# Precision not prediction: Body-ownership illusion as a consequence of online precision adaptation under Bayesian inference

Filip Novický[1,¤], Ajith Anil Meera[1,¤], Fleur Zeldenrust[1], Pablo Lanillos[1,2]

**1** Donders Institute for Cognition, Radboud University, Nijmegen, the Netherlands
**2** Cajal International Centre for Neuroscience, Spanish National Research Council, Madrid, Spain.

¤These authors share first authorship.
* filip.novicky@donders.ru.nl

## Abstract

Humans can experience body-ownership of new (external) body parts, for instance, via visuotactile stimulation. While there are models that capture the influence of such body illusions in body localization and recalibration, the computational mechanism that drives the experience of body-ownership of external limbs is still not well understood and under discussion. Here, we describe a mathematical model of this phenomenon via uncertainty minimization. Using the Rubber Hand Illusion (RHI) as a proxy, we show that to properly estimate one's arm position, an agent infers the least uncertain world model that explains the observed reality through online adaptation of the signals' relevance, i.e., the precision parameters (inverse variance of the prediction error signal). Our computational model describes that the illusion is triggered when the sensory precision quickly adapts to account for the increase of sensory noise during the physical stimulation of the rubber hand due to the real hand occlusion. This adaptation produces a change in the uncertainty of the body position estimates yielding to a switch of the perceived reality: the 'rubber hand is the agent's hand' becomes the most plausible model (i.e., it has the least posterior uncertainty). Overall, our theoretical account, along with the numerical simulations provided, suggests that while the perceptual drifts in body localization may be driven by prediction error minimization, body-ownership illusions are a consequence of estimating the signals precision, i.e. the uncertainty associated with the prediction error.

## Author summary

Understanding how body illusions occur is critical for comprehending body perception and embodiment. These illusions may hide powerful mechanisms of the brain and the body to handle uncertainty and conflicts in the sensory input and to adapt to body changes and world interactions. For instance, in less than one minute of visuotactile stimulation humans can experience body-ownership experiences of fake body parts or even experience that a different body is their own. Here, we present a possible explanation, through computational modeling, of why humans suddenly perceive a fake body part as their own. The proposed underlying process is the tracking of the uncertainty associated with the error between the predicted signals and the current sensory input. Our model describes that while the effect of body mislocalization during

body illusions may be driven by prediction error minimization, body ownership illusions are triggered by uncertainty resolution.

# Introduction

Humans navigate the environment and interact with it through their bodies. This happens with a precise balance between our body learnt representation [1] and sensorimotor related processes, such as body spatial localization and body ownership. Remarkably, these processes are "flexible", probably to handle uncertainty and conflicts in the sensory input and to adapt to body changes and world interactions. Strong experimental evidence reveal that body perception can be easily manipulated [2–4]. Body-ownership, i.e., "perceptual experience of body or body part as one's own" [5], has been associated with information integration in the ventral premotor cortex [6], right posterior insula, right frontal operculum as well as right temporoparietal junction [7]. However, the exact computational mechanisms behind and its intricate neural underpinnings are still not well understood. By developing computational models that replicate perceptual and motor effects of humans when subjected to multisensory conflicts, such as body illusions [8, 9], some inner characteristics of these hidden mechanisms may be revealed [3, 10]. In this sense, one of the most intriguing scientific questions is how we can experience the embodiment of external objects as a part of our own body. And, what is the mechanism behind the temporal dynamics of this body ownership illusion depending on the stimulation [11].

While the research literature already provides computational accounts that explain body mislocalizations and their temporal dynamics as an effect of body-illusions [3, 9, 10], there are no mathematical accounts – to our knowledge – that properly describe and explain the dynamics of body-ownership illusions. Here, we provide a computational model that explains the effect of body-ownership illusions as an indirect consequence of uncertainty minimization. In particular, our model describes that perceptual drifts in body localization are driven by prediction error minimization and body-ownership illusions are a consequence of online estimation of the signals' precision. Our model shows that the illusion is triggered when the sensory precision quickly adapts to account for the increase of sensory noise during the physical stimulation of fake body part due to the occlusion of the real body part. This adaptation produces a change in the uncertainty of the body localization estimates yielding to a change of the perceived reality: the 'rubber hand is the agent's hand' becomes the most plausible model—it is the hypothesis or model that has the least posterior uncertainty. Importantly, our model aligns with previous conceptual accounts of causal inference [10, 12] of body perception and the recent findings regarding the relation between the emergence of the illusion and the visual noise [13].

To introduce the proposed model we first summarize previous research on limb illusions and related computational models that are the backdrop of our body ownership model based on precision.

### Previous research on limb illusions and related computational models

There is an extensive body of literature on limb-body illusions and manipulations—see [3] and [8] for a review. Here, we briefly summarize the basic findings that point out two well-known effects: body-ownership of external/fake limbs and limb mislocalization (or perceptual drift).

Evidence shows that the majority of tested subjects—not all of them—perceive a fake limb as their own after a short period of visuotactile stimulation [3, 14]. The most well-studied body-illusion, the Rubber Hand Illusion (RHI) [14], can be seen as a

phenomenon where people falsely believe about the causes of tactile stimulation. Specifically, the participant sees a rubber hand that is being stroked (e.g., with a brush) at the same time that the real hand, which is occluded. In less than one minute, the participant has the subjective experience of owning the fake hand as part of the body [14]. This illusion comes with appearance and positional constraints and its intensity varies depending on the subject [15]. The illusion can be measured using subjective questionnaires objective measures, such as skin conductance, were introduced to evaluate the body's reaction to the process of the fake limb embodiment [16] and confirmed that external hands also produce physical responses.

The second well known effect is the perceptual drift, which can be seen as a mislocalization phenomenon due to sensory integration. Participants localize their own limb slightly drifted towards the fake limb [17,18]. By measuring the perceived location of the hand before and after the illusion is induced, researchers have shown that the closer participants estimate their real hand to the position of the rubber hand, the stronger the illusion may be [19]. While studies used this measure to elucidate the effects of the ownership effect [14,20], other studies highlighted that the correlation between the perceived ownership and the proprioceptive drift does not necessarily emerge [21–25], questioning the use of perceptual drifts to investigate the feeling of body ownership, and thus suggesting that perceptual drift and body-ownership illusion are a consequence of two interrelated but different processes.

Computational models of body perception should be able to account (at least) for these two well-known effects. There are already candidate computational models for the mislocalization effect (both perceputal and active drifts) based on the optimal integration of the multisensory signals [9,10,12,26,27]. The rationale is that the brain optimally integrates tactile, visual and proprioceptive cues to make them fit into an internal model of the body (learned through experience). In the presence of multisensory conflicts induced by the new visuotactile input, the adaptation mechanism within body localization forces the system to incorrectly estimate the limb in the space. The models differ on the mathematical framework used to integrate the signals [3]: optimal signal integration [28], Bayesian Causal Inference [12], Free Energy Minimization [9,10] and synaptic plasticity (e.g., spike-timing-dependent plasticity) [29].

The computational mechanism behind the body ownership illusion, however, is more elusive than the mechanism behind the mislocalization effect. One of the most influential proposed mechanisms, from a conceptual point of view, is that the brain, during the body illusion experiment, decides whether the information comes from the fake limb or the own limb [3,12]. This means that the subject decides between two models of the perceived reality. This may explain why the increase of presented cues (visuotactile synchronization, coherent poses, active movement) that support one of the two models results in a more significant illusion. Thus, according to this computational model causal inference plays a strong role in the body illusion. Still, the computational process, and its dynamics, behind the perceptual experience where the participant starts feeling a fake limb as its own is unknown. Bayesian Causal Inference as presented in [12] describes body ownership as a static probability computation where no explanation is given on the inner functional mechanism producing the dynamics of the perceptual drifts nor the illusion [11]. Importantly, in recent experiments, the notion of uncertainty as a driving mechanism due to (visual) noise induction [13] is gaining strength and aligns body ownership as a second-order decision-making process, where uncertainty monitoring modulates the experience [9]. For a more comprehensive summary of computational models of body illusions see the Appendix 0.1.

### A unified body-ownership model based on precision

Previous described models have varied significantly in their approaches to investigating specific phenomena observed in limb embodiment illusions. However, none have fully explained the mechanism and dynamics behind the ownership illusion and its relation to multisensory integration and localization drifts. We address this gap by presenting a novel mathematical description of the mechanism that may trigger ownership of a fake limb. Our model builds upon previous seminal Bayesian modeling accounts (i.e., Predictive Processing [9] and Bayesian Causal Inference [12]), supported by empirical investigations of illusory models.

Our hypothesis posits that body illusions are a consequence of resolving uncertainty by (online) adapting the precision (inverse variance) of prediction errors. Neurobiologically, precision primarily relates to encoding post-synaptic gain, which reports on prediction errors [30–34]. Interestingly, aberrant precision has been implicated in the causes of hallucinations [32, 35–38], and these precisions are modulated via NMDA receptors and neuromodulators such as dopamine, acetylcholine, serotonin, and noradrenaline [32, 39–42]. Essentially, as precision encodes the uncertainty of the prediction error, it may also model the experimental results observed when introducing visual noise in a body illusion experiment, where sensory uncertainty was found to be essential for the ownership illusion [13].

We propose a computational model of body ownership illusion—evaluated in the RHI paradigm—that uses precision and its adaptation to explain the perceptual switch of feeling a fake hand as part of the body. The model schematic, the paradigm and variables used are described in Figure 1A,B and C respectively. Similarly to the Bayesian Causal Inference approach [12], the brain decides between two different hypotheses or models $M$ of reality. However, our model will base its decision on the models posterior uncertainty. Model $M_1$ explains the observed data as "my hand is the real hand", while $M_2$ explains the data as "my hand is the rubber hand". The induced noise from covering the real hand prompts the system to adapt the signal noise. In conjunction with visuotactile stimulation, this adaptation makes the posterior estimates more uncertain. Consequently, hypothesis/model selection switches to the most probable one, which is to own the fake limb, thus producing the illusion.

