A Bayesian account of body ownership

Nathan Evans[1,2†], Danilo Jimenez Rezende[1,2,3†],

Wulfram Gerstner[3], Olaf Blanke[1,2,4]

[1] Center for Neuroprosthetics, École Polytechnique Fédérale de Lausanne, Lausanne, Switzerland

[2] Laboratory of Cognitive Neuroscience, School of Life Sciences, Brain Mind Institute, École Polytechnique Fédérale de Lausanne, Lausanne, Switzerland

[3] Laboratory of Computational Neuroscience, School of Computer and Communication Sciences and School of Life Sciences, Brain Mind Institute, École Polytechnique Fédérale de Lausanne, Lausanne, Switzerland

[4] Department of Neurology, University Hospital Geneva, Geneva, Switzerland

† These authors contributed equally to this manuscript

When looking at our hand, we simultaneously feel it based on tactile and proprioceptive cues. However, seeing a fake hand being touched while our real hand is touched (but hidden from view), we experience the fake hand as belonging to us (ownership) and recalibrate our perceived hand position. Using computational modelling and data collected from an automated, machine-controlled experimental setup, we extracted, on a subject-by-subject basis, the maximal distance between the real and fake hand (threshold) and the visuo-tactile stimulation conditions that subjects tolerate before the shift in their perceived hand position breaks down. The model predicts, and experiments confirm, that ownership breaks down discontinuously near this threshold, such that subjects sometimes perceive their hand close to the fake hand, and sometimes close to the real hand. By computing the limits of ownership and limb position perception, our model paves the way for computational approaches to the embodiment of limbs.

Introduction

Recent investigations into lower-level multisensory and sensorimotor aspects of self-consciousness have led researchers to define several specific bodily experiences[1–6]. One of these bodily experiences is the feeling that our body and its parts belong to us (body ownership). In particular, the perception of upper limb ownership has been extensively studied using the rubber hand illusion (RHI[7]). In the RHI, participants watch a fake hand being stroked in synchrony with stroking on their own (occluded) hand. This manipulation alters tactile perception and induces the illusion that touch is felt on the fake hand and that the fake hand feels like one's own hand[7–9]. These subjective effects are often accompanied by a shift in the perceived position of one's own hand towards the fake hand (referred to as localization error in this manuscript but also known as *drift* in the cognitive neuroscience literature) as well as physiological changes (*e.g.* body temperature changes) [10,11], which are absent or weaker when the stroking provided to the real hand and the fake hand is not synchronous[7–10,12]. The illusion is reduced or abolished when the fake hand does not match the real hand's posture[9], when the fake hand is placed too far from the real hand[13], or when the stroking is applied in different directions[14].

Although it has been speculated that illusory hand ownership and its associated shift in perceived hand position occur as the brain's perceptual systems attempt to interpret the conflicting visual, tactile, and proprioceptive information[15–17], there is currently no computational account of the RHI and no comprehensive understanding of the role that visual, tactile (stroking or vibrations), and proprioceptive stimulation parameters (*e.g.* duration, synchrony, and limb position) play on illusory hand ownership and perceived hand position. Since systematic changes in illusory hand ownership can also be induced in virtual environments[18], we used automated, machine-controlled stroking with a virtual-hand setup (Fig. 1) to investigate whether computational modeling can account for the measured localization errors of perceived hand position. We show that a Bayesian model of causal inference can predict the conditions under which humans fuse proprioceptive and visual information during the RHI. Fusion does not occur if the distance between the real hand and the fake hand is too large or if both the real and fake hands are stimulated asynchronously for extended periods of time. Our model and data suggest that, for a critical range of separations between the fake and real hand, perceived hand position switches discontinuously between a fused and a non-fused, proprioception-dominated position.

Results

Pilot experiment

We first tested if illusory hand ownership was induced using our automated experimental setup in a comparable fashion to that described in earlier studies using experimenter-applied stroking on a physical rubber hand[7,9,10] or on a virtual hand[18]. To this aim, we performed a pilot experiment employing a 2x2 factorial design with the factors Stroking and Posture[8] where visuo-tactile stroking was provided synchronous or asynchronously on a virtual hand with a congruent or incongruent orientation with respect to the real hand (Fig. 2A; see *Methods*). Statistical analysis on questionnaire scores relevant to illusory ownership revealed that participants experienced illusory ownership for the virtual hand during synchronous stroking in a congruent hand position ($p<0.01$, *Post-hoc Wilcoxon matched-pairs test*; Fig. 2B; Table 1), but not during asynchronous stroking or if the fake hand was in an incongruent position (all $p>0.05$).

Main experiment: Visuo-proprioceptive separation

Our automated setup enabled us to systematically vary, on a trial-by-trial basis, the distance between the virtual and real hand (visuo-proprioceptive separation) and the delay between tactile and visual stimulation via animations on the virtual hand (visuo-tactile delay). If participants were not at all influenced by the position of the seen virtual hand, then the subjects' perceived hand position would be independent to the position of the virtual hand. We found that the reported hand position exhibited a systematic localization error towards the virtual hand that increased with the magnitude of visuo-proprioceptive separation. For synchronous conditions (visuo-tactile delays < 0.2s), we observed a mean localization error of 5±3cm for separations from 0 to 10cm, and a mean localization error of 10±5cm for separations of 10 to 20cm. Additionally, we found a strong positive correlation between localization error and visuo-proprioceptive separation for the interval 0-20cm (*Pearson's product-moment test*: $t = 21.5$, $df = 766$, $p < 2e-16$, *correlation±95% confidence interval (CI)*: 0.615±0.045). For the interval 20-30cm, we found a negative correlation (*Pearson's product-moment test*: $t = -2.5$, $df = 584$, $p = 0.01$, *correlation±95% CI*: -0.10±0.08), suggesting that the participants' perceived hand position was influenced by visual information stemming from the virtual hand in these ranges (Fig. 3A). For the range 30-40cm, we found no significant correlation (*Pearson's product-moment test*: $t = 0.4$, $df = 191$, $p = 0.7$, *correlation±95% CI*: 0.03±0.14), suggesting a weak or nonexistent relationship between the perceived hand position and the virtual hand position.

Main experiment: Visuo-tactile stroking synchrony

How does visuo-tactile delay (Z) between tactile stimulation and the visual animation on the virtual hand modulate the perceived hand position? For visuo-proprioceptive separations of less than 10cm, we found that delays of $Z = 0$-1s had no significant influence on the perceived hand position. However, for separations between 20 and 30cm, the perceived hand position was significantly shifted towards the virtual hand for near-synchronous stimulation ($Z < 0.2$s) as compared to asynchronous stroking ($Z = 0.6$-1s; *Tukey multiple comparisons test*: $p = 0.003$, Fig. 3A). The largest localization error was found for near-synchronous stroking when the real and virtual hands were approximately 15 to 25cm apart. Importantly, under conditions of near-synchronous stroking

($Z < 0.2$s) and for visuo-proprioceptive separations smaller than 20cm, the perceived hand position localization error was influenced by visuo-proprioceptive separation (*linear regression*: $R^2 = 0.37$, $F = 450$, $df = 764$, beta = 0.56). By contrast, we found a broader distribution of perceived hand positions when the visuo-proprioceptive separation was between 20 and 30cm, indicating that hand localization was much less influenced by visuo-proprioceptive separation for larger separations (*linear regression*: $R^2 = 0.01$, $F = 6.4$, $df = 584$, beta = -0.3). The spread of localization errors for all participants is shown in Figure 3A in these different visuo-proprioceptive separation ranges.

Model: Rationale and Formulation

In order to understand the distribution of perceived hand positions, we developed a model of how subjects integrate sensory information from vision (arising from the virtual hand on the head-mounted display) and proprioception (hand position as estimated from proprioceptive signals from the subject's real hand). Additionally, we incorporated into our model how this integration is influenced by ownership as manipulated through additional visuo-tactile stimulation (stroking of the virtual and real hands with different visuo-tactile delays).

If the real hand is at position Q, the position of the hand as estimated by proprioceptive cues is formulated as $X_p = Q + \eta_p$, where the noise η_p is assumed to be Gaussian distributed with a standard deviation σ_p that reflects the lack of precision of the proprioceptive information. Analogously, we modeled the imprecision of visual cues with a Gaussian noise of standard deviation σ_v.. The results discussed below aggregates data from all stroking duration values, except where otherwise noted.

To test whether a Bayesian ideal observer model with access to visual, tactile, and proprioceptive cues provides a reliable explanation of the perceived hand position in the RHI, we hypothesized that the perceived hand position is based on a combination of prior beliefs (top-down influences) and sensory input to three sensory modalities (vision, proprioception, touch). The model relies on three assumptions: (i) incoming visual and proprioceptive information is independent, *i.e.* the firing of primary visual neurons and proprioceptive neurons are statistically independent for a given sensory stimulus; (ii) visuo-tactile synchrony conveys information about hand ownership; and (iii) prior to integrating visual and proprioceptive information, the subject makes a (likely unconscious) top-down decision as to whether the virtual hand is one's own (hand ownership). Assumptions (ii) and (iii) are based on findings from previous RHI studies demonstrating that visuo-tactile synchrony modulates hand ownership and that top-down information can modulate ownership.

Our RHI model is composed of two sub-models. First, a *perception model* describes how the participants form and maintain their internal percepts of their hand position as well as ownership of the virtual hand. Second, a *response model* captures subject reports when asked about his or her percept. In analogy to the model proposed in[19], our perception model is an encoder of sensory related information, while our response model produces meaningful decisions by decoding the representation formed by the sensory encoder.

Suppose that the subject believes that the virtual hand is his or her own hand and that their real hand is located at position Q. In this case, both the visual (X_v) and proprioceptive (X_p) position signals ought to be distributed around the real hand position. Using assumption (i), we define the stimulus likelihood:

$$p(X_p, X_v \mid Q) = N(X_v; Q, \sigma_v) N(X_p; Q, \sigma_p) \quad (1)$$

where $N(x;\mu,\sigma)$ is a Gaussian distribution evaluated at x, with mean μ and standard deviation σ. This simple model has been successfully used to explain visuo-auditory, visuo-spatial, and visuo-proprioceptive integration tasks[20–25].

We further characterized visuo-tactile synchrony as the delay Z between the visual and vibrotactile stroking patterns. Due to noise in the sensory systems, we modeled the perceived delay to fluctuate around 0s with a small variance σ_Z. Since visuo-tactile delay is always positive, we model its likelihood with an exponential distribution $E(Z,\sigma_Z)=\exp(-\frac{Z}{\sigma_Z})/\sigma_Z$. Taking into account this visuo-tactile likelihood term, we extend the visuo-proprioceptive likelihood in Eq. (1) to:

$$p(X_p,X_v,Z\,|\,Q)=N(X_v;Q,\sigma_v)N(X_p;Q,\sigma_p)E(Z,\sigma_Z) \qquad (2)$$

This equation defines the distribution of perceptual measurements from the three sensory systems (vision, proprioception and visuo-tactile delay) if he or she believes that their hand is at position Q and that the seen hand is their own hand. If one does not believe the seen hand to be his or her own hand, the visual position signal no longer fluctuates around the real hand position, but rather around a mean $\bar{Q}$, whose value is unknown to the subject. Analogously, the visuo-tactile delay also fluctuates around an unknown mean $\bar{Z}$. Because $\bar{Q}$ and $\bar{Z}$ are both unknown, we marginalized both variables over a large range in both the visual and visuo-tactile delay domains using a flat prior, $p(\bar{Q},\bar{Z})=\vartheta$, where ϑ is a constant. If the size of this range is much larger than the size of the visual modality's standard deviation (σ_v) and the visuo-tactile standard deviation (σ_Z), the marginalized likelihood can be well approximated by:

$$\iint p(X_p, X_v, Z, \bar{Q}, \bar{Z} \mid Q) d\bar{Q} d\bar{Z} \approx N(X_p; Q, \sigma_p)\vartheta .$$

Given this, the general likelihood function that accounts for both possible ownership beliefs is defined as:

$$p(X_p, X_v, Z \mid Q, own) = \begin{cases} N(X_v; Q, \sigma_v) N(X_p; Q, \sigma_p) E(Z, \sigma_Z) & \text{if } own = 1 \\ N(X_p; Q, \sigma_p)\vartheta & \text{if } own = 0 \end{cases} \quad (3)$$

where the binary variable '*own*' models the cognitive belief that one has regarding ownership of the virtual hand. Because the brain has neither direct access to the real hand position (Q) nor the ownership assignment of the virtual hand (*own*=0,1), it must deduce both values from sensory cues.

