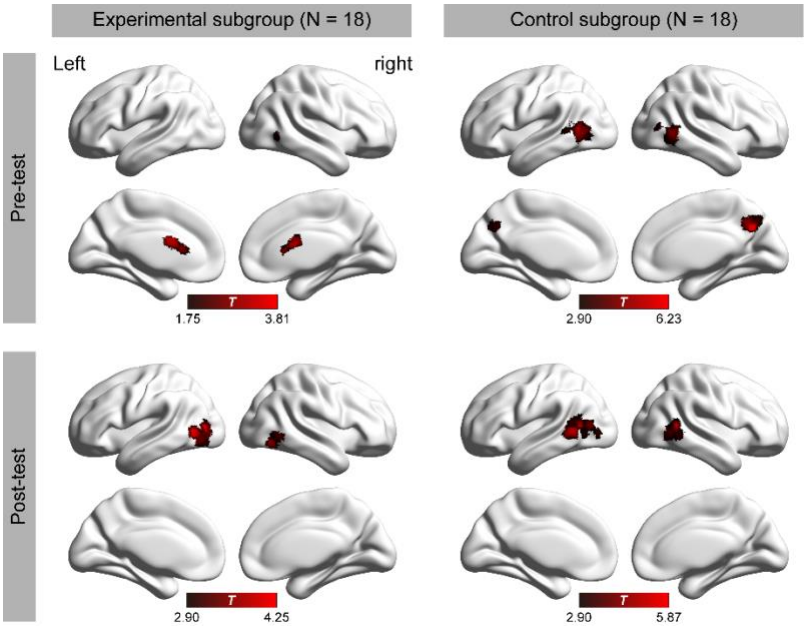

Figure S3. Whole-brain contrast images (visualized using BrainNet Viewer; Xia et al., 2013) of operant versus baseline under different phases (pre-test and post-test) in different subgroups (experimental subgroup and control subgroup). A lenient threshold was applied at voxel-wise uncorrected $p < .05$ and cluster size > 30 for the experimental subgroup in the pre-test, while the threshold of other three conditions was applied at voxel-wise $p < .005$ and cluster-level FWE correction at $p < .05$.

VR exposure-related neural changes

The whole-group analysis revealed a significant VR exposure-related neural effect centered in the angular gyrus (Figure 2c). To verify this effect in the subgroup, we calculated the neural changes of this cluster and conducted a Group*Phase mixed ANOVA, showing a significant Group*Phase interaction ($F_{1, 34} = 16.06$, $p < .001$), while the main effect of Group ($F_{1, 34} = 1.43$, $p = .241$) and Phase ($F_{1, 34} = .87$ $p = .358$) were not significant. Specifically, for the experimental subgroup, activation related to the operant versus baseline effects within this cluster increased from -3.36

(1.13) in the pre-test to 1.48 (1.24) in the post-test, while for the control subgroup, such activation decreased from 2.19 (1.13) in the pre-test to -.83 (1.24) in the post-test.

Correlation and regression analyses for neural changes and behavioral performance

For the cluster centered in the right angular gyrus, which showed a significant between-group effect (Table 2 and Figure 2c), the neural changes significantly correlated with the temporal binding changes (ρ_{34} = -.46, p = .005), but the correlation between the neural signal in the post-test and the explicit agency did not reach significance (ρ_{34} = -.24, p = .159).

Between-subgroup comparison in agency-related ROI analysis

Only the ROI at the right angular gyrus showed a significant difference (Table S1). Most of the Bayes factors were above 1/3 but below 3, indicating that it was insufficient to accept the null hypothesis of no difference between groups.

Between-subgroup comparison for action observation-related ROI signals

We did not obtain significant group differences in the ROIs associated with the observation of general hand actions and bilateral LOTC (All ps >.05 and all BF_{10} < 1; Table S2). Thus, there is no conclusive evidence for significant group differences.

Reference of supplementary

Caspers, S., Zilles, K., Laird, A. R., & Eickhoff, S. B. (2010). ALE meta-analysis of action observation and imitation in the human brain. *NeuroImage*, *50*(3), 1148–1167. https://doi.org/10.1016/j.neuroimage.2009.12.112

Seghezzi, S., Zirone, E., Paulesu, E., & Zapparoli, L. (2019). The Brain in (Willed) Action: A Meta-Analytical Comparison of Imaging Studies on Motor Intentionality and Sense of Agency. *Frontiers in Psychology*, *10*, 804. https://doi.org/10.3389/fpsyg.2019.00804

Xia, M., Wang, J., & He, Y. (2013). Viewer: A Network Visualization Tool for Human Brain Connectomics. *PLoS ONE*, *8*(7), e68910. https://doi.org/10.1371/journal.pone.0068910