Our computational model suggests that: *i*) body mislocalizations (e.g., proprioceptive drift) result from body state estimation, where prediction errors are minimized, and *ii*) body ownership illusions, i.e., perception of a fake limb/hand as one's own, result from uncertainty minimization, where precision adaptation to overcome induced noise triggers model selection switch.

## Methods

We introduce a mathematical model of the body-ownership illusion, which is based on the Bayesian brain hypothesis [43], particularly, on the active inference framework [44]. This framework proposes that the brain generates cognition and behaviour following a Bayesian optimization strategy. Active inference postulates that living beings show resilience over time, which is achieved via minimizing entropy (the second law of thermodynamics) in order to preserve their equilibrium state (i.e., to self-organize) [45, 46]. To do this, living systems embed a generative internal model that explains the world (i.e., perception) and adjust to the environment (by changing their belief or exerting actions) via minimizing the free energy functional (an upper bound on sensory surprisal) [47]. We first describe the the full model based on maximum posterior precision (minimum posterior uncertainty) and then we detail the computational sub-processes and the algorithm.

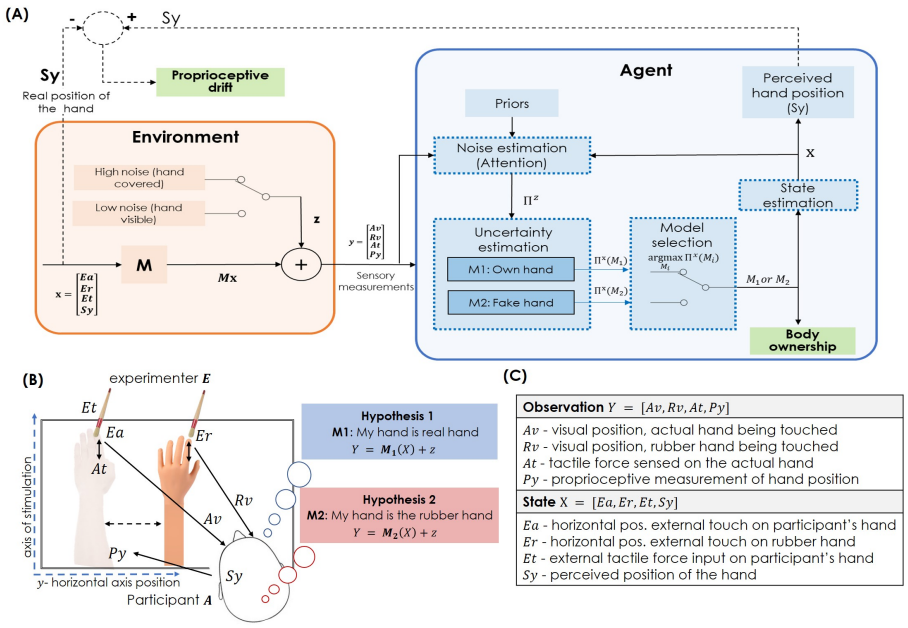


**Fig 1. Maximum posterior precision body-ownership model**. (A) Our proposed model consists of two main blocks: the environment and the agent. The state of the environment $x$ produces the output $y$ that is measured by the agent as sensory measurements, along with the observation noise $z$. The visual noise is high when the real hand is covered, and low when the real hand is visible. The agent estimates the noise via attentional mechanisms, i.e., precision priors, which leads to estimating the uncertainty of the $M_1$ and $M_2$ models. Through model selection, the agent selects the best suitable model, producing the body-ownership illusion. Using the selected model, the state estimation is performed, resulting in the state estimate $X$ that represents $x$. The proprioceptive drift is the difference between the perceived hand position $Sy$ and the real position of the hand **Sy** (that the agent do not have access to). (B) Experiment formalization. The brain is testing two competing models (i.e., hypotheses): $M_1$ stands for believing that "my hand is the real hand", while $M_2$ indicates that "my hand is the rubber hand". These two internal models maps (as a generative model) the states $X$ into the observations $Y$. (C) Notation used for the observations (sensory measurements) and the states.

## Maximum posterior precision body-ownership model

Figure (1A) shows our proposed model for the body-ownership illusion based on the posterior precision. It consists of two main parts: the environment and the agent. The environment consists of an external experimenter that manipulates on the human hand such that it generates signals or data that could make the participant/agent experience the illusion. The agent experiences the RHI under the influence of these signals, triggered by the experimenter.

The environment consists of an experimenter (E) that provides the participant (or the agent) with following stimuli (see Fig. (1C) for the notation used): stroking the real hand (**Ea**) and stroking the rubber hand (**Er**). The experimenter generates a force by the stroke (**Et**) on both the hands. The agent observes this and collects the visual signals **Av** and **Rv** that represent the position of the real hand and the rubber hand, respectively. The agent also feels a tactile sensation **At** on its real hand. The $y$ coordinate position of the real hand **Sy** is observed by the agent as $Py$ (proprioceptive

measurement). We stack these variables together as: i) environment state, from the experimenter, as $\mathbf{X} = [\mathbf{Ea\ Er\ Et\ Sy}]^T$ and ii) output measurements of the agent as $\mathbf{Y} = [\mathbf{Av\ Rv\ At\ Py}]^T$. The notation $E$ is used for the experimenter's inputs and, $A$ and $R$ are used for the signals on the actual hand and the rubber hand respectively.

We assume that there is a generative process that maps the sensory measurements from the state variables. We define this mapping as a linear function $\mathbf{M}$ between the state $\mathbf{X}$ and output (measurements) $\mathbf{Y}$ with measurement noise $\mathbf{z}$ as:

$$\mathbf{Y} = \mathbf{MX} + \mathbf{z} \rightarrow \begin{bmatrix} \mathbf{Av} \\ \mathbf{Rv} \\ \mathbf{At} \\ \mathbf{Py} \end{bmatrix} = \mathbf{M} \begin{bmatrix} \mathbf{Ea} \\ \mathbf{Er} \\ \mathbf{Et} \\ \mathbf{Sy} \end{bmatrix} + \mathbf{z} \tag{1}$$

The measurement noise $\mathbf{z}$ is composed of four components ($\mathbf{z^{Av}}, \mathbf{z^{Rv}}, \mathbf{z^{At}}, \mathbf{z^{Py}}$). The real hand coverage that obstructs the visual signal $\mathbf{Av}$ is modeled by adding a high noise $\mathbf{z^{Av}}$ to the signal. Similarly, for no coverage of the real hand, we use a low noise $\mathbf{z^{Av}}$ on the signal. The noise level $\mathbf{z^{Av}}$ relates to visual information, thus the agent either sees or does not see these signals, which translates to the amount of noise that is generated with it[1]. Thus, the covered hand is completely uncertain.

The agent processes the measurement $\mathbf{Y}$ to make sense of the world through estimation. However, the experimenter can manipulate $\mathbf{Y}$ to make the agent fall into the body-illusion. We model this process within the RHI paradigm using the following components, as shown in Figure 1A: i) noise estimation (attention), ii) state estimate (body localization), iii) posterior uncertainty estimation, and iv) model selection (body ownership). The next section will explain each of these computational blocks in detail.

## Computational sub-processes of the agent

At each time step, the agent estimates the state and noise in the environment, along with its uncertainty in estimation. This is done for two competing models or hypotheses of the perceived reality—either 'my hand is my real hand' or 'my hand is the rubber hand'—and the agent selects the model that best minimises its uncertainty in estimation. This section explains all subprocesses within the agent in Figure 1A.

### Competing models of perceived reality

The two competing models dictate the perceived reality during the illusion. The first model ($M_1$) relates to the common perception of owning the real hand. The second model ($M_2$) relates to perceiving the rubber hand as the agent's own. The agent tracks or estimates the uncertainty of both models of reality and select the one that minimizes the model's uncertainty about hidden states. In other words, the agent tries to explain the observed data as accurately as possible, given the structure (hidden states) of the model.

We design the agent's generative model, that defines the two models of perceived reality, as the mapping between the environment state variable $X$ and the measurements $Y$ (provided by the senses). We use the environment model given in Eq (2) as the first competing model to represent the "my hand is my real hand" model $M_1$,

[1]This follows from the solution to the dark-room problem introduced in the free-energy debate, which addresses why uncertainty minimizing agents do not stay all their lives in a dark environment. The argument is that darkness is completely uncertain [48].

where the brain believes that the effects observed fit to the real limb.

$$Y = M_1 X + z \rightarrow \begin{bmatrix} Av \\ Rv \\ At \\ Py \end{bmatrix} = \begin{bmatrix} 0.8 & 0 & 0.2 & 0 \\ 0 & 1 & 0 & 0 \\ -0.3 & 0 & 0.7 & 0 \\ Ay - Ry & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} Ea \\ Er \\ Et \\ Sy \end{bmatrix} + z \tag{2}$$

The sensory measurements on the actual hand ($Av$ and $At$) are influenced by the external manipulations applied on the actual hand ($Ea$ and $Et$), but are independent of the external information on the rubber hand $Er$ and of the hand position $Sy$. The sensory measurement on the rubber hand ($Rv$) is determined by the experimenter's actions on the rubber hand and is independent of the external manipulation on the actual hand ($Ea$ and $Et$) and the hand position $Sy$. This models the separation between the actual hand and the rubber hand in terms of body ownership. Finally, $Ay - Ry$ parametrizes the distance between the real and the fake hand in the horizontal axis. The perceived hand position $Py$ is influenced by the measured hand position $Sy$, and is biased by the distance between the actual hand and the rubber hand ($Ay - Ry$).

In the presence of a fake hand there is a competing model that describes that the participant may own the rubber hand, thus producing the illusion. We account for this information flow through a "rubber hand is my actual hand" model given by:

$$Y = M_2 X + z \rightarrow \begin{bmatrix} Av \\ Rv \\ At \\ Py \end{bmatrix} = \begin{bmatrix} 0.8 & 0 & 0.2 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & -0.3 & 0.7 & 0 \\ 0 & Ay - Ry & 0 & 1 \end{bmatrix} \begin{bmatrix} Ea \\ Er \\ Et \\ Sy \end{bmatrix}. \tag{3}$$

where the tactile force on the actual hand $At$ depends on the environment (experimenter) input on the rubber hand $Er$, and not on the experimenter manipulation on the real hand $Ea$. We hypothesize that when the real hand is covered, the noise in the visual sensory measurement $Av$ increases considerably, resulting in the brain switching from "my hand is the real hand" to the "rubber hand is my hand".