The resulting model (Eq. (3); Fig. 4A) can be seen as an extension of previously proposed models for causal inference[26–28]. However, in contrast to these previous models, our model accounts for three sensory modalities rather than two. Though our model is reminiscent of "window of integration" models[29–31], where the sensory delay (or "integration window") considers a delay between the percepts of two fused sensory modalities (*e.g.* visual and auditory), our model accounts for delay as a *third* sensory modality (tactile).

To an *ideal observer* receiving a multisensory stimulus $\{X_v, X_p, Z\}$, knowledge of variables Q and *own* is obtained via the posterior distribution using Bayes formula:

$$p(Q, own \mid X_p, X_v, Z) = \frac{p(X_p, X_v, Z \mid Q, own) p_0(Q, own)}{\sum_{own} \int p(X_p, X_v, Z \mid Q, own) p_0(Q, own) \; dQ} \quad (4)$$

where $p_0(Q, own)$ is the prior knowledge that the subject has about latent variables $\{Q, own\}$. The left side of the equation, $p(Q, own \mid X_p, X_v, Z)$, is also known as the *belief state*, as it provides a measure of how much an ideal observer believes that a particular pair of values $\{Q, own\}$ corresponds to the true hand position and the true ownership state. The model in Eq. (4) differs from standard multisensory integration models[20–25] because it incorporates the concept of ownership and takes into account three sensory modalities: vision, touch and proprioception.

Though participants in RHI experiments report their perceived hand position, we do not have access to their internal ownership assignment. By marginalizing over the ownership variable, we obtain:

$$p(Q \mid X_p, X_v, Z) = \frac{\sum_{own} p(X_p, X_v, Z \mid Q, own) p_0(Q, own)}{\sum_{own} \int p(X_p, X_v, Z \mid Q, own) p_0(Q, own) \; dQ} \quad (5)$$

We assume the prior $p_0(Q, own)$ to be flat for *Q,* but for consistency and generalization purposes, we consider the prior for *own* to be adaptive and parameterized as $p_0(Q, own = 1) = c$, where $c \in [0,1]$.

By substituting $p_0(Q, own)$ into Eq. (5) and integrating over Q, we obtain the final probability density function in the form of a Gaussian Mixture Model with a mixture coefficient α that is also a function of the perceptual stimuli:

$$
\begin{aligned}
p(Q|X_p, X_v, Z) &= \alpha N(Q; X_p, \sigma_p) + (1-\alpha) N(Q; \lambda X_v + (1-\lambda) X_p, 1/\sqrt{\frac{1}{\sigma_p^2} + \frac{1}{\sigma_v^2}}) \\
\alpha &= \left[1 + \frac{c}{\vartheta(1-c)} N(X_v - X_p; 0, \sqrt{\sigma_p^2 + \sigma_v^2}) E(Z, \sigma_Z)\right]^{-1} \in [0,1] \qquad (6) \\
\lambda &= \sigma_p^2 / (\sigma_p^2 + \sigma_v^2)
\end{aligned}
$$

The first term on the right-hand side in the first line of Eq. (6) accounts for the situation where the subject does not believe that the seen virtual hand is their own hand and therefore relies only on proprioceptive cues. This term contributes to the final estimate with a weight α. The second term (weight: $1-\alpha$) describes a Gaussian with a center that represents a weighted average between visual (weight λ) and proprioceptive information (weight: $1 - \lambda$). Note that the constants ϑ and c can be merged into a single parameter $\eta = c/[\vartheta(1-c)]$ without a loss of generality. Thus, the set of free parameters in our model is $\{\sigma_Z, \sigma_v, \sigma_p\}$.

The variable α has a particularly important meaning, as $1-\alpha$ is the *posterior probability of the ownership* of the virtual hand given the prior and the sensory input, namely: $1-\alpha = p(own = 1 | stimulus)$. If ownership of the virtual hand is certain, *i.e.* $p(own = 1 | stimulus) = 1$, then the position $\lambda X_v + (1-\lambda) X_p$ from Eq. (6) is equivalent to the *Maximum Likelihood Estimate* (MLE) of the perceived hand position when the subject fuses vision and proprioception.

Model: Unimodal sensory estimates

Estimating proprioceptive noise σ_p and visual noise σ_v ideally requires two additional, independent experiments designed specifically to measure these parameters. Due to the large amount of data required for our statistical analysis, performing these supplementary measurements would have substantially increased the time required per subject, which was already quite long (2 - 3 hours per subject; see *Methods*). Thus, rather than performing three independent experiments, we decided to extract the visual and proprioceptive perceptual noise from our data set based on observations from previous studies that have shown that (i) perceived hand position relies solely or mostly on proprioceptive signals if the fake hand is placed far away (*e.g.* >30cm) from the real hand[13] and (ii) visual and proprioceptive information is fused for small (*e.g.* <10cm) visuo-proprioceptive separations[21]. We therefore estimated the proprioceptive standard deviation ($\sigma_p \approx 7.2$ cm; see *Methods*) from data points with large visuo-proprioceptive separations (>30 cm) and the visual standard deviation ($\sigma_v \approx 3.8$cm; see *Methods*) from small visuo-proprioceptive separations (<10cm). To ensure that our fitting results were not dependent on the choice of the ranges selected for small and large visuo-proprioceptive separations, we simultaneously fit all parameters using the full data set without partitioning the data and found that the joint fit yielded results compatible with to those obtained with our splitting approach (σ_p: 6.8 ± 0.5cm; σ_v: 4.6 ± 0.4cm; *mean±95% CI*).

Ownership

The task of the subject is to infer the position of their real hand from the three sources of sensory information (*i.e.*, proprioception, vision, visuo-tactile delay). In contrast to classical sensor fusion paradigms[20–25], the RHI paradigm has the additional feature that the visual information source may or may not coincide with the participant's own hand. Thus, our model is formulated such that visual and proprioceptive sources are combined only if the subject has reason to believe that both vision and proprioception relate to the same object in the world. More precisely, we hypothesized that a probabilistic ownership variable is assigned for the seen virtual hand that can take one of two states: '*own*=1' indicates that the virtual hand is 'mine' (*i.e.* the participant's real hand) and '*own*=0' indicates that the virtual hand is 'not mine' (*i.e.*, not the participant's real hand; Fig. 4B, top right node). The perceived visuo-tactile delay was incorporated into the model under the assumption that if '*own*=1', the perceived visuo-tactile delay is small (*i.e.* on the order of the tactile delay), whereas if ‘*own*=0’, the perceived visuo-tactile delay is evenly distributed across a broad range.

Based on the unimodal estimates of the precision in the visual and proprioceptive channels (see above), our model assigns a probability of perceived ownership over the full range of visuo-proprioceptive separations from 0 to 40cm for both near-synchronous and asynchronous visuo-tactile stimulations (see *Methods*). This analysis revealed that for near-synchronous stroking at separations less than 20cm, the model reliably generates the percept of owning the virtual hand whereas separations greater than 30cm do not generally give rise to ownership (Fig. 4C). The ownership threshold, defined as the visuo-proprioceptive separation where subjects report ownership with 50% probability, was found to be approximately 25cm.

Model: Visuo-proprioceptive separation and stroking synchrony predictions

The model accurately predicts the observed distribution of localization errors in perceived hand position across the large range of visuo-proprioceptive separations that we measured in our experiment (Figs. S1A, S1B). It predicts that for near-synchronous stimulation and visuo-proprioceptive separations of 20 to 30cm, the distribution of localization errors has two peaks (Fig. 3B). The first, sharp peak accounts for large localization errors caused by trials where our model assigns hand ownership (and therefore fuses visual and proprioceptive cues). The second, broad peak around zero-localization error accounts for trials where our model does not assign ownership for the virtual hand.

Evidence for a double-peaked distribution was found by fitting a Gaussian Mixture Model (GMM) to the data for near-synchronous trials ($Z < 0.2$s) and testing a double-peaked GMM versus a single-peaked GMM using the Bayesian Information Criterion (BIC; see *Methods*) for the separation range of 20-30cm. This analysis was first performed on the group data, which showed the double-peaked model to better explain the distribution (single-peaked GMM: *log-likelihood*=-780, *BIC*=1571; double-peaked GMM: *log-likelihood*=-769, *BIC*=1560). For individual subject analyses, however, we found the single-peaked GMM to be better for 7 subjects, the double-peaked GMM better for 5 subjects, and for the 6 remaining subjects, there was not enough data in the visuo-proprioceptive separation range to reliably perform the analysis.

The double-peaked histogram might be caused by inter-individual differences such that at a separation of 25cm, some subjects assign ownership whereas others do not. Alternatively, it might arise intra-individually in subjects who, for the same stimulus, sometimes assign ownership and sometimes do not. We tested for each of these hypotheses and found that the ownership threshold varies both between and within single subjects (Fig. 4D), and that the ownership response is double-peaked for visuo-proprioceptive separations that are close to the threshold of 50% ownership probability for 5 subjects out of 18.

Model Comparisons

Finally, we confirmed the non-linearity of the relationship between the visual, proprioceptive, and tactile cues by comparing our proposed model with predictions from a linear sensory integration model that does not account for ownership or visuo-tactile cues. To this end, we used the Deviation Information Criterion (DIC, see *Methods*), a measure of model goodness that takes into account both the goodness-of-fit and the model complexity, and where smaller DIC values indicate better models. Our analyses showed the present model to better explain the empirical data than a two-sense linear model, both for the group data and for 10 out of 17 individual participants (Fig. 5A).

Single subject detailed analysis

To test that our model makes accurate predictions on a single-subject basis, we performed a detailed analysis for one of the double-peaked subjects (Fig. 6). In the visuo-proprioceptive separation range of 0 – 40cm, our analyses showed the model to not significantly differ from the observed data (*visuo-propriceptive separation 0-10cm*: df=18, p=0.09, N=11; *10-20cm*: df=12, p=0.30, N=6; *20-30cm*: df=104, p=0.13, N=50; *30-40cm*: df=28, p=0.27, N=6; χ^2 *test*).

Stroking duration

Previous RHI studies have manually applied stroking using a large variety of different stroking durations (Table 2), but little work to date has investigated whether longer trials with more strokes are more efficient than shorter trials in inducing illusory hand ownership and larger localization errors. We found no significant effect on localization error between short trials (fewer than 40 strokes) and long trials (more than 45 strokes) for near-synchronous stroking. However, for asynchronous stroking, long trials at large visuo-proprioceptive separations (20-30cm) induced significantly less localization error than short trials (*Tukey multiple comparisons test*, adjusted p-value = 0.012; Fig. 7). Thus, we found that subjects are more likely to detect the inconsistency between the visual and tactile cues (visuo-tactile delay) in long trails than in short trials.

