

Article

More Plausible Models of Body Ownership Could Benefit Virtual Reality Applications [†]

Moritz Schubert *  and Dominik Endres 

Theoretical Cognitive Science, Department Psychology, Philipps-Universität Marburg, 35032 Marburg, Germany; dominik.endres@uni-marburg.de

* Correspondence: moritz.schubert@uni-marburg.de

† This paper is an extended version of our paper published in IEEE VR 2021, 2nd Workshop on Seated Virtual Reality & Embodiment.

Abstract: Embodiment of an avatar is important in many seated VR applications. We investigate a Bayesian Causal Inference model of body ownership. According to the model, when available sensory signals (e.g., tactile and visual signals) are attributed to a single object (e.g., a rubber hand), the object is incorporated into the body. The model uses normal distributions with astronomically large standard deviations as priors for the sensory input. We criticize the model for its choice of parameter values and hold that a model trying to describe human cognition should employ parameter values that are psychologically plausible, i.e., in line with human expectations. By systematically varying the values of all relevant parameters we arrive at the conclusion that such quantitative modifications of the model cannot overcome the model's dependence on implausibly large standard deviations. We posit that the model needs a qualitative revision through the inclusion of additional sensory modalities.



Citation: Schubert, M.; Endres, D. More Plausible Models of Body Ownership Could Benefit Virtual Reality Applications. *Computers* **2021**, *10*, 108. <https://doi.org/10.3390/computers10090108>

Academic Editors: Daniel Zielasko and Bernhard Riecke

Received: 13 June 2021

Accepted: 12 August 2021

Published: 26 August 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Keywords: body ownership; bayesian modeling; bayesian causal inference; virtual reality

1. Introduction

In many virtual reality (VR) applications, the user is represented in the virtual environment by an avatar. Body ownership over the avatar is often helpful, e.g., to increase the perception of presence in the VR. We define body ownership as the experience of a body as one's own. A lack of body ownership would likely lead to a feeling of discomfort and reduce the appeal of the overall user experience. For example, the user might not feel comfortable in their "virtual skin".

A well-founded understanding of the mechanisms underlying the occurrence of body ownership can help VR application designers to create more appealing software for their customers. To this end, developing accurate computational models of body ownership is a promising approach, because they facilitate the prediction of changes in the modeled outcome. If a certain parameter of the model proves especially predictive of body ownership, a designer of embodied experiences might want to pay special attention to the construct this parameter represents.

We assume that the most useful model of this kind would approximate the data-generating process in the real world. In other words, we aim for a generative model of how internal (e.g., neural activity) and external (e.g., sensory input) factors cause body ownership percepts in humans. Our reasons for this assumption are two-fold: first, this aim is very much in line with the general project of body ownership research. In computational terms this research is an attempt to learn about the data-generating process. Second, a generative model generalizes to a much larger number of situations than situation-specific classifiers or similar data-driven approaches. Accordingly, a good approximation of the generating process should avoid both under- and overfitting and can be helpful in a wide variety of applications.

One attempt at finding such a generative model which has been gaining traction recently is the Bayesian Causal Inference of Body Ownership (BCIBO) model [1]. This model assumes a seated user who can change the position of their arms, but not the position of their torso. Therefore, it should be easily applicable to seated VR problems in the real world.

This paper is an extended version of our proceedings article [2] presented at the 2nd Workshop on Seated Virtual Reality & Embodiment. In it, we are going to give a brief overview of the BCIBO model and the experimental paradigm it is trying to explain (Section 2). Our main contribution is an analysis of a flaw in the BCIBO model's assumptions (Section 2.2). We report on our attempts of correcting that flaw (Section 3). Finally, we discuss our findings (Section 4) and speculate on promising future research (Section 4.1).

2. Theory

2.1. The Rubber Hand Illusion

One of the most widespread paradigms to study body ownership is the rubber hand illusion (RHI) [3] (for an illustration see Figure 1). In the classic version of the experiment the participant is seated with one of their arms resting on a table in front of them. A rubber hand is placed on the table in an anatomically plausible position. The real hand is hidden from view and the shoulder covered by a blanket, out of which the rubber hand protrudes. Therefore, at first sight it might look to the participant as if the hand in front of them is their real hand.