The rubber hand model from Eq (3) just differs from $M_1$ Eq (2) in two aspects: *i)* $At$ takes information from $Er$ instead of $Ea$; and analogously to $M_1$ *ii)* we also included a term $Ay - Ry$ that models the distance from the actual hand to the rubber hand and modulates, in this case, the position of the rubber hand $Er$.

The proprioceptive signal of the hand position ($Py$) can be used to calculate the proprioceptive drift from the actual position ($Sy$) by the product of two terms: *i)* the stimulation on the actual hand ($Ea$) and *ii)* the distance between the actual hand and the rubber hand ($Ay - Ry$). Intuitively, this models an increase in proprioceptive drift when the noise on the visual observation of the actual hand position ($z^{Av}$) is high, through an inaccurate estimation of $Ea$, depending on the distance between the actual hand and the rubber hand ($Ry - Ay$).

In summary, the two models or hypotheses of reality differ slightly, but provide an important behavioural change when it comes to simulating the RHI. Particularly, the real hand model considers the information from the experimenter touching the real hand and also the pressure (or force) applied to the real hand as more important compared to the alternative model, which values the rubber hand more as a source of these stimuli and thus can allow itself to slightly ignore the force $Et$.

### Role of precision in the computational model

In our computational model, we realize precision using three different ways: *i)* noise precision $\Pi^z$ that is parameterized using $\lambda$ (details in appendix 0.2) , *ii)* prior precision on estimates $P^\lambda$, and *iii)* posterior precision on estimates $\Pi^X$ and $\Pi^\lambda$. While the noise

precision ($\Pi^z$) is learned online by our model using the agent's prior precision ($P^\lambda$), the posterior precision ($\Pi^X$) is computed alongside the state estimates $X$. The noise precision learning contributes to the identification of the noise levels in the sensory data, while the posterior precision computation contributes to the decision (model selection) of which model to rely on ($M_1$ or $M_2$). The following sections will detail the mathematical formulations behind both processes.

### Model selection by uncertainty minimization

First, we describe the overall mechanism of model selection[2] that produces the body ownership illusion—depicted as a pseudocode in algorithm 1—and then we detail the different computations, such as noise precision estimation, body state estimation and uncertainty estimation. The code and the parameters for replication can be found in Appendix 0.3.

**Algorithm 1** Model selection by uncertainty minimization

**Require:** Models $M = \{M_1, M_2\}$
  **for** $t = 0 : dt : T$ **do**
    $\mathbf{Y} = DataGeneration()$ ▷ Environment generates the data
    **for** $M_i \in M$ **do** ▷ Iterate through possible models
      $\Delta\lambda = NoiseEstimation(\mathbf{Y}, \lambda, M_i)$ ▷ Compute $\Delta\lambda$ by minimizing $F$
      $\lambda_{t+1}(M_i) = \lambda_t(M_i) + \Delta\lambda$ ▷ Update the noise parameter
      $\Delta x = StateEstimation(\mathbf{Y}, x, M_i)$ ▷ Compute $\Delta x$ by minimizing $F$
      $x_{t+1}(M_i) = x_t(M_i) + \Delta x$ ▷ Update the state estimation
      $\Pi^\lambda(M_i) = EvaluatePrecision()$ ▷ Precision on $\lambda$
      $\Pi^x(M_i) = EvaluatePrecision()$ ▷ Precision on $X$
      $\Sigma^x(M_i) = (\Pi^x(M_i))^{-1}$ ▷ Compute uncertainty in state estimation
    **end for**
    $M^* = \arg\min_{M_i} trace(\sqrt{\Sigma^x(M_i)}\,)$ ▷ Model selection
    $x_{t+1} = x_{t+1}(M^*)$ ▷ Select the state
    $\lambda_{t+1} = \lambda_{t+1}(M^*)$ ▷ Select the noise
  **end for**

For all competing models we compute the posterior precision of states ($\Pi^X$), along with the estimated state ($X$). We propose a model selection criterion for the RHI, using the posterior precision of state estimation, for both the competing models as:

$$M = \arg\min_{M_i} trace\big[(\Pi^X)^{-\frac{1}{2}}\big]. \tag{4}$$

The core idea is that the agent selects the model that best minimises the uncertainty in state estimation. We define it as the trace (sum of diagonal elements) of the square root of the posterior covariance matrix (inverse of precision matrix, $\Sigma^X = (\Pi^X)^{-1}$) of state estimation . According to the proposed model, the RHI will begin at the moment where $M_2$ shows lower uncertainty, compared to the non-illusory model. Once the RHI starts, the distinction between the actual hand and the rubber hand is hypothesised to begin to fade from the agent's point of view, resulting in a flow of information from external inputs on the rubber hand ($Er$) to the sensory perception about the actual hand ($At$ and $Py$). This can be observed by comparing the differences in the last two rows of equations (2) and (3), as the agent switches from model $M_1$ to model $M_2$, once the RHI begins. In summary, model selection through uncertainty minimization (maximum posterior precision) is the core mechanism to model the body ownership illusion.

[2]See Discussion section for the reason of using model selection instead of model averaging, which is usually followed in computational neuroscience.

### Joint State and Noise Precision Estimation

We model body state estimation as inferring the internal states of the body from the observations $Y$, given a model of reality. The agent keeps track of the states (the activities of the experimenter - $Ea, Er, Et$, and the perceived hand position $Sy$). This estimate is computed in the agent by approximate Bayesian inference through the minimization of the free energy [49]. The free energy of the agent that tries to predict the outputs $\mathbf{Y}$, under state $X$ and model $M$, and learns the measurement noise parameter $\lambda$, with prior mean and precision $\eta^{\lambda}$ and $P^{\lambda}$ respectively, can be written down as (refer [50] for full derivation):

$$F = \frac{1}{2}(\lambda - \eta^{\lambda})^T P^{\lambda}(\lambda - \eta^{\lambda}) + \frac{1}{2}(\mathbf{Y} - MX)^T \Pi^z (\mathbf{Y} - MX) - \log|\Pi^z|. \tag{5}$$

Here, the noise precision $\Pi^z$ is parametrized using $\lambda$ (refer appendix 0.2). The agent performs state estimation under a selected model using free energy minimization as [50]:

$$X(t + \Delta t) = X(t) + \left(e^{-k\frac{\partial^2 F}{\partial X^2}\Delta t} - I\right)\left(\frac{\partial^2 F}{\partial X^2}\right)^{-1}\frac{\partial F}{\partial X}. \tag{6}$$

The update rule uses the first two gradients of $F$ in Eq 5, which can be calculated by differentiating Eq (5) with respect to $X$. This mathematical routine updates $X$ such that it minimises $F$ in Eq 5 with respect to $X$. The agent performs online noise precision learning by updating $\lambda$ such that it minimises the free energy, using its first two free energy gradients as [50]:

$$\lambda(t + \Delta t) = \lambda(t) + \left(e^{-k\frac{\partial^2 F}{\partial \lambda^2}\Delta t} - I\right)\left(\frac{\partial^2 F}{\partial \lambda^2}\right)^{-1}\frac{\partial F}{\partial \lambda}. \tag{7}$$

The gradients of $F$ can be calculated by differentiating Eq (5) with respect to $\lambda$. Note that the update of $\lambda$ changes the learned observation noise $\Pi^z$ as per Eq (10), which directly enters $F$ in Eq (5) and influences the state estimation in Eq (6) in the next time step. Therefore, the noises and its estimation directly influence the state estimation. Under high noises, the agent can have illusory body state estimation. During the RHI experiment, the agent's state estimate may be disrupted with regard to $Sy$, due to high noise in the visual data $\mathbf{z^{Av}}$, leading to the experience of a proprioceptive drift where the perceived hand position $Sy$ shifts from the real hand position $Ay$, towards the rubber hand position $Ry$.

### Uncertainty in estimation

According to FEP, the posterior precision on the estimate (the second order statistic reflecting the confidence on estimation [51]) is obtained using the second order gradient of $F$. We use this to evaluate the uncertainty in estimation by taking the inverse of the posterior precision on state estimates as [50]:

$$\Sigma^X = (\Pi^X)^{-1} = \left(\frac{\partial^2 F}{\partial X^2}\right)^{-1}. \tag{8}$$

$\Sigma^X$ denotes the agent's uncertainty in making predictions about the states, given the observed data. The agent computes the precision for both the competing models $\Pi^X(M_1)$ and $\Pi^X(M_2)$ to make a decision on the model selection. It is used directly in the model selection criterion given in Eq (4). The core idea is that the agent decides to believe in a model that best minimizes (or resolves) its uncertainty in estimation. The hypothesis is that once the visual signals of the hand $(Av)$ is highly noisy (hand covered), the condition $trace(\sqrt{\Sigma^X(M_1)}) > trace(\sqrt{\Sigma^X(M_2)})$ becomes true and the agent switches its model selection from a non-illusory model $(M_1)$ to an illusory model $(M_2)$, resulting in the experience of RHI.

## Evaluative measures for the model

Here, we describe how to model illusion sensitivity and measure proprioceptive drift. Note that these measures are not placed inside the agent in Figure 1, but uses them for evaluations and comparisons.

### Modeling individual differences and illusion sensitivity of the agent

Previous research shows a high variability between individual susceptibility to the body illusion. Moreover, depending on the manipulation (e.g., temporal synchrony) body illusion sensitivity varies [15]. To account for the agent sensitivity to experience the illusion, we define infinite models that exist between 'non-illusory model' ($M_1$) and 'illusory model' ($M_2$), which is modulated by a parameter $\alpha$ that defines an agent's sensitivity to illusion:

$$SM = \frac{1}{1+e^{-\alpha}} M_1 + \frac{e^{-\alpha}}{1+e^{-\alpha}} M_2. \tag{9}$$

In the case of more than two models of perceived reality, a softmax operator could be used. Technically, the resulting subjective model $SM$ then spans between these two models, based on the $\alpha$ value. Heuristically, the higher the $\alpha$, the more it resembles the non-illusory model and vice versa. Thus, Eq (9) captures the variability of body ownership sensitivity, which will also tunes the experienced proprioceptive drift from different participants.