Discussion

A fully automated RHI setup allowed us to perform a systematic analysis of the relationship and relative importance of visual, proprioceptive, and ownership cues (manipulated through an additional visuo-tactile stimulus) in their contribution to perceived hand position. Earlier work studied has ownership and perceived hand position with virtual hands presented on a distanced, rear projection screen[18], a monitor[32], or a video-projector[33]. Here, we built upon these earlier approaches and projected an immersive virtual reality scenario in a head-mounted display where the virtual hands were seen as extending from our participants' bodies in stereoscopic vision[34].

In contrast to previous RHI studies using binary, factorial designs[7,9,10], our design tested the effects of visuo-proprioceptive separation[13], delay and duration in a continuous fashion across a large range of values (Table 2). Importantly, the present data show that the visuo-proprioceptive separation of approximately 15-25cm is optimal to induce changes in perceived hand position that depend on visuo-tactile delay, in correspondence with these previous studies that use this separation range to induce the illusion[7,9,10]. At small separations (<10cm), we found that localization errors were not influenced by delay[35], whereas large separations (>30cm) induced small localization errors with a large variability that reflects the unreliability of proprioceptive signals[13]. Extending data from a recent behavioral RHI study[35], we also found that prolonged stimulation did not boost, but rather significantly decreased localization error in perceived hand position during asynchronous stimulation for large visuo-proprioceptive separations (20-30cm), whereas no effects of duration were observed during near-synchronous stimulation.

Although the perceived hand position was on average biased towards the virtual hand for all separations between 0-40cm, large localization errors were induced most reliably under conditions of near-synchronous stroking at separations of 10-20cm. For separations of 20-30cm, the localization error was also large, but unreliable (Fig. 3B) and subject-dependent (Fig. 4D). Our hierarchical Bayesian inference model accounts not only for the average localization error in perceived hand position as a function of separation and delay, but also for the observed variance, or unreliability of the localization errors within and across subjects[7].

Our model predicts that the distribution of localization errors in perceived hand position should be double-peaked for visuo-proprioceptive separations that correspond to an ownership probability around 50%. Our experimental data confirmed this prediction, showing this critical visuo-proprioceptive separation threshold to be approximately 20-25cm when averaged across subjects (Fig. 3B). Importantly, this threshold can be extracted for each individual subject (Fig. 4D), representing significant progress beyond previous experimental settings that relied on averaging across large subject samples[9,35]. This threshold is important for three reasons. First, perceived hand position is strongly and reliably influenced by visuo-tactile delay for separations around the threshold. Second, for separations around the individual threshold for a given subject, our model shows that the subject assigns hand ownership and therefore fuses visual and proprioceptive information for some trials (peak around large localization error values; Fig 3B; Fig 6), but generates zero localization error and refuses ownership for other trials (shoulder around smaller localization error values). Third, for separations significantly below this threshold, our model systematically assigns hand ownership and therefore predicts reliable fusion of visual and proprioceptive information. On the other hand, for separations significantly above the threshold, it predicts a refusal of ownership and relies exclusively on proprioceptive signals. Previous computational models of visuo-proprioceptive integration tasks[21,23] have not incorporated the possibility of assigning body ownership and therefore fail to explain the variability of the localization error and changes in the mean localization error across a large range of visuo-proprioceptive separations.

The unimodal variances (σ_p and σ_v) that we observed are substantially higher than values previously reported using a setup that did not use virtual reality[21]. However, direct comparison between our study and this previous work is difficult due to several methodological differences. First, in the present setup, participants were asked to report their perceived hand position verbally whereas in the previous work, participants were asked to point to their unseen hand using their other hand. Recent research has described differences in hand localization error depending on whether hand motor movements or verbal estimations were employed[12]. Second, the previous work only manipulated visual and proprioceptive cues whereas we estimated the unimodal variances from a dataset where a third sensory modality was implicated (the visuo-tactile stimulus). It remains to be studied whether estimating unimodal variances from data obtained while manipulating additional sensory modalities leads to inflated unimodal estimates. Finally, the present study used a virtual reality based-setup (requiring that participants internally calibrate their hand position in real-space coordinates to those in the virtual space), whereas the data reported by van Beers and colleagues were acquired using robotic stimulation without virtual reality. Though we took great care to align and calibrate the real and virtual environments, the internal mappings for each participant may be imprecise and thereby alter single-sense estimates. Nevertheless, despite the magnitude differences in visual and proprioceptive estimates across these studies, we note that all experimental conditions in our study were carried out under the same experimental settings and thus that all observed statistical differences remain valid.

The present model is reminiscent of a causal inference model that was developed to explain illusory perceptions during the ventriloquist effect, a visuo-auditory illusion[26,36]. Both the RHI and the ventriloquist effect involve misperception of the location of an object. In the former, the object is the position of the subjects' own touched hand in relation to the fake hand that is seen being touched (visuo-proprioceptive conflict). In the latter, it is the position of the ventriloquist's mouth with respect to the seen "speaking" puppet (visuo-auditory conflict). In both illusions, the observer has to decide whether the different sensory signals (visual, proprioceptive for the RHI; auditory, visual for the ventriloquist effect) arise from a single cause (fake hand; "speaking" puppet) or from two separate causes (real or fake hand; ventriloquist or puppet). The parallels between the present

RHI model and previous causal inference models[26,36], and the analogy to earlier models of cue integration and fusion[20–25], suggest that probabilistic inference processes are powerful tools to understand multisensory perception and subjective experience. However, our RHI model additionally includes distinct features of bodily processing related to the self as the misperceived object is part of the observer's body and the occluded real hand gives rise to additional tactile signals that are not available in the ventriloquist effect.

One of the advantages of a computational model for illusory hand ownership is that it might eventually be used to design and extract the optimal stimulation parameters to induce hand ownership for artificial limbs for individual patients. A major goal of neuroprosthetics[37–40] is to design artificial limbs that feel and move, ideally like real limbs. Most research, however, has focused on movement control of artificial limbs[37,38,41], although for a limb to be functionally useful, one must also be able to perceive somatosensory signals from the artificial limb such as touch and proprioception[39,40,42,43]. Recently, artificial limbs have been interfaced to the peripheral nervous system[43,44] or somatosensory cortex[45] in order to provide somatosensory feedback[46–48]. Yet, despite these achievements, many amputees continue to reject current artificial limbs because they rely on visual instead of tactile and proprioceptive signals to interact with objects[49]. The potential importance of inducing body ownership for prosthetic limbs was recently demonstrated by showing that upper limb amputees experience an artificial hand as part of their own body when synchronous touch was applied to an artificial hand and their (occluded) stump[50]. These findings were later extended using a robotic tactile interface allowing for greater stimulus control and repeated conditions[51]. More work is needed to experimentally test some of these issues in patients.

Based on these previous and the present findings, we argue that illusory hand ownership and hand position perception using automated visuo-tactile stimulation on the prosthetic hand and the stump or chest regions containing skin regions with referred hand sensations[43,52] may contribute to the design of artificial limbs that feel like real limbs. Automated visuo-tactile feedback as presented here and interfaced with the skin[51], the peripheral[53], or the central nervous system[40,41] may generate ownership for a prosthesis. We speculate that the combination of visual and somatosensory feedback with ownership automation will boost tactile perception in amputees[54], induce the sensation that the prosthesis is part of the amputee's body, and may decrease the rejection rate of current artificial limbs due to the feeling that they are too heavy and alien.

Materials & Methods

Ethics statement

The studies were undertaken in accordance with the ethical standards as defined in the Declaration of Helsinki and was approved by the local ethics research committee at University of Lausanne.

Participants

18 healthy, right-handed participants (10 females; aged: 24 ± 5.8 years; *mean*±*SD*) were recruited for the main study. In addition, for a pilot study where we investigated illusory touch and hand ownership using the current setup, 11 healthy right-handed participants (4 females, aged: 23.5 ± 4.9 years; *mean*±*SD*) were recruited. All participants reported having normal or corrected-to-normal vision and provided informed consent prior to partaking in the two studies.

Visuo-tactile Stimulation

The general experimental setup is shown in Fig. 1. Tactile stimulation was provided via a set of four button-style vibration motors (Precision Microdrives, London, UK) affixed in a line to the top of the participants' right hand. The vibration motors were 12mm in diameter, with a weight of 1.7g, and vibrated at a maximum of 9000rpm. The motors were programmed to vibrate in sequence to simulate a continuous, stroke-like movement lasting 600ms (100ms per motor and a 50ms pause between motor vibrations). This type of sequence was chosen to automate the stroking patterns that are generally used to manually stroke the participants' hand during the RHI[7,8] and was based on previous work[34] The direction of the stroking sequence was either to the left or to the right (randomized across participants). An inter-stroke interval of 600ms was inserted between strokes to aid in perceiving the sequence of vibrations as a single motion (Fig. 1E, top). Visual stimuli were rendered with XVR (VRMedia, Pisa, Italy) on a Fakespace Wide5 head-mounted display (HMD; Fakespace Labs, Mountain View, CA, USA). The HMD displayed a stereoscopic virtual scene with a tabletop and four spheres on a virtual right hand (Fig. 1B), representing the four vibration motors on the real right hand (Fig. 1A). Visual “vibrations” were represented by animating the virtual motor to jitter and by changing its color from white to red. Synchronization between visual and tactile stimuli was controlled with a custom-made program where the vibration motors were controlled at a precision of approximately 0.1ms. The full experimental loop, including updates to the display had an overall precision of approximately 50ms.

General Procedure

Participants were seated in a fixed chair approximately 10cm in front of a table and saw a 3D virtual hand on the screen of the HMD while the skin of the participants’ real hand (resting on a table in front of them) was stimulated by a set of four small electric vibrators. We addressed the issue of whether the hand is perceived at the position of the real hand, at the position of the virtual hand, or somewhere in between (Fig. 1D). In order to create a close perspective correspondence between the real and virtual scenes, the HMD was individually fit to each subject such that the real and virtual tables were aligned. The head was restrained with a chin rest to stabilize the virtual scene and the HMD fully blocked the participants’ vision of the table, their real hand, and the rest of the room. To eliminate the possibility that participants perceived auditory cues from the vibrators, white noise was provided through a set of headphones. The participants’ right hand (palm down) was placed on the table with the tip of the middle finger at one of three pre-defined proprioceptive hand positions. A virtual hand was projected at different positions on the virtual tabletop (see below). Participants were asked to fixate on the virtual hand and to remain still while visuo-tactile stimulation was administered.

Pilot Experiment

To verify that our experimental setup led to the induction of illusory touch and hand ownership for the virtual hand, we performed a pilot experiment employing a 2x2 factorial design with the factors Stroking and Posture. The pilot study was designed to closely follow well-established RHI experimental protocols[8]. Visuo-tactile stroking was applied for one minute either synchronously (no visuo-tactile delay) or asynchronously (visuo-tactile delay of 400ms). The virtual hand position was fixed 17cm to the left of the real hand[9,13] and its orientation was either congruent or incongruent to the real hand posture[8] (Fig. 2A). Immediately following the visuo-tactile stimulation, participants were given a questionnaire composed of ten questions (7-item Likert scale) to gauge the strength of the RHI in each condition (see Table 1 for the complete list of questions). The order of the experimental conditions was randomized and balanced across subjects.