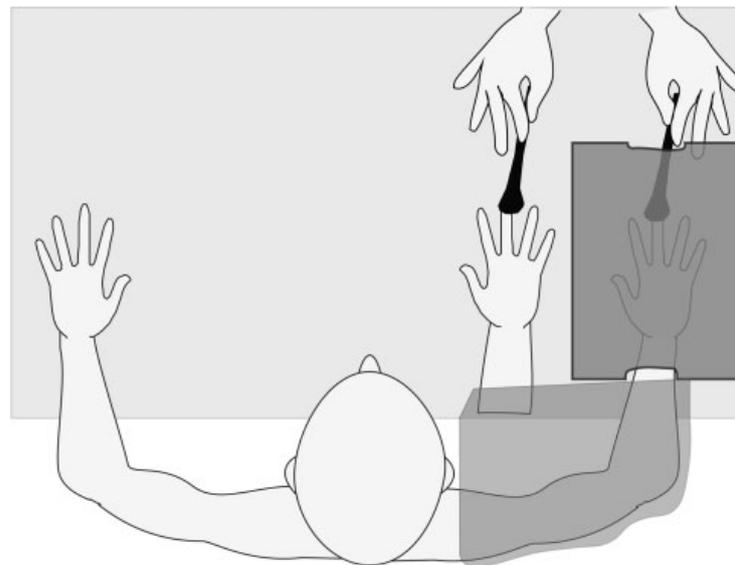


Figure 1. Setup of the classic rubber hand illusion. The participant's real hand is hidden inside a box and a blanket is spread across their shoulder. From the blanket protrudes a rubber hand. Rubber hand and real hand are stroked by the experimenter in synchrony with a brush. The image is from Neustadter et al. [4] and was released under a Creative Commons Attribution Non-Commercial License.

Rubber hand and real hand are stroked in synchrony by the experimenter with a brush. This often results in referral of touch [5], i.e., feeling the touch of the brush on the rubber hand instead of the real one. Most of the time, referral of touch is accompanied by a body ownership illusion (BOI) towards the rubber hand [3,6]. BOI is measured by questionnaire responses, physiological variables and involuntary protective actions towards the rubber hand. In many RHI experiments, this so-called synchronous condition is accompanied by an asynchronous control condition. In the latter, the series of brush strokes on the two hands are out of synchrony. This leads to a lack of BOI [1,3].

2.2. The BCIBO Model

Samad et al. [1] explain the RHI with the BCIBO model, a Bayesian causal inference model applied to the RHI paradigm. Bayesian inference is a statistically optimal method for updating current knowledge considering new observations. In the following we will briefly outline how Bayesian inference works.

Constructs of interest are represented by random variables. In the case of the RHI, the variable H (for hypothesis) indicates the occurrence of a BOI while D refers to sensory data. The Bayesian framework represents the uncertainty inherent in our knowledge about the world in the form of probability distributions over random variables. Perception is the act of updating uncertain knowledge when sensory signals D (here: vision, touch) become available. In statistical terms, this update is an inference. Inference is accomplished using Bayes' theorem:

$$p(H|D) = \frac{p(D|H)p(H)}{p(D)} \quad (1)$$

where $p(H)$, the prior distribution, represents our knowledge of the world before seeing any of the data. The likelihood $p(D|H)$ is the conditional probability of the data under our different hypotheses. The marginal likelihood $p(D)$ is the probability of our data, irrespective of any of the hypotheses under consideration. By multiplying the prior with the likelihood and normalizing it by the marginal likelihood, we arrive at $p(H|D)$, the posterior, which represents our knowledge about the world updated by the (sensory) data.

Bayesian causal inference applies Bayes' theorem to the search for the causes of events (such as sensory input) [7]. If two sensory inputs are assumed to have a common cause, then an optimal inference will integrate both inputs into one percept. For example, let us say that a person sees a dog opening and closing its mouth and at the same time they hear a barking sound coming from the direction of the dog. If they assume a common cause for both sensations, they will bind the auditorily perceived barking to the visually perceived movement of the dog's mouth, and perceive a barking dog.

The Bayesian causal inference framework codifies this search for common causes in the form of a decision between competing models. In the context of the BCIBO model, these competing models are: first, the common cause model (C_1), which supposes a single cause for the sensory percepts. Second, the separate causes model (C_2), which supposes a separate cause for each percept. A high degree of spatiotemporal congruency provides evidence for C_1 . Spatiotemporal disparity provides evidence for C_2 (pages 102–106 in Hohwy [8]). In other words, the closer two percepts are in space and time, the more likely they are assumed to stem from a common cause. Consequently, two spatiotemporally close events are integrated into one percept with high probability.

The BCIBO model [1] explains the RHI as the participant's inference of such a common cause of multisensory input. The model abstracts the sensory input of the brain during the RHI into two categories: spatial information which indicates the position of the rubber and/or real hand. In addition, temporal information which indicates the time points at which the brushes touch both (or one) of the hands. The latter models the synchronicity of the brush stroking the hands. We will refer to these two categories of sensory information as dimensions.