### Measuring the proprioceptive drift

The agent will not be able to report the proprioceptive drift during RHI because it is in illusion and does not have access to the true signal $\mathbf{Sy}$. The proprioceptive drift in this paper is computed by taking the difference between the real position of the hand $\mathbf{Sy}$ and the agent's estimated hand position $Sy$ as shown in Figure 1.

# Results

We focused on analysing the proposed precision-based mechanism for modeling body ownership illusions. Specifically, we synthetically simulated the RHI as being one of the most paradigmatic body perceptual manipulations, where a participant experiences the illusion that an isolated fake hand becomes an agent's hand. The parameters used to generate the results are described in Appendix 0.3. This section shows the following results:

1. We validate the proposed model in a simulated RHI experiment (Fig. 2), where we show how posterior uncertainty minimization under the two competing models of reality yields to a perceptual switch that produces the illusion of owning the fake hand.

2. We analyse the precision adaptation mechanism of the proposed model, which is underneath the perceptual switch (Fig. 3). The introduction of visual noise by covering the hand is one of the factors for producing the effect. The online adaptation to the sensory noise through precision estimation appears as an essential property to experience the illusion.

3. To further validate our model, we show that breaking the precision learning – where the noise over $Av$ is not learned – leads to a state where switching current model of reality to a different one is implausible, hence no illusion can occur (Fig. 4).

4. We evaluate the proprioceptive drift generated depending on the selected competing model (Fig. 5), where $M_1$ produces noisy estimations but not drift, while when selecting $M_2$ the model produces the experimentally observed proprioceptive drift.

5. We validate the effect of the inter-hand distance on the proprioceptive drift dynamics giving that the illusion is happening (Fig. 6).

6. Finally, using the modeling of artificial participants with individual differences, we explore the effect of the participant susceptibility to experience the RHI and the inter-hand distance on the generated proprioceptive drift (Fig. 7). The results show resemblance with the profiles found in a recent experiment in humans and macaques [52].

## Body-ownership illusion: model selection based on the uncertainty minimization

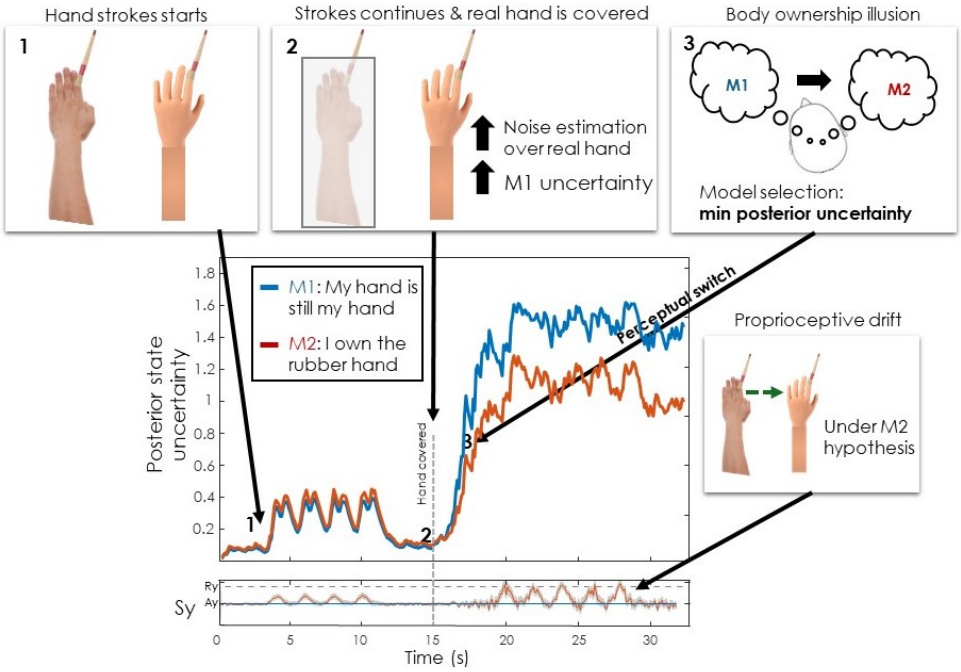


**Fig 2. Body-ownership perceptual switch.** We show the experiment sequence and its result under our model: 1) The both hands are being simultaneously stroked; 2) The real hand is covered and the strokes continue; 3) The participant starts experiencing the RHI. In the bottom graph, it is shown the competition of the two models and their uncertainty during the experiment. After the hand is covered, there is a switch between these two models, as the $M_1$ starts being more precise. This then leads to believing that the rubber hand is a participant's hand, thus leading to a proprioceptive drift depicted in the graph below.

With the proposed model selection algorithm in mind—as described in algorithm (1)—we can now consider what happens when the uncertainty is estimated to select one or the other model, as well as to face-validate the proprioceptive drift happening with the model selection. This is shown in Figure (2). First, the posterior uncertainty in the beginning of the trial – when both hands are visible and being stroked by the researcher

(Figure (2(1)) – favors the non-illusory model ($M_1$). Once the hand becomes covered (Figure 2(2)) there is an immediate increase in $M_1$ uncertainty which is due to the increase of noise estimate of the real hand. After a few moments, there is a switch between the two models regarding the posterior uncertainty (Figure 2(3)). This leads to a participant believing that they own the rubber hand, as the $M_2$ hypothesis now provides lower uncertainty compared to the $M_1$. As this is happening, we can see that in the $M_2$ model, there is also an observable proprioceptive drift driving away from $Py$=0 towards the location of $Ry$=0.8. This is further described in section 'Proprioceptive drift in competing models' below. Ultimately, this shows that there are two distinct – but highly interrelated – computational mechanism responsible for body ownership (minimizing uncertainty) and proprioceptive drift or mislocalization (state estimation as multisensory integration).

There are some limitations that are worth considering here. First, the model is mostly driven by the increase of the noise from the real hand, and does not consider the simultaneous spatio-temporal synchronization of strokes on the real and rubber hands. This implies that covering the fake hand is sufficient for experiencing the body-ownership illusion. Our approach thus does not meet the acknowledged finding that to induce the rubber hand, the strokes are fundamental, and they also have to be synchronous for both hands [53]. Our response to this is that we use RHI as a proxy for the entire body-ownership illusion, thus focusing only on the mechanism of competing models, therefore allowing ourselves to simplify the modeled experiment and mechanisms of RHI significantly.

## Estimating noise through precision adaptation

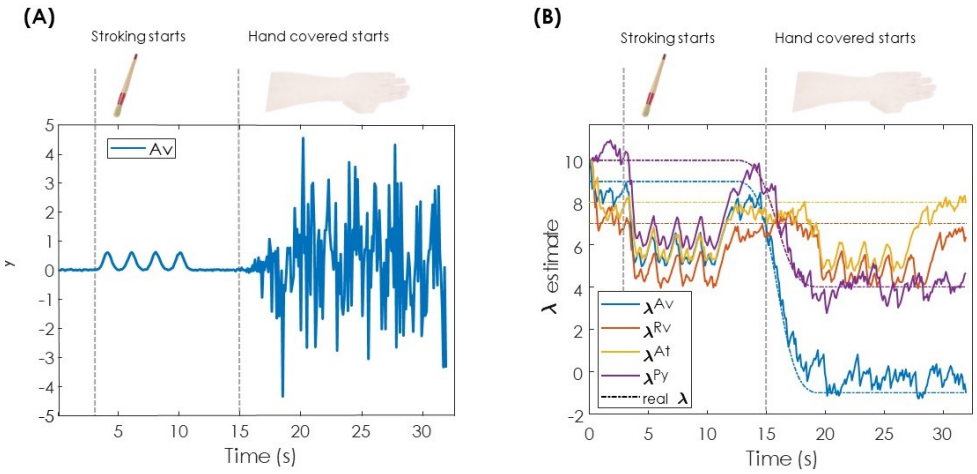


**Fig 3. Precision adaptation (attention).** **(A)** Visualization of the sensory measurement data. The *Av* signal in blue (visual signal from the real hand) is highly noisy after the real hand is covered (at time = 15seconds). This represents the highly uncertain nature of the visual data for the hand position, once it is covered.**(B)** shows is the result of the noise estimate block from our model in Fig. 1, given the sensory measurements on the left figure. The colored plots track the true noise parameter $\lambda$ (in dashed black), showing that our model is capable of correctly estimating the precision levels of all sensory signals. In other words, the uncertainty of the sensory signals is fully captured by our model in real time.

First, it is important to understand how reliable our model is when estimating the

noise. This is depicted in Figure (3B). This figure shows that both proprioceptive (purple line) and visual signals from the real hand (blue line) become noisy after the hand is covered at the 15th second. This is because, based on our assumption, covering the real hand relates to generating more noise. For the proprioceptive signal, the noise increases because of lower sensory integration (due to decreased Av signals). This signal is for clarity visualized in Figure (3A). In this sense, the proprioceptive signal should be regarded as a higher level computation happening in the brain already, rather than pure afferent signal coming from the hand position per se. In other words, it is a *felt* proprioception. In addition, the parameters follow the real noise (the black line) in all signals. Ultimately, all precisions are estimated correctly.

## Validating the model via breaking precision

We hypothesised that precision learning is a necessary precondition for model switching in Figure (2). In other words precision learning is crucial for the model switching shown in Figure (2), and if this is broken (by providing such priors that inhibits learning), the switch will not occur. To test this, we disrupted the precision learning process and observed its effects, as illustrated in Figure (4). Figure (4A) shows improperly learned precision over Av (blue curve) due to a high prior on $\lambda$. Consequently, Figure (4B) demonstrates that the posterior state uncertainty of $M_1$ remains consistently lower than $M_2$, regardless of hand coverage. This contrasts with our earlier findings in Figure (2). These results provide face-validation for the necessity of precision learning in achieving the body-ownership illusion, highlighting its critical role in our model.