Main Experiment

Our setup allowed us to control the position of the seen hand and administer computer-controlled, automated visuo-tactile stimulation across a large range of stroking durations as well as visuo-tactile delays (from near-synchronous to many different levels of asynchronous stimulation). We tested perceived hand position following the modulation of three stimulation parameters: (1) visuo-tactile delay, (2) duration of stimulation, and (3) visuo-proprioceptive separation. We adopted a continuous experimental design, in which each trial was defined by fixing a value for these three parameters. All parameters were selected randomly on a trial-by-trial basis with uniform probability. The visuo-proprioceptive separation ranged from 0 to 40cm, trial duration from 5 to 90 strokes (with 1 stroke + inter-stroke interval = 1.2s), and visuo-tactile delay from 0 to 0.8s. Throughout the experiment, the proprioceptive hand position was fixed by the experimenter and changed each five trials by displacing the subjects' right hand to one of three randomly selected positions (17, 26, or 35cm to the right of the body midline). These values were determined from pilot studies with the aim of focusing the collected data on regions where the subjects were found to be more sensitive.

Each trial involved a visuo-tactile stimulation period followed by a darkened virtual scene (1s) without the virtual table and the virtual hand. Next, the virtual scene reappeared with a virtual ruler (with centimeter precision) spanning the virtual table (Fig. 1C). Participants were instructed to verbally provide the label of the tick on the virtual ruler corresponding to the perceived position of the tip of the real right hand's middle finger[9]. Response times of varied from 2s to 5s. Labels for the ticks on the virtual ruler were randomly selected on a per-trial basis to diminish biases.

Of the eighteen participants recruited for this experiment, fifteen performed 62 ± 4 ($mean \pm SD$) trials and the remaining three participants performed 163 ± 37 trials. We recorded a larger number of trials for these three participants in order to contrast our proposed model with competing models on an individual subject basis. For the group analysis, a total of 1341 trials were pooled across all eighteen participants.

Response Model and Parameter Optimization

Our perception model is defined by the distribution $p(Q|X_v,X_p,Z)$, that is, the distribution of perceived hand positions Q given the multisensory input as described by Eq. (6). We additionally specify how a subject makes their decision when reporting their perceived hand position. For this, we assume that the subject draws a single sample from the distribution $p(Q|X_v,X_p,Z)$ and reports the resulting Q-value.

In general, a given subject's decision-making strategy depends on their individual cost function[55]. For instance, if the cost function of the subject is based on the mean squared error (MSE), the optimal policy consists of reporting the *posterior mean* of the belief state. However, our results suggest that a mixture of *two* Gaussians provides a better explanation than a single Gaussian for the localization errors at visuo-proprioceptive separations in the range 20-30cm (as measured by the *Bayesian Information Criterion*, see *Significance Tests* below). Alternatively, if the cost function is taken to be a Dirac-delta on the correct answer, the optimal policy is to report the *maximum a posteriori* (MAP). As it has been shown that subjects may use approximations to the MAP estimate based on few samples from the posterior[56], our response model can be interpreted as an approximation to the MAP strategy based on a single sample from the posterior distribution of perceived hand positions.

In order to fit a perceptual MAP model to behavioral data we would have to marginalize the MAP of the posterior distribution defined in Eq. (6) with respect to the perceptual noise and then maximize the likelihood of the resulting distribution of MAP-responses to the data[19]. This procedure cannot be easily applied to our model due to the complicated dependency of the MAP on the single percepts Eq. (6). Rather, we adopted the following approximation: we set the percepts X_v, X_p and Z to their real values and took the participant's response as a sample from the posterior distribution defined by Eq. (6). Consistency can be checked by performing a quadratic expansion of log-likelihood of the Gaussian mixture Eq. (6) around its closest mode given the fixed percepts. Note also that this approximation gives the same result as the correct, but intractable fitting procedure (up to quadratic order).

The four free parameters $\{\sigma_Z, \sigma_v, \sigma_p, \eta\}$ in Eq. (6) were fit using the following step-by-step procedure for both the group and individual data sets:

(i) For visuo-proprioceptive separations larger than 30cm, we first measured the standard deviation of localization errors from the data. This defined the parameter σ_p, which remained fixed throughout the rest of the fitting procedure.

(ii) For visuo-proprioceptive separations smaller than 10cm, we measured the standard deviation of localization errors from the data. This defined the standard deviation $\sigma_{\text{fused}} = 1/\sqrt{\frac{1}{\sigma_v^2}+\frac{1}{\sigma_p^2}}$ under the assumption of fusion of two independent Gaussian signal channels[21], represented by vision and proprioception in our setup. Since (i) provided the value of the parameter σ_p, we could deduce σ_v, which also remained fixed during subsequent steps.

(iii) In the case that the delay parameter σ_Z was irrelevant for a given fitting analysis (*i.e.* for Fig. 3B and Fig. 6), we only considered one free parameter (η), which accounted for the Bayesian prior in the full model described in Eq. (6). We determined a distribution $p(\eta \mid Data)$ using a Markov Chain Monte Carlo (MCMC) procedure[57], under the assumption of a flat prior $p_0(\eta \mid Data)$ on a finite interval ($\eta \in [0, 4000]$). We then computed 10,000 steps of MCMC resulting in 10,000 "particles" for the parameter η where a specific value η_k appears with a probability $p(\eta \mid Data)$. This probability is itself proportional to $p(Data \mid \eta, \sigma_v, \sigma_p)$, where '*Data*' represents the reported localization errors Q for a given X_p and X_v for the real and virtual hands, respectively, and η, σ_p, σ_v defines the set of free parameters of the model described in Eq. (6). We proceeded analogously for all fits that consider visuo-tactile delay, except that we treated the two free parameters $\theta = \{\sigma_Z, \eta\}$ in parallel with a flat prior over σ_z in the interval 0 - 1s.

(iv) The full model from Eq. (6) was evaluated at the data points from each range (*e.g.*, visuo-proprioceptive separations between 10cm and 20cm), by adding the contributions of the 10,000 choices of θ:

$$p(Q \mid X_v, X_p) = \int d\theta p(Q \mid \theta, X_v, X_p) p(\theta \mid Data) \approx \frac{1}{S} \sum_{s=1}^{S} p(Q \mid \theta_s, X_v, X_p) \quad (7)$$

We analogously computed the inferred ownership $(1-\alpha) = p(own = 1 \mid stimulus)$ as a function of visuo-proprioceptive separation, summing over 10,000 MCMC particles $\theta = \{\sigma_z, \eta\}$. Individual ownership thresholds were determined by computing the ownership curve $(1-\alpha) = p(own = 1 \mid stimulus)$ for each of the 10,000 particles and then extracting the value of visuo-proprioceptive separation (threshold) for which the ownership curve passes through 0.5.

Localization distribution peak tests

In order to test whether a Gaussian Mixture Model with one or two components better explained the localization errors in the visuo-proprioceptive separation interval of 20-30cm, we fit both models and compared them using the *Bayesian Information Criterion*[58]. The BIC is defined in terms of the model's log-likelihood $p(Data \mid \theta)$, the number of free parameters in the model d, and the amount of data seen by the model N as: $BIC = -2 \ln p(Data \mid \theta) + d \ln N$. Models with smaller BIC values indicate better, more parsimonious descriptions of the data. The *Bayes Factor* $BF = p(M_1 \mid Data) / p(M_2 \mid Data)$ between two models M_1 and M_2 can be approximated with their BIC values as: $BF \approx \exp\left[\frac{1}{2}(BIC_2 - BIC_1)\right]$. Finally, preference between models M1 and M2 is made by choosing M1 if BF > 1 or M2 if BF < 1.

Model comparisons

Model comparison between our model and a two-sense linear model[20–25] was performed using the *Deviation Information Criterion* (DIC), defined as:

$$-2[2\langle \ln p(Data \mid \theta) \rangle_{p(\theta|Data)} - \ln p(Data \mid \langle \theta \rangle_{p(\theta|Data)})]$$

where $\langle f(\bullet)\rangle_{p(\bullet)}$ is the expectation of function $f(x)$ under density $p(x)$ and θ is the set of free parameters of the model[58]. DIC values are estimated using a MCMC procedure that produces samples from $p(\theta \mid Data)$ to estimate DIC values (10^4 samples, using 10^3 burning steps and a thinning of 10 steps). Models with smaller DIC are preferred to models with larger DIC values. Note that for the model comparisons in our per-subject analyses, one subject's DIC estimator did not converge, resulting in individual model comparisons for 17 subjects.

Significance Tests

Each visuo-proprioceptive separation range defined by our splitting procedure consisted of a finite number of data samples, *N*. We computed whether a sample of *N* data points drawn from the model statistically differed from the observed *N* experimental data points. 4000 samples were drawn from the model's predictive distribution as defined in Eq. (7) for each visuo-proprioceptive separation range (*e.g.* 10-20cm, 20-30cm, *etc.*). These samples were then compared to the experimentally observed distribution of localization errors in the same range. Finally, we created histograms (bin size=3cm) and performed a two-sample χ^2-test.

REFERENCES

1. Berllucchi, G. & Aglioti, S. The body in the brain: neural bases of corporeal awareness. *Trends in Neurosci.* **20,** 560–564 (1997).

2. Botvinick, M. Probing the Neural Basis of Body Ownership. (2004).

3. Damasio, A. *The Feeling of What Happens: Body and Emotion in the Making of Consciousness. New:Harcourt Brace* (New York: Harcourt Brace, 2000).

4. Jeannerod, M. *Motor cognition: What actions tell the self.* (Oxford University Press, 2006).

5. Jeannerod, M. Being oneself. *J. Physiol. Paris* **101,** 161–8 (2007).

6. Vogeley, K. & Fink, G. R. Neural correlates of the first-person-perspective. *Trends Cogn. Sci.* **7,** 38–42 (2003).

7. Botvinick, M. & Cohen, J. Rubber hands 'feel' touch that eyes see. *Nature* **391,** 756 (1998).

8. Ehrsson, H. H., Spence, C. & Passingham, R. E. That's my hand! Activity in premotor cortex reflects feeling of ownership of a limb. *Science* **305,** 875–877 (2004).

9. Tsakiris, M. & Haggard, P. The rubber hand illusion revisited: visuotactile integration and self-attribution. *J. Exp. Psychol. Hum. Percept. Perform.* **31,** 80–91 (2005).

10. Armel, K. C. & Ramachandran, V. S. Projecting sensations to external objects: evidence from skin conductance response. *Proc. Biol. Sci.* **270,** 1499–1506 (2003).

11. Moseley, G. L. *et al.* Psychologically induced cooling of a specific body part caused by the illusory ownership of an artificial counterpart. *Proc. Natl. Acad. Sci. U. S. A.* **105,** 13169–13173 (2008).

12. Kammers, M. P. M., de Vignemont, F., Verhagen, L. & Dijkerman, H. C. The rubber hand illusion in action. *Neuropsychologia* **47,** 204–211 (2009).

13. Lloyd, D. M. Spatial limits on referred touch to an alien limb may reflect boundaries of visuo-tactile peripersonal space surrounding the hand. *Brain Cogn.* **64,** 104–109 (2007).

14. Costantini, M. & Haggard, P. The rubber hand illusion: sensitivity and reference frame for body ownership. *Conscious. Cogn.* **16,** 229–240 (2007).

15. Kammers, M., Mulder, J., de Vignemont, F. & Dijkerman, H. The weight of representing the body: addressing the potentially indefinite number of body representations in healthy individuals. *Exp Brain Res* **204,** 333–342 (2009).

16. Tsakiris, M. My body in the brain: a neurocognitive model of body-ownership. *Neuropsychologia* **48,** 703–712 (2010).