The spatial and temporal dimensions encompass two sources of sensory information, respectively. The spatial information is provided by vision (χ_v) and proprioception (χ_p). A glossary for the abbreviations and symbols used in this paper can be found on page 17. χ_v can only provide information about the rubber hand (since the real hand is hidden from view) and χ_p only about the real hand. Temporal information is provided by vision (τ_v), i.e., seeing the brush strokes on the rubber hand, and tactile signals (τ_t), i.e., feeling the brush strokes on the (hidden) real hand. Again, τ_v can only provide information about the rubber hand and τ_t only about the real hand.

C_1 postulates that the rubber hand causes all the sensory input. C_2 postulates the true state of affairs, namely that the rubber hand causes χ_v and τ_v and the real hand χ_p and τ_t (compare Figure 2). If C_2 is strongly favored, the participant feels as if the real hand

belongs to them and the rubber hand is an external object. If the evidence strongly favors C_1 instead, the participant incorporates the rubber hand into their body model in place of their real hand, leading to a BOI.

In the synchronous condition the congruency in the temporal dimension is very high, because the experimenter applies the brush strokes as synchronously as possible. This provides evidence for a common cause. At the same time, there is a considerable distance between the real and rubber hand (see Figure 1), leading to a discrepancy in the spatial dimension. This is evidence for separate causes. If the evidence in favor of C_1 from the temporal dimension overrides the evidence in favor of C_2 in the spatial dimension, then the participant experiences a BOI.

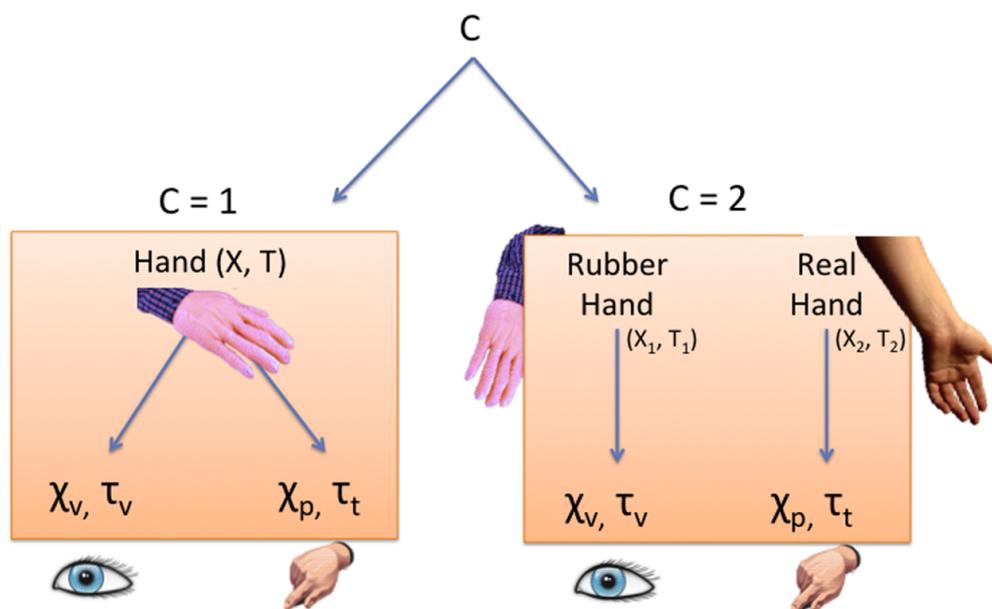


Figure 2. RHI as the decision between a common cause ($C = 1$, left) and two separate causes ($C = 2$, right). If $C = 1$, all sensory input is caused by the rubber hand. If $C = 2$, visual input is caused by the rubber hand and the proprioceptive and tactile inputs are caused by the real hand. Since a common cause only assumes a single hand, there is no need to distinguish between the two hands. Hence, under a common cause the rubber hand is simply referred to as “Hand”. X : position of hand, T : time points of brush strokes, χ_v : spatial visual input, τ_v : temporal visual input, χ_p : proprioceptive input, τ_t : tactile input. The image is from Samad et al. [1] and was released under a Creative Commons Attribution License.

2.3. Related Works

Bayesian causal inference models have been successfully employed to explain a wide variety of cognitive phenomena. For example, the paradigm has been used to model multisensory integration in stimulus localization [7] and speech perception [9]. In these studies, Bayesian causal inference models are usually employed as ideal observers [10], i.e., agents that make the best possible use of sensory information. This is also the assumption of the BCIBO model. Modeling humans as near-optimal agents is justifiable in cases where evolutionary adaptation has solved some important perceptual problem in a near-optimal fashion [7]. Arguably, determining which objects belong to one’s body is such an important problem.