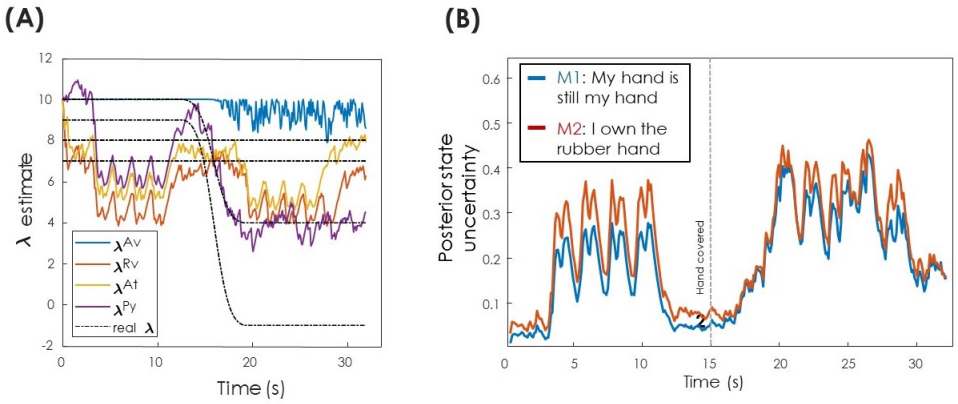


**Fig 4. Breaking precision leads to no illusion**. **(A)** This figure is similar to Figure 3B, except for the fact that the agent is now biased towards believing that the signal **Av** is highly precise although it is as noisy as in Figure 3A ($\lambda^{Av}$ in blue stays high around 10 instead of tracking the real $\lambda$ in black). **(B)** We show that if an agent is (incorrectly) overconfident over the visual signal Av, it will lead to a behaviour where the two competing models will not switch—red curve always stays above blue in comparison with a switching in Figure 2.

## Proprioceptive drift in competing models

After getting an intuition of the model data and the precision terms, we introduce the simulation of proprioceptive drift in more detail compared to Figure 2 where it is mentioned only briefly as an independent mechanism happening along with the model selection. Here, we show that the proprioceptive drift happens only in the illusory model $M_2$ (Figure 5B) but not in the non-illusory model $M_1$ (Figure 5A). These plots represent

a single trial of the experiment described in Figure 2, where in the first half when the real hand is still visible, the agent perceives all signals, and can therefore correctly estimate the touch of the fake hand ($Er$), of the real hand ($Ea$), and the tactile sensation ($Et$). The proprioceptive signal ($Sy$) remains the same for the non-illusory model. The illusory model already shows a small proprioceptive drift. The reason why the proprioceptive drift is not felt in the first half is because the model selection prefers the non-illusory model. When there is no switch of the models to the illusory one, the proprioceptive drift nor illusory tactile perception do not occur (Figure 5A).

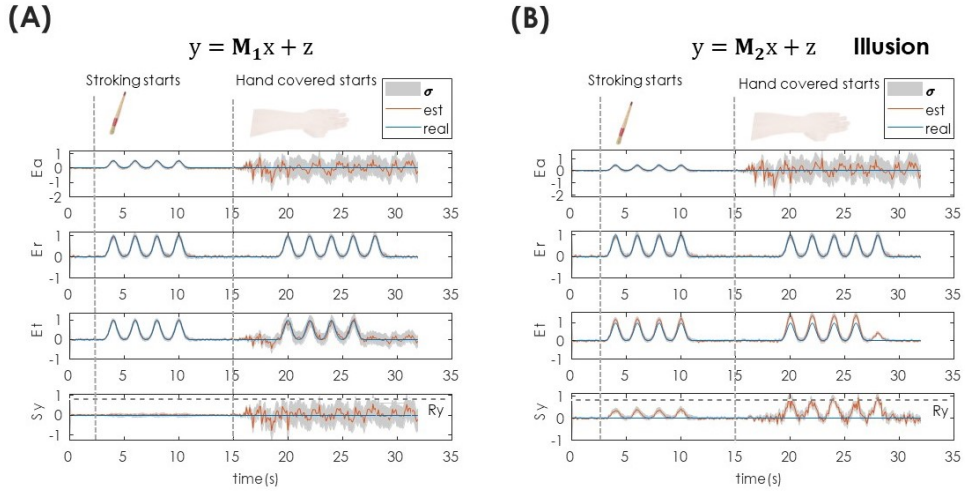


**Fig 5. Replicating the proprioceptive drift. (A)** State estimation of all the variables under the non-illusory model $M_1$. **(B)** the same state estimation when the illusory model $M_2$ is winning and been chosen. Both figures show state estimation of all input data $Ea$, $Er$, $Et$, and $Sy$. Only in the illusory model on the right, the proprioceptive drift ($Sy$) and the related touch ($Et$) are observed. Also note the increase of noise in $Ea$ in under both models when the hand is covered. The distance between the rubber hand and real hand is 0.8 in this case. The distance dependence on the proprioceptive drift is evaluated in Figure 7.

## Dependence of proprioceptive drift on inter-hand distance

The dependence of the proprioceptive drift on the distance has been observed empirically in many experiments [54, 55]. Here, we also validate our model with this logic. In Figure (6), we compare two illusory models where the distance of the two hands is increased. In Figure (6A), the inter-hand distance is smaller compared to the one in Figure (6B). In the latter figure, the estimate in $Er$ is incorrect, which leads to a smaller and shorter proprioceptive drift compared to Figure (6A). Ultimately, the proprioceptive drift is diminished for the completely illusory model when the distance increases. This observation is consistent with previous results that report decreased proprioceptive drift with increased distance [55]. Interestingly, though, this dependence is reversed in results of [52]. We therefore modulate the strength of both the illusory and non-illusory models via the softmax algorithm in Eq (9) – see methods for more detail. Doing so we assume that we modulated the subjective models of the competing models. Changing these, we show that the proprioceptive drift is strengthening when slowly increasing the distance between the two hands, and then finding its plateau (Figure (7)). In this figure, the higher $\alpha$ is associated with the model being less illusory and vice versa. In this figure, to simplify, we take the maximum value of the

proprioceptive drift, which is represented on the y-axis, while the distance is on the x-axis. In addition, we also show that modulating the subjective strength of the illusory model can provide different tendencies of the proprioceptive drift, as shown in [52]. This, from our perspective, also links the difference between the selection of the illusory model and feeling the drift, and not selecting the illusory model and still feeling the drift. In this way we have touched on the topic of the proprioceptive drift criticism as an indicator of the RHI, as they relate to distinct computational mechanisms [21–24] – state estimation and precision maximisation.

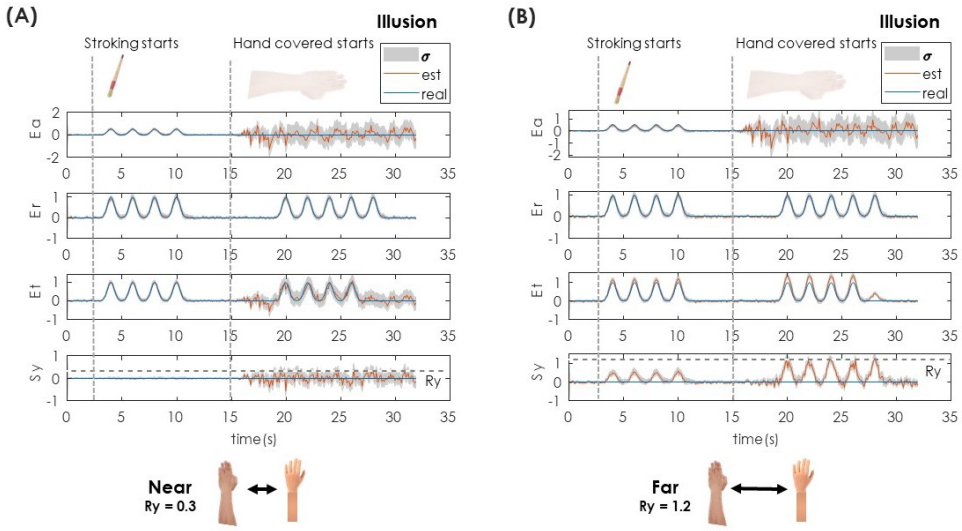


**Fig 6. Inter-hand distance influence on the preceptual drift.** The same template was used in Figure 5. Here, however, both figures represent the state estimation under the illusory model, where we varied the interhand distance. **(A)** Near. Effect of $M_2$ when the interhand distance is small (Ry = 0.3). **(B)** Far. effect of $M_2$ when interhand distance is increased (Ry = 1.2). This shows that with the increase of the distance between the hands, the proprioceptive drift strengthens.

# Discussion

In this theoretical paper, we have proposed a novel model that can explain body illusion dynamics and showed it in the RHI experiment. We built a general model that selects between two competing models of reality, and the switch between these two is computed by maximizing the precision (minimizing the uncertainty) of the state estimation (Fig. 2). The model was face-validated on previous experimental results via replicating the proprioceptive drift (Fig. 5), and its dependence on the inter-hand distance (Figs. 5 and 6). We showed that the model can also reproduce the proprioceptive drift due to the subjective differences in the structure of the models (Fig. 7). Overall, our model provides useful insights into *i*) the process dynamics during the RHI experiment due to the online precision adaptation; and *ii*) a unified mathematical account that can distinguish between the perceptual illusion and the proprioceptive drift. Thus, going beyond previous computational models that consider the illusion to be static.