17. Makin, T. R., Holmes, N. P. & Ehrsson, H. H. On the other hand: dummy hands and peripersonal space. *Behav. Brain Res.* **191,** 1–10 (2008).

18. Slater, M., Perez-Marcos, D., Ehrsson, H. H. & Sanchez-Vives, M. V. Towards a digital body: the virtual arm illusion. *Front. Hum. Neurosci.* **2,** 6 (2008).

19. Stocker, A. A. & Simoncelli, E. P. Noise characteristics and prior expectations in human visual speed perception. *Nat. Neurosci.* **9,** 578–85 (2006).

20. Cheng, K., Shettleworth, S. J., Huttenlocher, J. & Rieser, J. J. Bayesian integration of spatial information. *Psychol. Bull.* **133,** 625–637 (2007).

21. van Beers, R. J., Sittig, a C. & Gon, J. J. Integration of proprioceptive and visual position-information: An experimentally supported model. *J. Neurophysiol.* **81,** 1355–1364 (1999).

22. Knill, D. C. Robust cue integration: a Bayesian model and evidence from cue-conflict studies with stereoscopic and figure cues to slant. *J Vis* **7,** 1–24 (2007).

23. Ernst, M. O. & Banks, M. S. Humans integrate visual and haptic information in a statistically optimal fashion. *Nature* **415,** 429–433 (2002).

24. Prsa, M., Gale, S. & Blanke, O. Self-motion leads to mandatory cue fusion across sensory modalities. *J. Neurophysiol.* **108,** 2282–91 (2012).

25. Ernst, M. O. & Bülthoff, H. H. Merging the senses into a robust percept. *Trends Cogn. Sci.* **8,** 162–169 (2004).

26. Körding, K. P. *et al.* Causal inference in multisensory perception. *PLoS One* **2,** e943

(2007).

27. Wozny, D. R. & Shams, L. Computational characterization of visually induced auditory spatial adaptation. *Front Integr Neurosci* **5,** 75 (2011).

28. Beierholm, U. R., Körding, K. P., Shams, L. & Ma, W. J. Comparing Bayesian models for multisensory cue combination without mandatory integration. in *Advances in Neural Information Processing Systems* **20,** 81–88 (2008).

29. Colonius, H. & Diederich, A. The optimal time window of visual-auditory integration: a reaction time analysis. *Front. Integr. Neurosci.* **4,** 11 (2010).

30. Di Luca, M., Machulla, T.-K. & Ernst, M. O. Recalibration of multisensory simultaneity: cross-modal transfer coincides with a change in perceptual latency. *J. Vis.* **9,** 7.1–16 (2009).

31. Roseboom, W., Nishida, S. & Arnold, D. H. The sliding window of audio-visual simultaneity. *J. Vis.* **9,** 4.1–8 (2009).

32. Hägni, K. *et al.* Observing virtual arms that you imagine are yours increases the galvanic skin response to an unexpected threat. *PLoS One* **3,** e3082 (2008).

33. IJsselsteijn, W. a, de Kort, Y. a. W. & Haans, A. Hand I See Before Me? The Rubber Hand Illusion in Reality, Virtual Reality, and Mixed Reality. *Presence Teleoperators Virtual Environ.* **15,** 455–464 (2006).

34. Evans, N. & Blanke, O. Shared electrophysiology mechanisms of body ownership and motor imagery. *Neuroimage* (2012). doi:10.1016/j.neuroimage.2012.09.027

35. Rohde, M., Luca, M. Di & Ernst, M. O. The Rubber Hand Illusion : Feeling of Ownership and Proprioceptive Drift Do Not Go Hand in Hand. **6,** (2011).

36. Beierholm, U. R., Quartz, S. R. & Shams, L. The ventriloquist illusion as an optimal percept. *J Vis* **5,** 647–647 (2010).

37. Carmena, J. M. *et al.* Learning to control a brain-machine interface for reaching and grasping by primates. *PLoS Biol.* **1,** E42 (2003).

38. Schwartz, A. B. Cortical neural prosthetics. *Annu Rev Neurosci* **27,** 487–507 (2004).

39. Navarro, X. *et al.* A critical review of interfaces with the peripheral nervous system for the control of neuroprostheses and hybrid bionic systems. *J. Peripher. Nerv. Syst.* **10,** 229–258 (2005).

40. Lebedev, M. a & Nicolelis, M. a L. Brain-machine interfaces: past, present and future. *Trends Neurosci.* **29,** 536–46 (2006).

41. Hochberg, L. R. *et al.* Neuronal ensemble control of prosthetic devices by a human with tetraplegia. *Nature* **442,** 164–171 (2006).

42. London, B. M., Jordan, L. R., Jackson, C. R. & Miller, L. E. Electrical stimulation of the proprioceptive cortex (area 3a) used to instruct a behaving monkey. *IEEE Trans. Neural Syst. Rehabil. Eng.* **16,** 32–6 (2008).

43. Kuiken, T. A. *et al.* Targeted reinnervation for enhanced prosthetic arm function in a woman with a proximal amputation: a case study. *Lancet* **369,** 371–80 (2007).

44. Colgate, J. E., Santos-Munne, J. J., Makhlin, A. & Peshkin, M. A. On the Design of Miniature Haptic Devices for Upper Extremity Prosthetics. *IEEE/ASME Trans. Mechatronics* **15,** 27–39 (2010).

45. O'Doherty, J. E. *et al.* Active tactile exploration using a brain–machine–brain interface. *Nature* **479,** 228–231 (2011).

46. Dhillon, G. S. & Horch, K. W. Direct neural sensory feedback and control of a prosthetic arm. *IEEE Trans. Neural Syst. Rehabil. Eng.* **13,** 468–72 (2005).

47. Schultz, A. E., Marasco, P. D. & Kuiken, T. A. Vibrotactile detection thresholds for chest skin of amputees following targeted reinnervation surgery. *Brain Res* **1251,** 121–9 (2009).

48. Sensinger, J. W., Schultz, A. E. & Kuiken, T. A. Examination of force discrimination in human upper limb amputees with reinnervated limb sensation following peripheral nerve transfer. *IEEE Trans. Neural Syst. Rehabil. Eng.* **17,** 438–44 (2009).

49. Wright, T. W., Hagen, A. D. & Wood, M. B. Prosthetic usage in major upper extremity amputations. *J Hand Surg Am* **20,** 619–22 (1995).

50. Ehrsson, H. H. *et al.* Upper limb amputees can be induced to experience a rubber hand as their own. *Brain* **131,** 3443–3452 (2008).

51. Marasco, P. D., Kim, K., Colgate, J. E., Peshkin, M. A. & Kuiken, T. A. Robotic touch shifts perception of embodiment to a prosthesis in targeted reinnervation amputees. *Brain* **134,** 747–58 (2011).

52. Ramachandran, V. S., Rogers-Ramachandran, D. & Stewart, M. Perceptual correlates of massive cortical reorganization. *Science* **258,** 1159–60 (1992).

53. Micera, S., Carpaneto, J. & Raspopovic, S. Control of hand prostheses using peripheral information. *IEEE Rev Biomed Eng* **3,** 48–68 (2010).

54. Zopf, R., Harris, J. A. & Williams, M. A. The influence of body-ownership cues on tactile sensitivity. *Cogn Neurosci* **2,** 147–154 (2011).

55. Paninski, L. Nonparametric inference of prior probabilities from Bayes-optimal behavior. in *Advances in Neural Information Processing Systems 18* (2005).

56. Vul, E., Goodman, N., Griffiths, T. & Tenenbaum, J. One and Done? Optimal Decisions From Very Few Samples. in *Proceedings of the 31st Annual Conference of the Cognitive Science Society* (2009).

57. Andrieu, C., De Freitas, N., Doucet, A. & Jordan, M. I. An introduction to MCMC for machine learning. *Mach Learn* **50,** 5–43 (2003).

58. Spiegelhalter, D. J., Best, N. G., Carlin, B. P. & Linde, A. Van Der. Bayesian measures of model complexity and fit. **64,** 583–639 (2010).

59. Sanchez-Vives, M. V, Spanlang, B., Frisoli, A., Bergamasco, M. & Slater, M. Virtual hand illusion induced by visuomotor correlations. *PLoS One* **5,** e10381 (2010).

60. Tsakiris, M., Costantini, M. & Haggard, P. The role of the right temporo-parietal junction in maintaining a coherent sense of one's body. *Neuropsychologia* **46,** 3014–8 (2008).

61. Hohwy, J. & Paton, B. Explaining away the body: experiences of supernaturally caused touch and touch on non-hand objects within the rubber hand illusion. *PLoS One* **5,** e9416 (2010).

62. Durgin, F. H., Evans, L., Dunphy, N., Klostermann, S. & Simmons, K. Rubber hands feel the touch of light. *Psychol. Sci. a J. Am. Psychol. Soc. / APS* **18,** 152–157 (2007).

63. Ehrsson, H. H., Holmes, N. P. & Passingham, R. E. Touching a rubber hand: feeling of body ownership is associated with activity in multisensory brain areas. *J. Neurosci.* **25,**

10564–10573 (2005).

64. Morgan, H. L. *et al.* Exploring the impact of ketamine on the experience of illusory body ownership. *Biol. Psychiatry* **69,** 35–41 (2011).

65. Shimada, S., Fukuda, K. & Hiraki, K. Rubber hand illusion under delayed visual feedback. *PLoS One* **4,** e6185 (2009).

66. Dummer, T., Picot-Annand, A., Neal, T. & Moore, C. Movement and the rubber hand illusion. *Perception* **38,** 271–280 (2009).

67. Ocklenburg, S., Rüther, N., Peterburs, J., Pinnow, M. & Güntürkün, O. Laterality in the rubber hand illusion. *Laterality* **16,** 174–187 (2011).

68. Schütz-Bosbach, S., Mancini, B., Aglioti, S. M. & Haggard, P. Self and other in the human motor system. *Curr. Biol.* **16,** 1830–1834 (2006).

69. Zopf, R., Truong, S., Finkbeiner, M., Friedman, J. & Williams, M. A. Viewing and feeling touch modulates hand position for reaching. *Neuropsychologia* **49,** 1287–1293 (2011).

70. Tsakiris, M., Hesse, M. D., Boy, C., Haggard, P. & Fink, G. R. Neural signatures of body ownership: a sensory network for bodily self-consciousness. *Cereb. Cortex* **17,** 2235–2244 (2007).

71. Lopez, C., Lenggenhager, B. & Blanke, O. How vestibular stimulation interacts with illusory hand ownership. *Conscious. Cogn.* **19,** 33–47 (2010).

Acknowledgments

Funding: This work was partially supported by the VERE project: FP7-ICT-2009-5 — Project 257695 and the Bertarelli Foundation (N.E., O.B.), European Science Foundation, Systems X (D.R.), Swiss National Science Foundation (W.G.)

Author Contributions

N.E., D.R. and O.B. were responsible for the study conception; N.E. and D.R. implemented and executed the experiments; N.E. and D.R. performed the data analyses; N.E., D.R., W.G, and O.B contributed to the interpretation, methodology and manuscript preparation.

Additional Information

Supplementary information accompanies this paper

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

Figure Legends

Figure 1: Automated experimental setup to induce hand ownership for a virtual hand. (**A**) The participant's right hand is stimulated by small vibrotactile motors while they wear a head-mounted display that occludes vision of their hand. (**B**) Subjects see a virtual hand with virtual representations of the vibrotactile motors on their real hand. (**C**) Following visuo-tactile stimulation, a virtual ruler was presented for participants to report perceived hand position. (**D**) Illustration of a case where visual, proprioceptive, and perceived hand positions differ from one another. (**E**) Temporal sequence of visuo-tactile stimulation on the hand for a representative trial. Vibrotactile motors "stroked" the hand with a sequence of four vibrations either in synchrony with a visual counterpart (synchronous) or with an injected delay Z (asynchronous). Individual motor colors added for graphical representation only. Note that all virtual scenes are shown in a monocular view though participants saw stereoscopic scenes.

Figure 2: Experimental design and self report scores for the induction of illusory ownership. **(A)** Hand positions for the real hand and the virtual hand in congruent and incongruent posture conditions of the pilot study. **(B)** Illusory hand ownership scores (item Q3 from Table 1) and scores for a control question (item Q5). Scores from the 7-item Likert scale were normalized between -3 and 3. Post-hoc Wilcoxon matched-pair tests revealed a body-selective, synchrony-dependent modulation of illusory hand ownership ($p = 0.005$). Importantly, this result was absent for the control conditions with incongruent visual hand postures, as well as for the control question (all $p > 0.05$).

Figure 3: Empirical data show that perceived hand position depends on visuo-proprioceptive separation and visuo-tactile delay.

(A) Each point represents an individual trial with a given visuo-proprioceptive separation (y-axis) and the localization error as reported by the participants (x-axis; localization error = +/- absolute value of the ’reported hand position’ - ‘real hand position’ where positive localization errors indicate a shift of perceived hand position towards the virtual hand). To avoid overlap of trials with identical parameter settings, a small Gaussian jitter ($\mu = 0cm$ and $\sigma = 0.2cm$) was added for visualization purposes. Small visuo-tactile delays (0s < Z < 0.2s) are shown in red, large visuo-tactile delays (0.8s > Z > 0.6s) are in blue. Colored vertical bars (red: synchronous stroking; blue: asynchronous; shaded regions: standard error) indicate mean localization error averaged over all trials within a visuo-proprioceptive separation range of 0-10cm, 10-20cm, 20-30cm, and 30-40cm (from bottom to top). The vertical solid line at zero-localization error indicates perceived hand positions that are independent from visual cues (and thus rely exclusively on proprioceptive cues). The dashed diagonal line indicates responses based exclusively on visual cues. **(B)** Histograms of the experimental distribution of localization errors for each of the four visuo-proprioceptive ranges in (A). The solid black curve overlaying the histograms indicates our model’s prediction of the distribution. Note that the broad distribution for visuo-proprioceptive separations >20cm that well captured by the model. The vertical solid line and dashed diagonal lines represent the visual and proprioceptive dominated responses, as in (A). For separations of 0 – 30cm, the model does not significantly differ from the empirical data (*0-10cm*: df=112, p=0.18, N=97; *10-20cm*: df=180, p=0.03, N=105; *20-30cm*: df=312, p=0.16, N=222; *30-40cm*: df=234, p=0.1, N=100; χ^2-test).

Figure 4: Three-sense Bayesian model of the rubber hand illusion.

(A) Our generative model of perception depicted as a standard causal diagram. Variables (perceived hand position, ownership, visual hand position, *etc*…) are represented as circles (nodes), and conditional dependencies between variables are indicated by arrows connecting the nodes. The conditional distribution of sensory variables (proprioceptive position, visual position, visuo-tactile delay) indicated in the lower part of the diagram depends on whether the visual hand is the subject's real hand or not (ownership node, *upper right*) and on the real hand position (Q node, *upper left*). (**B**) Illustration of a hand position inference (*top left*) and ownership probability (*top right*) from sensory variables X_p, X_v, and Z. **(C)** Model-predicted ownership probability for the virtual hand as a function of visuo-proprioceptive separation for asynchronous ($Z = 0.7$s, blue) and synchronous ($Z = 0.1$s, red) visuo-tactile stimulation. Ownership thresholds (dashed arrows) are defined as the separation that leads to a probability of ownership of 0.5 (dashed horizontal line). (**D**) Ownership thresholds (mean and standard deviation) for asynchronous ($Z = 0.7$s, blue) and synchronous ($Z = 0.1$s, red) visuo-tactile stimulation for individual participants, as extracted by the model. Inset: Model predicted probability of ownership as in (C) for three individual participants.

Figure 5: Comparison of our proposed model with a linear model that always fuses vision and proprioception.

(A) Relative model performance between our model and a linear model (lower values indicate *higher performance* of our model relative to the linear model), where $\text{performance} = (DIC_{\text{proposed}} - DIC_{\text{linear}}) / \sigma_{\text{diff}}$ and σ_{diff} is the estimated standard deviation of the difference between DIC scores (see *Methods*). Relative performance measures are shown for data collapsed across participants (group data) and for each individual subject. *NA* indicates subject data that led to ill-defined DIC scores. We found our model to outperform the linear model (negative relative performance values) in the large majority of individual subjects and at the group level. **(B)** Empirical localization error measurements for an individual participant (black dots; subject 18) with the model prediction for synchronous (red curve) and asynchronous (blue curve) trials. Shaded regions indicate the predicted standard deviation of the mean. **(C)** Linear model predictions for synchronous and asynchronous trials on the same participant's data. Note that without the third sensory modality (tactile), no distinction is made between predictions for synchronous and asynchronous trials.

Figure 6: Single subject detailed analysis

Histograms of the empirical distribution of localization errors and the associated model prediction (as in Fig. 3B) for an individual participant (subject 18). For near-synchronous stimulation and in the visuo-proprioceptive separation range of 0 – 30cm, the model did not significantly differ from the observed data.

Figure 7: Effect of trial (stroking) duration on perceived hand position

Long trials (>45 strokes, dark blue) vs. short trials (<40 strokes, light blue) for trials with asynchronous stroking**.** Each point represents an individual trial with the visuo-proprioceptive separation on the y-axis and the induced localization error as reported by the participants on the x-axis (as in Fig. 3A). Colored vertical bars indicate mean localization error averaged over all trials within visuo-proprioceptive ranges 0-10cm, 10-20cm, 20-30cm, 30-40cm (from bottom to top; shaded region: standard deviation). For separations of 20-30cm, long-duration, asynchronous stroking led to significantly smaller localization error ($p = 0.0032$, *two-tailed T-test*) than for short-duration stroking trials.

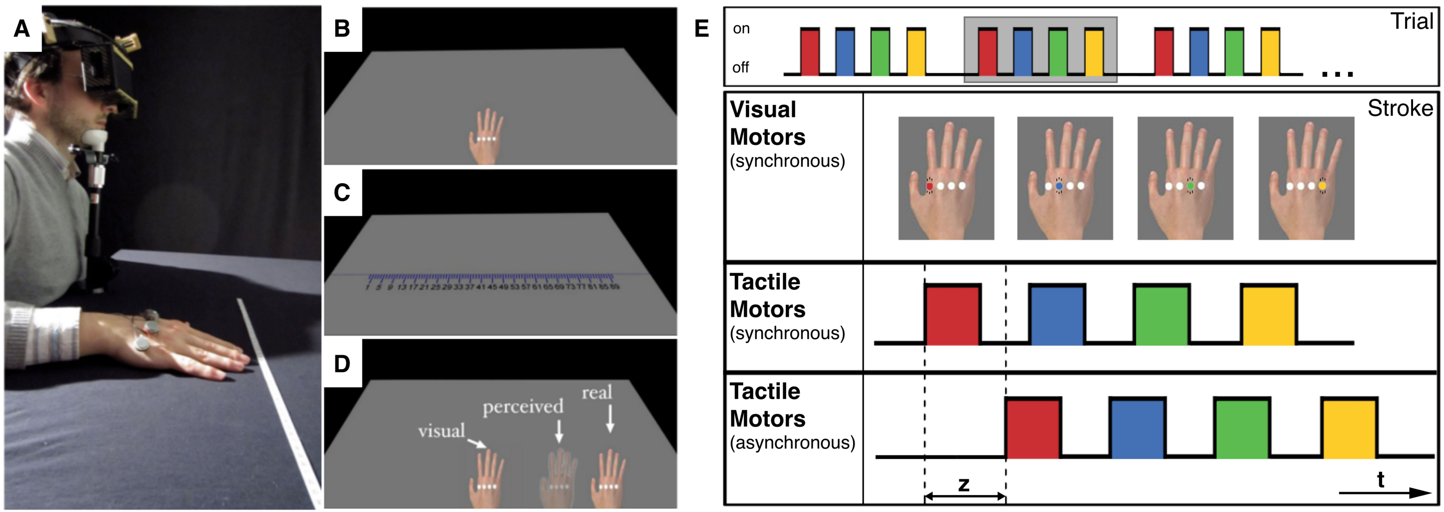
Tables

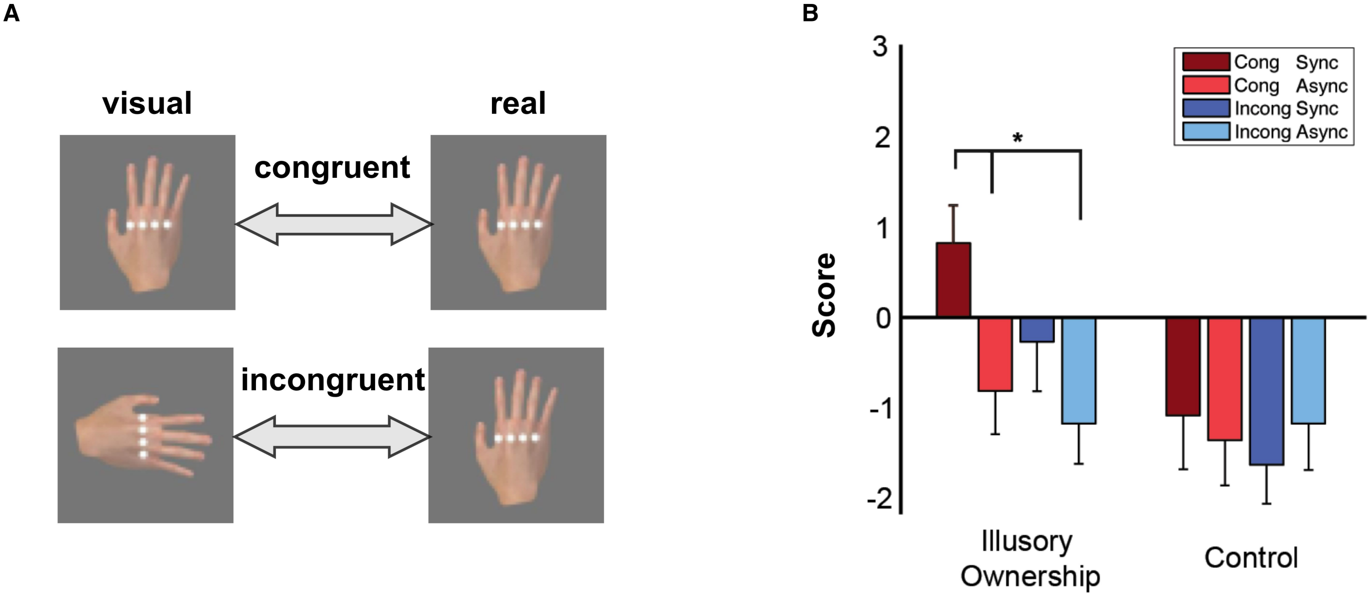
Question	Synchronous Congruent	Asynchronous Congruent	Synchronous Incongruent	Asynchronous Incongruent
Q1: It seemed as if I were feeling the vibrations in the location where I saw the virtual hand being vibrated.	2.09 ± 0.37	-0.82 ± 0.57	2.23 ± 0.29	-0.55 ± 0.57
Q2: It seemed as though the virtual vibrations I felt were caused by the vibrations I saw on the virtual hand.	1.27 ± 0.39	-0.91 ± 0.51	0.55 ± 0.59	-0.73 ± 0.55
Q3: I felt as if the virtual hand were my own hand.	0.82 ± 0.44	-0.82 ± 0.48	-0.27 ± 0.55	-1.18 ± 0.44
Q4: It felt as if my (real) hand was moving/drifting towards the virtual hand's position.	-1.09 ± 0.45	-1.55 ± 0.45	-1.18 ± 0.49	-1.64 ± 0.41
Q5: It seemed as if I might have more than one right hand or arm.	-1.09 ± 0.60	-1.36 ± 0.50	-1.64 ± 0.43	-1.18 ± 0.51
Q6: It seemed as if the vibrations I felt originated from somewhere between my own hand and the virtual hand.	-0.91 ± 0.51	-0.91 ± 0.65	-1.27 ± 0.58	-0.27 ± 0.59
Q7: It felt as if my (real) hand was becoming 'virtual'.	0.64 ± 0.45	-0.82 ± 0.60	-0.36 ± 0.58	-0.64 ± 0.53
Q8: It appeared (visually) as if the virtual hand was drifting towards my (real) hand.	-0.55 ± 0.39	-1.82 ± 0.44	-1.45 ± 0.45	-0.55 ± 0.57
Q9: The virtual hand began to resemble my own (real) hand in terms of shape, skin tone, freckles, or some other visual feature.	-1.0 ± 0.53	-1.18 ± 0.49	-0.91 ± 0.49	-1.73 ± 0.32
Q10: I felt as if I were fully immersed in the virtual environment.	0.27 ± 0.47	-0.64 ± 0.49	-0.09 ± 0.55	-0.36 ± 0.62