In this section, we provide a neuroscientific context for the proposed computational

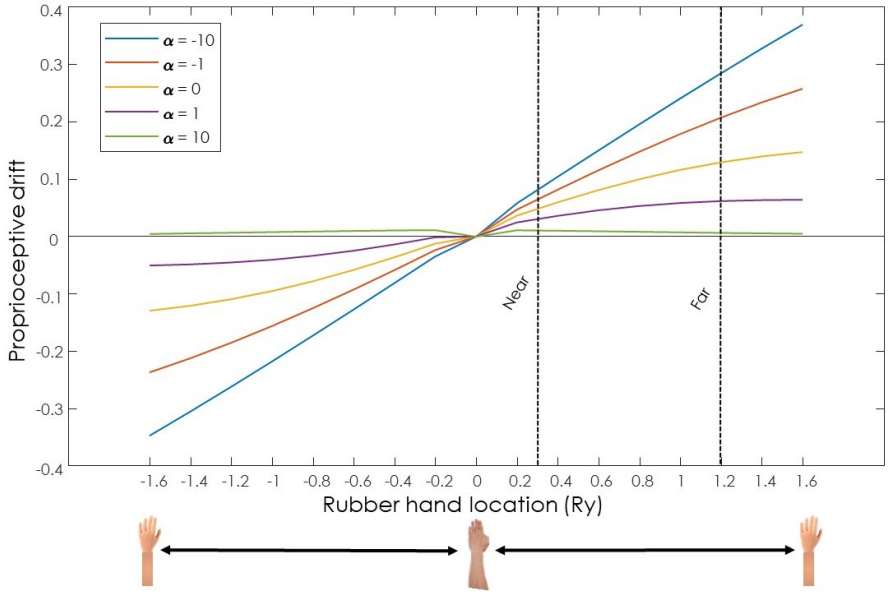


**Fig 7. Dependence of the inter-hand distance and the participant's susceptibility to experience the illusions on the proprioceptive drift.** Inspired on [52], we show a similar trend from previous empirical findings where the modulation of the inter-hand distance and different and subjective experience to the illusion affects the proprioceptive drift. Every line in the plot can be related to individual differences when transiting from $M_1$ non-illusory to $M_2$ illusory models (i.e., susceptibility to experience the illusion). This is modulated through the parameter alpha. Here, $\alpha = 10$ and $\alpha = -10$ are equal to models from equations (2) and (3), respectively. We measured the proprioceptive drift as the average of Sy between the 17th and the 30th second. Note that the dashed lines represent the distances for near and far conditions from Figure (6).

model. First, we highlight the importance of precision of sensory signals, and how this can benefit in studying the modulation of RHI via attention. Next we will speculate about the potential neural region that can provide such a computational mechanism.

## Potential neural underpinnings of the model

There are a few brain regions that are closely linked to the RHI, and body ownership illusion in general: right temporoparietal junction (rTPJ), ventral premotor cortex (VPC), and intraparietal sulcus (IPS) [11, 56]. The main debate is whether body ownership illusion is due to in perceptual-motor coordination (VPC-IPS) or it recruits areas connected to attention and information processing (rTPJ). Here, we focus on the rTPJ, while noticing that is only one of the alternatives.

The connection of rTPJ and RHI has been first reported in [56]. In this study, researchers found that disrupting this region by transcranial magnetic stimulation (TMS) led to worse performance in detecting whether the fake hand belonged to the participant's body. Authors also suggested the rTPJ is crucial for the sense of agency. This idea of connecting sense of agency to rTPJ has also be corroborated in several other works [57–63]. Along with this, functional neurological disorder associated with impaired sense of agency has also been linked to rTPJ [64, 65]. These findings strongly support the view that the rTPJ is central for the tasks involving body-awareness, in particular, the sense of agency. Its connection with the RHI and body awareness suggest

that the rTPJ may have an important role in body-ownership illusions. More importantly, rTPJ is also known to be a part of the attention network focusing on orienting objects [66,67]. This is also nicely corroborated by a lesion study, where a rTPJ lesion led to a higher disengagement from attending to a particular location [68]. There is also a meta-analysis on the rTPJ and attention concluding a significant connection between these two [69]. However, note that there is a debate whether rTPJ should be included in the attention network or not [70] or whether it is stimulus or context-driven [71].

Given these two quite heavily studied findings regarding the rTPJ, and the strong connection between our precision adaptation approach and attention, we speculate that the rTPJ could embed our proposed algorithm. This is to weigh the precision of distinct and competing models of body awareness. Hence, disrupting this region would lead to switching the models because their competing precision values would be altered, along with its related behavioural homologue (under the Bayesian brain perspective) – attention.

## Testing our model

**Model selection versus model averaging.** Although we focused on model selection as the decision-making mechanism for body ownership, we can adapt it to model averaging—the most widely accepted computational account to combine different uncertain models in the brain. Actually, we used model averaging in the modeling of individual differences and the illusion sensitivity. For the model selection, we took inspiration from the affordance competition hypothesis that states that there are two (or more) competing neural correlates for two distinct actions in the parieto-dorsal stream [72–74]. Once a decision is made, only one such neural activation can remain. This leads to the final action referred to as the winner-take-all algorithm. It slightly resembles our modeling effort but the differences have to be highlighted. First, our approach does not consider a 'simple' hand movement, but rather a more complicated model that unites body ownership together via an intricate network of multisensory integration. Thus, different models of body ownership compete, rather than actions that are generated from these models. Ultimately, though, we can consider a similar switch in brain regions showing a similar competition.

**Where to look for precision.** Since we know that attention relates to the density of neural population activity and their decorrelation [75–77] – also known as neural variability – and that attention may be mapped to precision, we speculate that we could see the switch of our two (or more) models in the rTPJ by localizing the density of distinct neural activity and its decorrelation in subregions of rTPJ. At least two such subregions should exist, with their attractor. We hypothesize that the illusion would happen when the high density of activated neurons would decrease, while in the other subregion we could perceive the opposite, with a more random neural activity. This switch of neural density, if observed, would corroborate our model, that the precision maximizing algorithm relates to the sensation of body ownership.

## Limitations

**Uncertain signals versus no signal:** One of the main assumptions of our body-ownership illusion model states that although the real hand is "invisible" it still produces a signal that is highly noisy – from the agent's perspective – rather than not producing a signal at all. This is an important part of the switch between the two models and can be used to analyse the experiments where instead of fully hiding the hand, the degree of visibility is manipulated [13]. We formulated our RHI model to

emphasize that the primary driving mechanism stems from the precision of the signal that triggers the switch between hypotheses or models of reality– following the free energy fundamental assumptions.

**Unstable proprioceptive drift:** In our simulations (Figures 5 and 7) the proprioceptive drift appears to precisely follow the touch of the stimulation, and dissipates when the strokes no longer occur. This, in empirical settings, would result that a participant perceives the proprioceptive drift only during an episode of touch, resulting in a feeling of moving the hand. This is obviously wrong and it is one of our limitations. This, however, can be solved by introducing an additional algorithm that slows down the process of estimating new information rapidly, and rather averages the signal across a long period. This would then lead to a steady belief that the proprioceptive drift is happening in moments when there is no tactile stimulation.

**Temporal synchronicity:** The main limitation of our model is that it does not account for the temporal synchronization [15] of the strokes, a well-studied aspect of the RHI. While we can model the participant sensitivity to the body-illusion through the *alpha* parameter that blends the two competing models, visuotactile spatio-temporal patterns are not modeled. This should be certainly investigated in the future.

**Peripersonal space and implausible body configuration and appearance:** Our model cannot predict the effect of erroneous fake limb configurations (e.g., 180° rotated hands) where the illusion breaks, replicate the large decay on the proprioceptive drift when the rubber hand is outside of the peripersonal space, and account for hand appearance modulators. Only appearance can be modeled through the body-illusion sensitivity parameter. To seamlessly integrate these characteristics into the model, we need to learn these priors. Some work has been done using simulated realistic visual input [26].

**Free energy principle:** The FEP, despite its explanatory power, remains a theoretical framework that may not fully capture the complexity of neural processes. One limitation is its assumption of a single overarching principle governing brain function, which may oversimplify the diverse mechanisms at play. Alternative approaches could provide complementary or even contrasting perspectives. For instance, predictive coding models, while related to FEP, offer a more specific computational framework that might yield more testable hypotheses. Another alternative is the use of reinforcement learning models, which could better account for the role of reward and punishment in shaping the model switching mechanism.

**Future work:** Future research should focus on experimental validation to demonstrate the dynamics of switching between competing models, as proposed in our study. In our context, researchers could create scenarios where multiple states of body ownership are simultaneously induced, allowing only one to predominate. To extend our model, we suggest implementing a probabilistic framework that incorporates competing hypotheses. This could be achieved by using a softmax selection function or through model averaging techniques. Such an approach would enable the study of probabilistic transition dynamics between different competing perceptions of reality. These modifications could potentially explain phenomena such as double-limb and supernumerary limb illusions, providing valuable insights into the mechanisms of body ownership.

# Acknowledgments

FN is funded by the Serotonin & Beyond programme (953327 to F.N.). AM is funded by the EU Metatool project (Grant agreement 101070940) under the EIC Pathfinder program. The research has been partially funded by the DEEPSELF project (467045002 DFG SPP The Active Self).


# Appendix

## 0.1 Related computational models of body illusions

Previous computational models of body illusions approached the topic as a problem of multisensory conflicts, especially the RHI [10, 14, 78] and its virtual reality (VR) version [9, 79]. One of the first seminal works on modeling RHI comes from [12], where a model driven by Bayesian inference is proposed, focusing on the synchronization of stimuli in time. The authors went beyond the optimal integration model [52], by introducing causal inference, where two potential models of reality are presented: either the tactile information is coming from the subject own hand or from the fake hand. However, this account presents body ownership as a static probability computation where no explanation is given on the inner functional mechanism producing the dynamics of the perceptual drifts nor the illusion [11]. Bayesian causal inference was revisited in [13], focusing on the impact of the input noise to the emergence of the RHI.