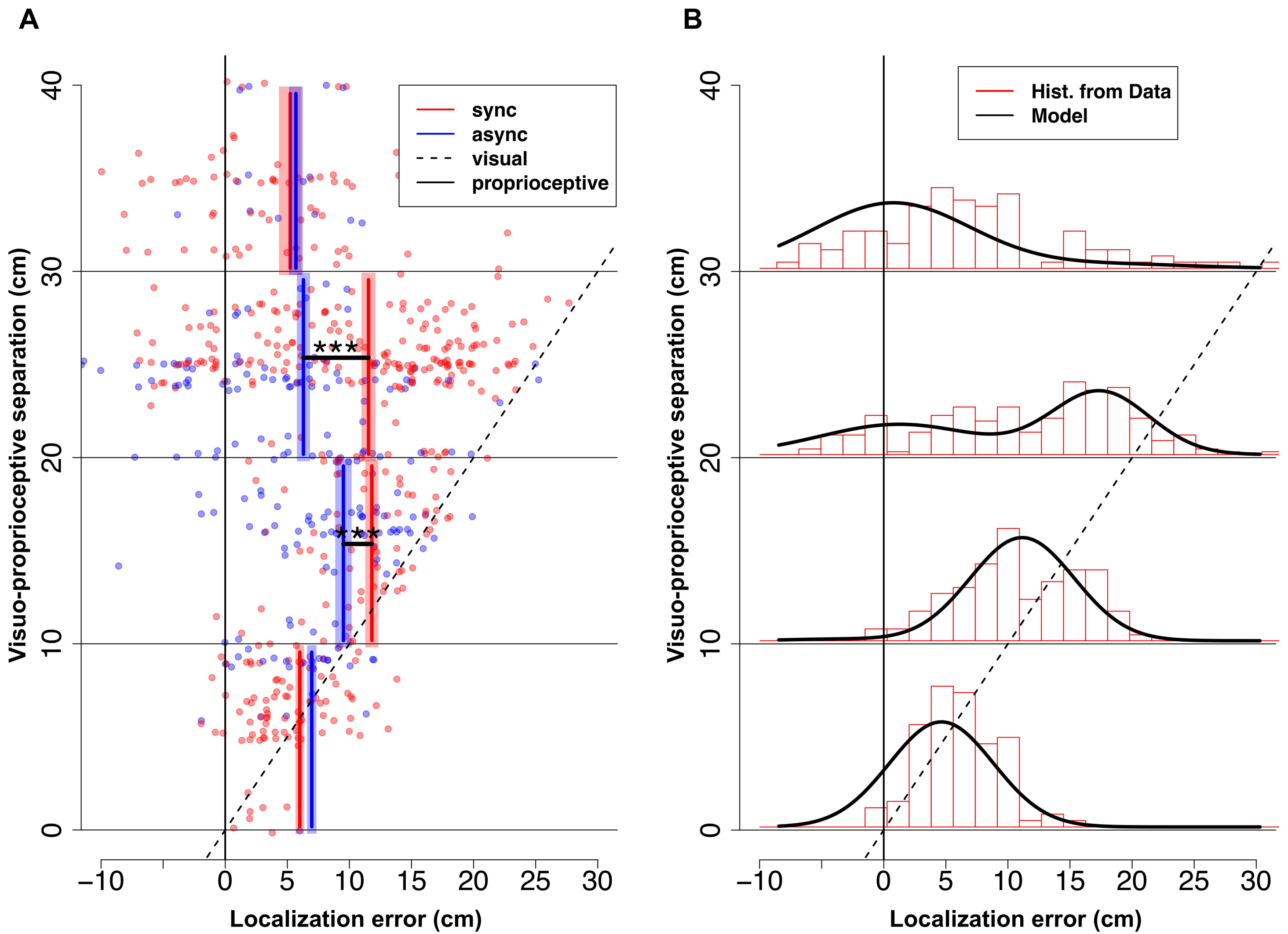
Table 1. Questionnaire scores from the illusory ownership pilot experiment. Scores correspond to a 7-item Likert scale normalized between -3 and 3. Questions were adapted from classical questionnaires gauging illusory effects during the RHI [7,18].

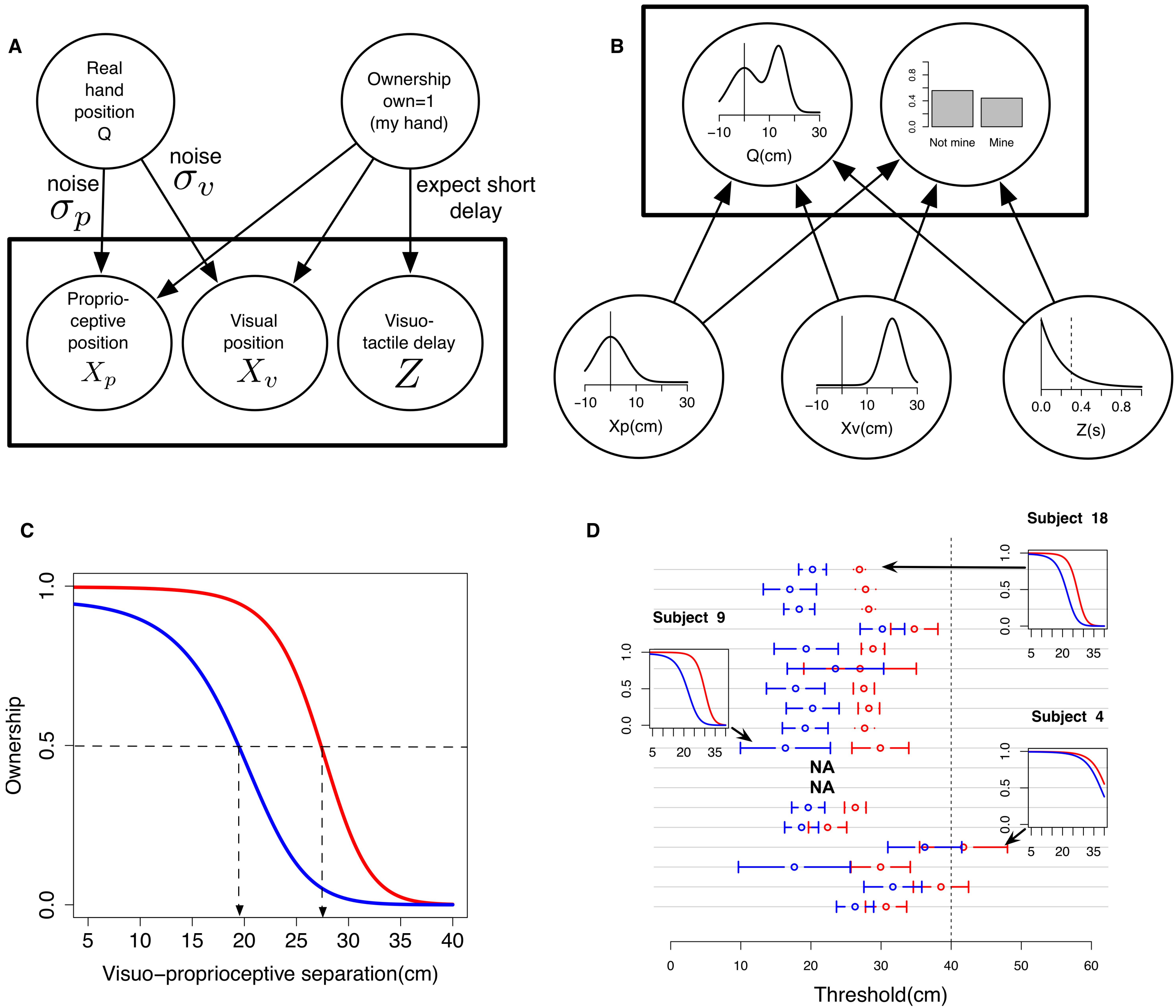
Paper	Duration of visuo-tactile stimulation (s)	Visuo-tactile delay (ms)	Visuo-prorioceptive separation (cm)
Present work	{15.7,31.5,56.7 85.0,113.4,141.7 170.1,226.8, 283.5}	[0,800]	[0,35]
Botvinick et al. (1998) [7]	NA	NA	NA
Tsakiris & Haggard (2005) [9]	240	[500,1000]	17.5
Moseley et al. (2008) [11]	450	NA	NA
Kammers et al. (2009)[12]	90	NA	15
Lloyd (2007) [13]	60	NA	{17.5,27.5,37.5,47.5,57.5, 67.5}
Ehrsson et al. (2008) [50]	60	NA	26
Slater et al. (2008) [18]	300	NA	20
Rohde et al. (2011) [35]	{420,[0,10,40,120]}	NA	17
Sanchez-Vives et al. (2010) [59]	NA	NA	20
Tsakiris et al. (2008) [60]	2.3	NA	17.5
IJsselsteijn et al., (2006) [33]	450	NA	30
Hohwy et al. (2010) [61]	{10,30,60}	[500,1000]	NA
Durgin et al. (2007) [62]	120	NA	15
Ehrsson et al. (2005) [63]	{30,60}	NA	15
Morgan et al. (2011) [64]	300	NA	15
Shimada et al. (2009) [65]	180	[100,600]	15
Dummer et al. (2009) [66]	600	NA	NA
Ocklenburg et al. (2010) [67]	180	NA	17.5
Schütz-Bosbach et al. (2006) [68]	NA	NA	NA
Zopf et al. (2011) [69]	120	NA	20
Tsakiris et al. (2007) [70]	125	[500,1000]	15
Lopez et al. (2010) [71]	60	NA	24.5
Mean ± SD	174 ± 168	650 ± 200	25 ± 14
{Min, Max}	{2.3,600}	{100,1000}	{15,67.5}

Table 2. Comprehensive summary of experimental parameter ranges used in previously reported RHI setups including the present work. Note that the range of parameters used in the present study encompasses most of the previous setups. *NA* indicates that the corresponding information was not provided in the article or was unclear from its methods description. The bottom row summarizes the distribution of the parametric ranges (ignoring *NA* values).