Another line of research inspired by the predictive processing approach to perception [80–82], yielded in a set of computational models of body illusions that follow the prediction error minimization as the inner mechanism. [10] aimed to investigate how the distance between the rubber and real hands modulates the proprioceptive drift. The work compared the data obtained from a human RHI experiment and a humanoid RHI experiment. It used the learnt generative model of the sensory signals (proprioceptive, visual and visuotactile) to dynamically estimate the location of the real hand. Mathematically, the model used approximated Bayesian inference (i.e., free energy minimization) to infer the body pose given the input signals. Whilst, this work supports the separation of body-ownership illusion and proprioceptive drift, as these seem to be split into two interconnected processes, as suggested by [25, 83], it only provided the computation for replicating the proprioceptive drift and not the body-ownership. In a follow-up model [9, 27], the focus shifted from the passive perceptual process to the active component of the illusion. The model and human evidence suggests that participants actively move their body position (or generate forces) to minimize the discrepancy between proprioceptive signals coming from the real hand, and visual signals coming from the rubber hand. As an aside result, the model explained that perceptual drifts are not necessarily linked to tactile stimulation, but is achievable with visual and proprioceptive signals in VR [84].

The models described assumed that the location of the hand is already processed. The work by [26] introduced deep neural networks to allow the body illusion model to use large-scale visual inputs. The architecture implemented the deep active inference framework for continuous time variables [85]. This model was able to account for the perceptual and active drifts dynamics during the illusion using synthetic images generated from VR. Furthermore, it introduced synchronicity of the input signals as a variable. However, this model did not contribute to the biological underpinnings of body ownership. There are other connectivity based models, such as [29], where the neural network connections are inspired by brain functional segregation.

## 0.2 Precision parametrization

Precision learning follows the online learning of measurement noise precision (or the inverse noise covariance matrix) $\Pi^z$. Intuitively, this refers to the brain's attentional mechanism of gauging the right amount of noise in the environment so as to make inferences about the world under high noise (uncertainties)—see [86, 87]. We postulate that an accurate attentional mechanism (online precision learning) is critical to experiencing the RHI. We use FEP for online precision learning. This follows two steps: modeling and learning. The noise precision $\Pi^z$ is modeled through an exponential parametrization with respect to parameter $\lambda = [\lambda^1\ \lambda^2\ ...\ \lambda^m]^T$ as:

$$\Pi^z = \begin{bmatrix} e^{\lambda^1} & 0... & & \\ 0 & e^{\lambda^2} & 0... & \\ 0 & 0 & e^{\lambda^3} & ... \\ ... & ... & ... & \end{bmatrix}_{m\times m} . \tag{10}$$

This parametrization was chosen to ensure that $\Pi^z$ always remains positive definite ($\Pi^z \succ O$) throughout the estimation. Intuitively, this ensures the noise covariance to be a positive real number.

## 0.3 Code and parameters used for simulations

The MATLAB code [3] for the simulation is available: `https://github.com/ajitham123/precision_RHI/`. This section lists the parameters that were used in our simulations. The generative process uses the model given in Eq (2). The same model is used as $M_1$ with $Ay = 0$ and $Ry = 0.8$, while the model in Eq (3) is used as $M_2$. The $\alpha$ in Eq (9) is set as 10 for low sensitivity to illusion, and -10 for high sensitivity to illusion. The generative process is simulated for 32 seconds with a sampling time of 0.1s. The observation noise for each signal is parametrized using $\lambda = \begin{bmatrix}\lambda^{Av} & \lambda^{Rv} & \lambda^{At} & \lambda^{Py}\end{bmatrix}^T = [9, 7, 8, 10]$ for hand visible, and $\lambda = [-1, 7, 8, 4]$ for hand covered (see Figure (3) for the generated data). The causes (hand strokes) $\mathbf{X}$ (plotted in blue in Figure (5)) comprises of i) four Gaussian bumps centred around $t = 4, 6, 8, 10$ seconds for hand visible case and ii) five Gaussian bumps centred around $t = 20, 22, 24, 26, 28$ seconds for hand covered case. The magnitude of these bumps for $\mathbf{Ea}, \mathbf{Er}, \mathbf{Er}, \mathbf{Sy}$ are $0.5, 1, 1, 0$, respectively (when stroked). The estimation starts with priors on noise parameters (in Eq (5)) as $\eta^\lambda = [10, 10, 10, 10]$ with prior precision $P^\lambda = diag([e^{-4}, e^{-4}, e^{-4}, e^{-4}])$ where $diag(.)$ converts an array into a diagonal matrix. For breaking the precision learning in figure 4, a higher prior precision ($P^\lambda = diag([e^{10}, e^{-4}, e^{-4}, e^{-4}])$) for the signal $Av$ was used. The state estimation starts at 0 with a unit precision. All the learning rates in Section  are set to $k = 4$ for a quick learning. All figures are generated using these parameters as the basic reference for the simulation, unless another value is mentioned for a few variables within the text.

# References

1. Longo MR, Haggard P. An implicit body representation underlying human position sense. Proceedings of the National Academy of Sciences. 2010;107(26):11727–11732.

2. Ehrsson HH. The Concept of Body Ownership and Its Relation to Multisensory Integration. In: The New Handbook of Multisensory Processing. The MIT Press; 2012.Available from: `https://doi.org/10.7551/mitpress/8466.003.0067`.

[3] The code will be made public upon paper acceptance

3. Kilteni K, Maselli A, Kording KP, Slater M. Over my fake body: body ownership illusions for studying the multisensory basis of own-body perception. Frontiers in human neuroscience. 2015;9:141.

4. Hide M, Ito Y, Kuroda N, Kanda M, Teramoto W. Multisensory integration involved in the body perception of community-dwelling older adults. Scientific Reports. 2021;11(1):1581.

5. Ehrsson HH. Multisensory processes in body ownership. Multisensory perception. 2020; p. 179–200.

6. Ehrsson HH, Spence C, Passingham RE. That's my hand! Activity in premotor cortex reflects feeling of ownership of a limb. Science. 2004;305(5685):875–877.

7. Tsakiris M, Hesse MD, Boy C, Haggard P, Fink GR. Neural signatures of body ownership: a sensory network for bodily self-consciousness. Cerebral cortex. 2007;17(10):2235–2244.

8. Ehrsson HH. Bodily illusions. The Routledge Handbook of Bodily Awareness. 2022;.

9. Lanillos P, Franklin S, Maselli A, Franklin DW. Active strategies for multisensory conflict suppression in the virtual hand illusion. Scientific Reports. 2021;11(1):22844.

10. Hinz NA, Lanillos P, Mueller H, Cheng G. Drifting perceptual patterns suggest prediction errors fusion rather than hypothesis selection: replicating the rubber-hand illusion on a robot. In: 2018 Joint IEEE 8th International Conference on Development and Learning and Epigenetic Robotics (ICDL-EpiRob). IEEE; 2018. p. 125–132.

11. Finotti G, Garofalo S, Costantini M, Proffitt DR. Temporal dynamics of the Rubber Hand Illusion. Scientific Reports. 2023;13(1):7526.

12. Samad M, Chung AJ, Shams L. Perception of body ownership is driven by Bayesian sensory inference. PloS one. 2015;10(2):e0117178.

13. Chancel M, Ehrsson HH, Ma WJ. Uncertainty-based inference of a common cause for body ownership. Elife. 2022;11:e77221.

14. Botvinick M, Cohen J. Rubber hands 'feel'touch that eyes see. Nature. 1998;391(6669):756–756.

15. Lanfranco RC, Chancel M, Ehrsson HH. Quantifying body ownership information processing and perceptual bias in the rubber hand illusion. Cognition. 2023;238:105491.

16. Armel KC, Ramachandran VS. Projecting sensations to external objects: evidence from skin conductance response. Proceedings of the Royal Society of London Series B: Biological Sciences. 2003;270(1523):1499–1506.

17. Tsakiris M, Haggard P. The rubber hand illusion revisited: visuotactile integration and self-attribution. Journal of experimental psychology: Human perception and performance. 2005;31(1):80.

18. Guterstam A, Petkova VI, Ehrsson HH. The illusion of owning a third arm. PloS one. 2011;6(2):e17208.

19. Tosi G, Mentesana B, Romano D. The correlation between proprioceptive drift and subjective embodiment during the rubber hand illusion: A meta-analytic approach. Quarterly Journal of Experimental Psychology. 2023; p. 17470218231156849.

20. Costantini M, Haggard P. The rubber hand illusion: sensitivity and reference frame for body ownership. Consciousness and cognition. 2007;16(2):229–240.

21. Holmes NP, Crozier G, Spence C. When mirrors lie:"Visual capture" of arm position impairs reaching performance. Cognitive, Affective, & Behavioral Neuroscience. 2004;4(2):193–200.

22. Holmes NP, Snijders HJ, Spence C. Reaching with alien limbs: Visual exposure to prosthetic hands in a mirror biases proprioception without accompanying illusions of ownership. Perception & psychophysics. 2006;68(4):685–701.

23. Rohde M, Di Luca M, Ernst MO. The rubber hand illusion: feeling of ownership and proprioceptive drift do not go hand in hand. PloS one. 2011;6(6):e21659.

24. Rohde M, Wold A, Karnath HO, Ernst MO. The human touch: skin temperature during the rubber hand illusion in manual and automated stroking procedures. PloS one. 2013;8(11):e80688.

25. Abdulkarim Z, Ehrsson HH. No causal link between changes in hand position sense and feeling of limb ownership in the rubber hand illusion. Attention, Perception, & Psychophysics. 2016;78:707–720.

26. Rood T, Gerven Mv, Lanillos P. A deep active inference model of the rubber-hand illusion. In: International Workshop on Active Inference. Springer; 2020. p. 84–91.

27. Maselli A, Lanillos P, Pezzulo G. Active inference unifies intentional and conflict-resolution imperatives of motor control. PLoS computational biology. 2022;18(6):e1010095.

28. Ernst MO, Banks MS. Humans integrate visual and haptic information in a statistically optimal fashion. Nature. 2002;415(6870):429–433.

29. Zhao Y, Lu E, Zeng Y. Brain-inspired bodily self-perception model for robot rubber hand illusion. Patterns. 2023;4(12).

30. Mumford D. On the computational architecture of the neocortex: II The role of cortico-cortical loops. Biological cybernetics. 1992;66(3):241–251.

31. Feldman H, Friston KJ. Attention, uncertainty, and free-energy. Frontiers in human neuroscience. 2010;4:215.

32. Adams RA, Stephan KE, Brown HR, Frith CD, Friston KJ. The computational anatomy of psychosis. Frontiers in psychiatry. 2013;4:47.

33. Friston K. Computational psychiatry: from synapses to sentience. Molecular psychiatry. 2023;28(1):256–268.

34. Anil Meera A, Novicky F, Parr T, Friston K, Lanillos P, Sajid N. Reclaiming saliency: Rhythmic precision-modulated action and perception. Frontiers in Neurorobotics. 2022;16:896229.

35. Brown H, Adams RA, Parees I, Edwards M, Friston K. Active inference, sensory attenuation and illusions. Cognitive processing. 2013;14:411–427.

36. Jeganathan J, Breakspear M. An active inference perspective on the negative symptoms of schizophrenia. The Lancet Psychiatry. 2021;8(8):732–738.

37. Feeney EJ, Groman SM, Taylor JR, Corlett PR. Explaining delusions: reducing uncertainty through basic and computational neuroscience. Schizophrenia bulletin. 2017;43(2):263–272.

38. Jardri R, Denève S. Computational models of hallucinations. The neuroscience of hallucinations. 2013; p. 289–313.

39. Parr T, Friston KJ. Uncertainty, epistemics and active inference. Journal of the Royal Society Interface. 2017;14(136):20170376.

40. Parr T, Benrimoh DA, Vincent P, Friston KJ. Precision and false perceptual inference. Frontiers in integrative neuroscience. 2018;12:39.

41. Novicky F, Parr T, Friston K, Mirza MB, Sajid N. Bistable perception, precision and neuromodulation. Cerebral Cortex. 2023;34(1):bhad401. doi:10.1093/cercor/bhad401.

42. Novicky F, Offergeld J, Janssen S, Lanillos P. Robotic Active Tactile Sensing Inspired by Serotonergic Modulation Using Active Inference. In: Conference on Biomimetic and Biohybrid Systems. Springer; 2023. p. 33–55.

43. Doya K, Ishii S, Pouget A, Rao RP. Bayesian brain: Probabilistic approaches to neural coding. MIT press; 2007.

44. Friston K, FitzGerald T, Rigoli F, Schwartenbeck P, Pezzulo G. Active inference: a process theory. Neural computation. 2017;29(1):1–49.

45. Friston K. The free-energy principle: a unified brain theory? Nature reviews neuroscience. 2010;11(2):127–138.

46. Parr T, Pezzulo G, Friston KJ. Active inference: the free energy principle in mind, brain, and behavior. MIT Press; 2022.

47. Friston K. The free-energy principle: a rough guide to the brain? Trends in cognitive sciences. 2009;13(7):293–301.

48. Friston K, Thornton C, Clark A. Free-energy minimization and the dark-room problem. Frontiers in psychology. 2012; p. 130.

49. Meera AA, Wisse M. Free energy principle based state and input observer design for linear systems with colored noise. In: 2020 American Control Conference (ACC). IEEE; 2020. p. 5052–5058.

50. Meera AA, Lanillos P. Adaptive Noise Covariance Estimation under Colored Noise using Dynamic Expectation Maximization. In: 2023 62nd IEEE Conference on Decision and Control (CDC). IEEE; 2023. p. 165–171.

51. Meera AA, Lanillos P. Confidence-Aware Decision-Making and Control for Tool Selection. arXiv preprint arXiv:240303808. 2024;.

52. Fang W, Li J, Qi G, Li S, Sigman M, Wang L. Statistical inference of body representation in the macaque brain. Proceedings of the National Academy of Sciences. 2019;116(40):20151–20157.

53. Valenzuela Moguillansky C, O'Regan JK, Petitmengin C. Exploring the subjective experience of the "rubber hand" illusion. Frontiers in human neuroscience. 2013;7:659.

54. Preston C. The role of distance from the body and distance from the real hand in ownership and disownership during the rubber hand illusion. Acta psychologica. 2013;142(2):177–183.

55. Kalckert A, Perera ATM, Ganesan Y, Tan E. Rubber hands in space: the role of distance and relative position in the rubber hand illusion. Experimental Brain Research. 2019;237(7):1821–1832.

56. Tsakiris M, Costantini M, Haggard P. The role of the right temporo-parietal junction in maintaining a coherent sense of one's body. Neuropsychologia. 2008;46(12):3014–3018.

57. Farrer C, Frith CD. Experiencing oneself vs another person as being the cause of an action: the neural correlates of the experience of agency. Neuroimage. 2002;15(3):596–603.

58. Farrer C, Franck N, Georgieff N, Frith CD, Decety J, Jeannerod M. Modulating the experience of agency: a positron emission tomography study. Neuroimage. 2003;18(2):324–333.

59. Schnell K, Heekeren K, Schnitker R, Daumann J, Weber J, Heßelmann V, et al. An fMRI approach to particularize the frontoparietal network for visuomotor action monitoring: detection of incongruence between test subjects' actions and resulting perceptions. Neuroimage. 2007;34(1):332–341.

60. Yomogida Y, Sugiura M, Sassa Y, Wakusawa K, Sekiguchi A, Fukushima A, et al. The neural basis of agency: an fMRI study. Neuroimage. 2010;50(1):198–207.

61. Nahab FB, Kundu P, Gallea C, Kakareka J, Pursley R, Pohida T, et al. The neural processes underlying self-agency. Cerebral cortex. 2011;21(1):48–55.

62. Zito GA, Anderegg LB, Apazoglou K, Müri RM, Wiest R, grosse Holtforth M, et al. Transcranial magnetic stimulation over the right temporoparietal junction influences the sense of agency in healthy humans. Journal of Psychiatry and Neuroscience. 2020;45(4):271–278.

63. Hughes G. The role of the temporoparietal junction in implicit and explicit sense of agency. Neuropsychologia. 2018;113:1–5.

64. Maurer CW, LaFaver K, Ameli R, Epstein SA, Hallett M, Horovitz SG. Impaired self-agency in functional movement disorders: a resting-state fMRI study. Neurology. 2016;87(6):564–570.

65. Bühler J, Weber S, Loukas S, Walther S, Aybek S. Non-invasive neuromodulation of the right temporoparietal junction using theta-burst stimulation in functional neurological disorder. BMJ neurology open. 2024;6(1).

66. Raz A, Buhle J. Typologies of attentional networks. Nature reviews neuroscience. 2006;7(5):367–379.

67. Fan J, McCandliss BD, Fossella J, Flombaum JI, Posner MI. The activation of attentional networks. Neuroimage. 2005;26(2):471–479.

68. Friedrich FJ, Egly R, Rafal RD, Beck D. Spatial attention deficits in humans: a comparison of superior parietal and temporal-parietal junction lesions. Neuropsychology. 1998;12(2):193.

69. Krall SC, Rottschy C, Oberwelland E, Bzdok D, Fox PT, Eickhoff SB, et al. The role of the right temporoparietal junction in attention and social interaction as revealed by ALE meta-analysis. Brain Structure and Function. 2015;220:587–604.

70. Markett S, Nothdurfter D, Focsa A, Reuter M, Jawinski P. Attention networks and the intrinsic network structure of the human brain. Human Brain Mapping. 2022;43(4):1431–1448.

71. Geng JJ, Vossel S. Re-evaluating the role of TPJ in attentional control: contextual updating? Neuroscience & Biobehavioral Reviews. 2013;37(10):2608–2620.

72. Cisek P, Kalaska JF. Neural correlates of reaching decisions in dorsal premotor cortex: specification of multiple direction choices and final selection of action. Neuron. 2005;45(5):801–814.

73. Cisek P. Cortical mechanisms of action selection: the affordance competition hypothesis. Philosophical Transactions of the Royal Society B: Biological Sciences. 2007;362(1485):1585–1599.

74. Sakreida K, Effnert I, Thill S, Menz MM, Jirak D, Eickhoff CR, et al. Affordance processing in segregated parieto-frontal dorsal stream sub-pathways. Neuroscience & Biobehavioral Reviews. 2016;69:89–112.

75. Cohen MR, Maunsell JH. Attention improves performance primarily by reducing interneuronal correlations. Nature neuroscience. 2009;12(12):1594–1600.

76. Mitchell JF, Sundberg KA, Reynolds JH. Spatial attention decorrelates intrinsic activity fluctuations in macaque area V4. Neuron. 2009;63(6):879–888.

77. Arazi A, Yeshurun Y, Dinstein I. Neural variability is quenched by attention. Journal of Neuroscience. 2019;39(30):5975–5985.

78. Limanowski J, Friston K. Active inference under visuo-proprioceptive conflict: Simulation and empirical results. Scientific reports. 2020;10(1):4010.

79. Slater M, Pérez Marcos D, Ehrsson H, Sanchez-Vives MV. Inducing illusory ownership of a virtual body. Frontiers in neuroscience. 2009; p. 29.

80. Rao RP, Ballard DH. Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. Nature neuroscience. 1999;2(1):79–87.

81. Hohwy J. The predictive mind. OUP Oxford; 2013.

82. Clark A. Surfing uncertainty: Prediction, action, and the embodied mind. Oxford University Press; 2015.

83. Abdulkarim Z, Hayatou Z, Ehrsson HH. Sustained rubber hand illusion after the end of visuotactile stimulation with a similar time course for the reduction of subjective ownership and proprioceptive drift. Experimental brain research. 2021;239(12):3471–3486.

84. Maselli A, Slater M. The building blocks of the full body ownership illusion. Frontiers in human neuroscience. 2013;7:83.

85. Sancaktar C, van Gerven MA, Lanillos P. End-to-end pixel-based deep active inference for body perception and action. In: 2020 Joint IEEE 10th International Conference on Development and Learning and Epigenetic Robotics (ICDL-EpiRob). IEEE; 2020. p. 1–8.

86. Reynolds JH, Pasternak T, Desimone R. Attention increases sensitivity of V4 neurons. Neuron. 2000;26(3):703–714.

87. Reynolds JH, Heeger DJ. The normalization model of attention. Neuron. 2009;61(2):168–185.