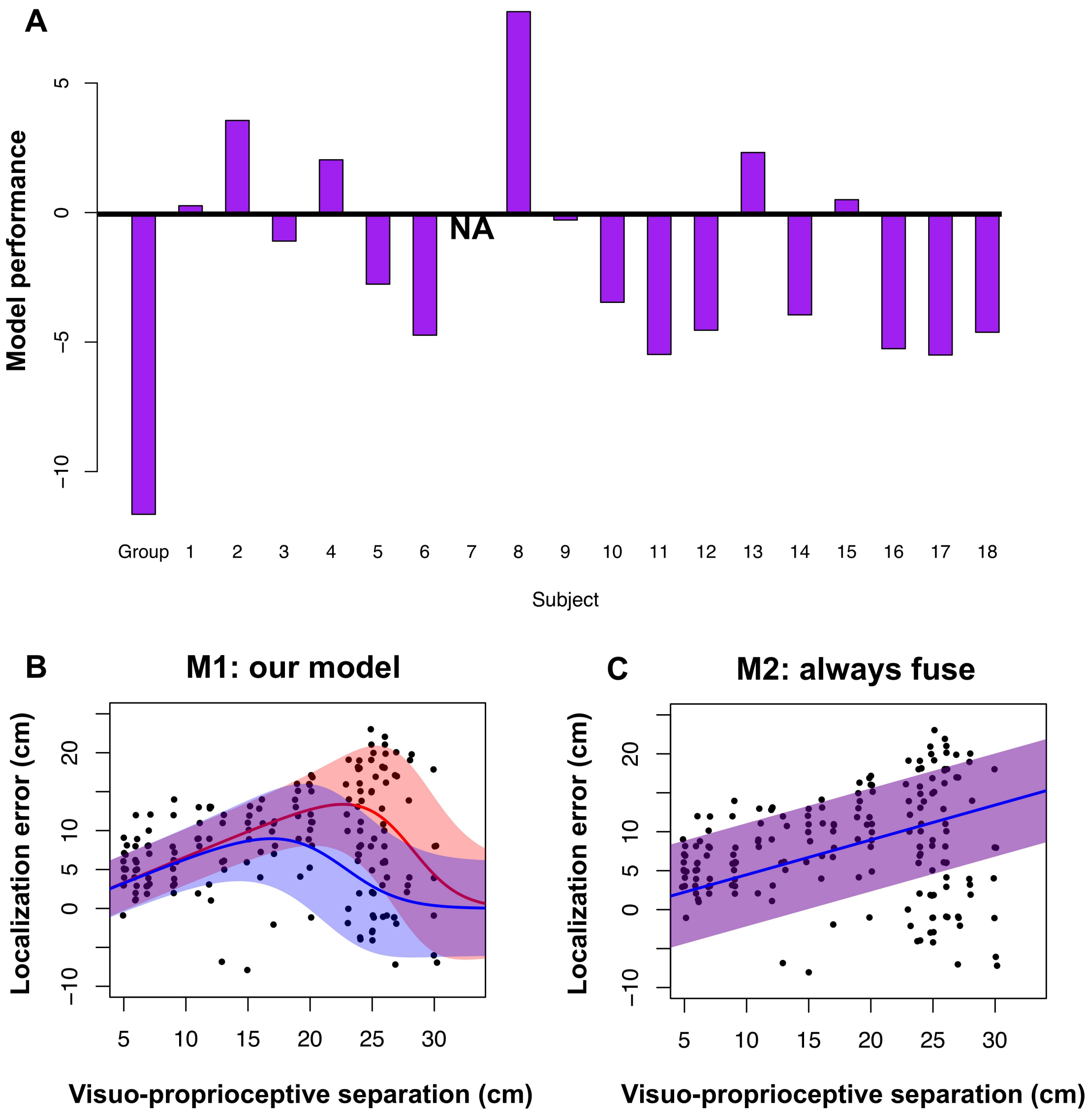

A
B
C
1 5 9 13 17 21 25 29 33 37 41 45 49 53 57 61 65 69 73 77 81 85 89
D
perceived
real
visual
E
Trial
on
off
...
Stroke
Visual
Motors
(synchronous)
Tactile
Motors
(synchronous)
Tactile
Motors
(asynchronous)
z
t

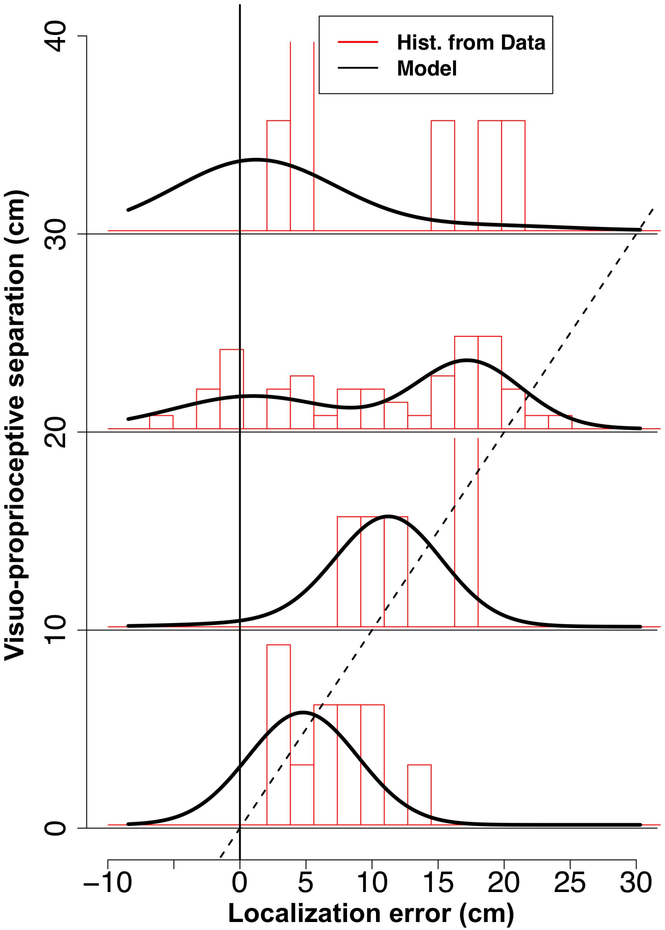

A
visual
real
congruent
incongruent
B
Cong Sync
Cong Async
Incong Sync
Incong Async
*
Score
3
2
1
0
-1
-2
Illusory Ownership
Control


A
B
sync
async
visual
proprioceptive
Hist. from Data
Model
***
***
Visuo-proprioceptive separation (cm)
Localization error (cm)
0
10
20
30
40
−10
0
5
10
15
20
25
30

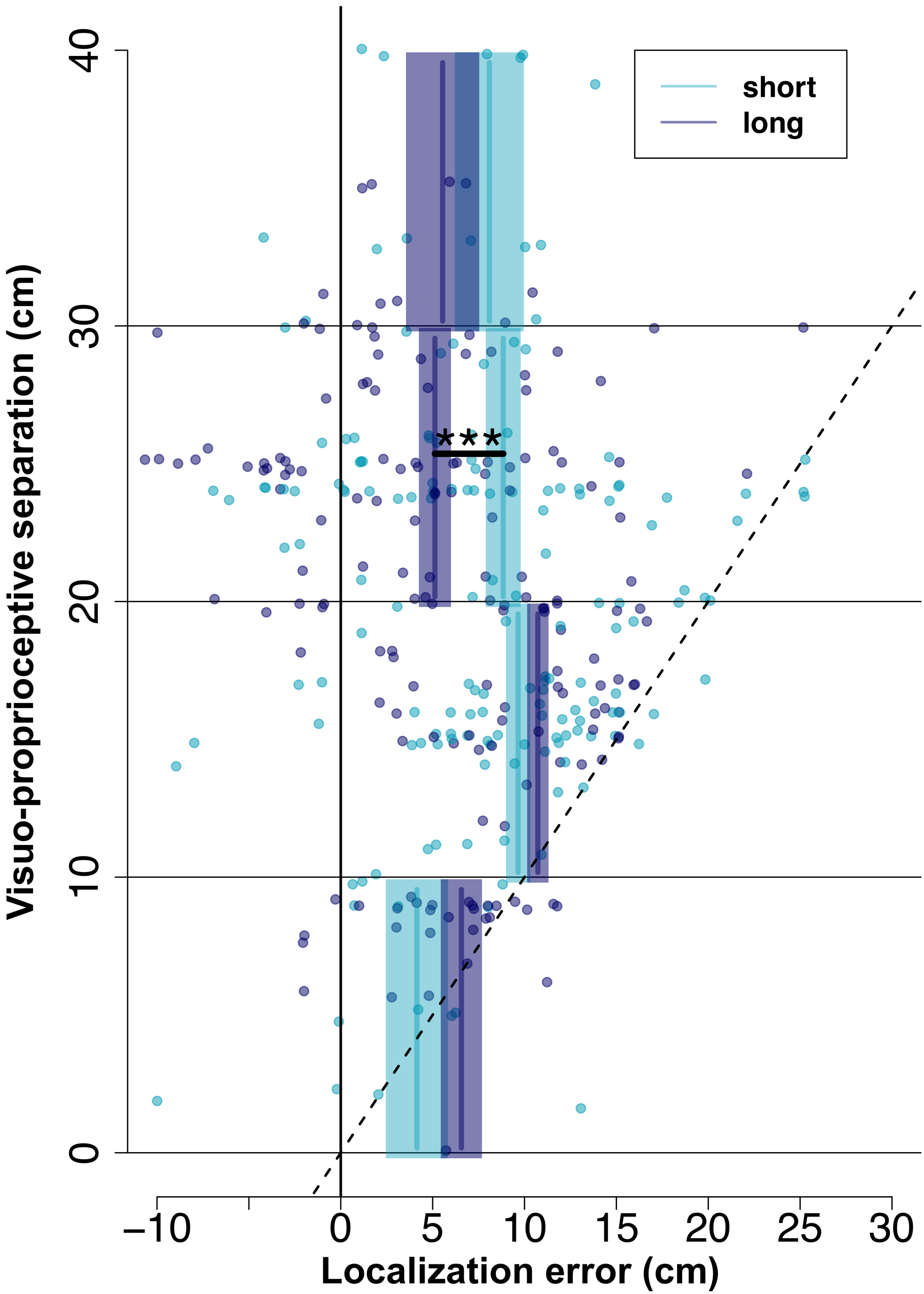


A
Real hand position Q
Ownership own=1 (my hand)
noise σ_p
noise σ_v
expect short delay
Proprio-ceptive position X_p
Visual position X_v
Visuo-tactile delay Z
B
Q(cm)
Not mine
Mine
Xp(cm)
Xv(cm)
Z(s)
C
Ownership
Visuo−proprioceptive separation(cm)
D
Subject 18
Subject 9
Subject 4
NA
NA
Threshold(cm)


A
Model performance
5
0
−5
−10
NA
Group
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
Subject
B
M1: our model
C
M2: always fuse
Localization error (cm)
20
10
5
0
−10
Visuo-proprioceptive separation (cm)
5
10
15
20
25
30



Hist. from Data
Model
Visuo-proprioceptive separation (cm)
0
10
20
30
40
−10
0
5
10
15
20
25
30
Localization error (cm)



short
long
***
Visuo-proprioceptive separation (cm)
0
10
20
30
40
Localization error (cm)
−10
0
5
10
15
20
25
30