The BCIBO model can account for a variety of well replicated observations in RHI experiments. First, referral of touch (see Section 2.1) is explained by the integration of τ_t into the rubber hand under the common cause model. The integration of χ_p into the rubber hand captures an aspect of the RHI called proprioceptive drift. Proprioceptive drift is defined as the difference between the estimated location of the hand before and during the illusion. Participants typically must indicate the perceived position of their hand with

their eyes closed, hence they must rely solely on proprioception for the task [5]. Commonly, participants in the synchronous condition report a proprioceptive drift towards the rubber hand [1,3,11]. It should be noted though that the drift typically does not “reach” the rubber hand. This is indicated by an average reported proprioceptive estimate that is often 15–30% of the distance between the real and the rubber hand [12].

In addition to referral of touch and proprioceptive drift, the BCIBO model can account for the synchronicity effect: the observation that synchronous stroking induces the illusion while asynchronous stroking does not. This effect has been replicated numerous times [1,13,14]. Since the BCIBO model postulates the congruency on the temporal dimension as the driving factor behind the RHI, it follows that the temporal discrepancy of the asynchronous condition would not induce a BOI. In the synchronous condition the model’s predictions for χ_p are close to the rubber hand, i.e., it predicts proprioceptive drift. In the asynchronous condition they are close to the real hand, i.e., no multisensory integration occurs [1]. Furthermore, the model predicts a BOI probability close to one for the synchronous condition, and a probability close to zero for the asynchronous condition [1].

To our knowledge, there is only one study beside this one that has approached the BCIBO model from a computational perspective: Schürmann et al. [15]. Chancel et al. [16] also implemented a Bayesian causal inference model for body ownership, but it differs from Samad et al.’s [1] model in significant ways—the most prominent of which might be that it only has a temporal and no spatial dimension. In contrast to our paper, which focuses on the posterior distribution of the probability for a common cause, Schürmann et al. [15] looked at the posterior predictive distribution of the sensory signals. A posterior predictive distribution describes the predictions of future data given a model’s posterior.

Another difference between our study and Schürmann et al. [15] is that we focused on the RHI, while they applied the BCIBO model to the rubber foot illusion [17,18]. As the name suggests, rubber foot illusion experiments try to induce body ownership over a rubber foot instead of a rubber hand. However, in both cases synchronous visuotactile stimulation is usually the driving factor behind the illusion.

Schürmann et al. [15] adapted the BCIBO model [1] to the rubber foot illusion and termed it the uniform model. They compared it with an empirically informed model. For the latter they sampled the mean of χ_p ’s sensory prior from a real-world data set [19], while keeping the standard deviation constant and identical to Samad et al. [1]. Another data set taken from Flögel et al. [18] provided the ground-truth proprioceptive drift. They compared the posterior predictive distributions of the position of the rubber hand (i.e., X , see Figure 2) of the two competing models with the empirical distribution of Flögel et al. [18]. The empirically informed model strongly outperformed the uniform model, as indicated by Bayes factors. The uniform model (i.e., BCIBO model) in its current form overestimated both the strength (i.e., the mean) and the precision of the proprioceptive drift as reported in Flögel et al. [18].

2.4. Specification of the BCIBO Model

In this subsection we are going to describe the BCIBO in greater detail, to provide a basis for our modifications of the model.

As explained in Section 2.2, if the probability of C_1 is high, the model predicts the occurrence of a BOI. The posterior probabilities of C_1 and C_2 can be calculated by applying Bayes’ Theorem:

$$p(C|\chi_v, \chi_p, \tau_v, \tau_t) = \frac{p(\chi_v, \chi_p, \tau_v, \tau_t|C)p(C)}{p(\chi_v, \chi_p, \tau_v, \tau_t)} \quad (2)$$

where C is a binary variable with $C = C_1$ indicating a common cause and $C = C_2$ indicating separate causes.

The BCIBO model represents the hands’ perceived positions (χ_v and χ_p) in millimeters on a horizontal line relative to the body midline. It is assumed that the body and the table are roughly parallel to each other. The perceived timing of the brush stroke sequence (τ_v and τ_t) is represented by the time of the first brush stroke (in milliseconds) after the

beginning of the trial. Assuming that all the brush strokes are separated by the same time interval (e.g., 1000 milliseconds), the time point of the first brush stroke provides enough information to represent the entire time series of strokes. The closer τ_v to τ_t , the higher the synchronicity of the brush strokes.

In the following we are going to list all the distributions that are part of the model and establish some other important terminology. We will also interpret what these distributions mean on a psychological level.

X and T denote the position of a hand and the time point of the first brush stroke, respectively. The likelihoods for the spatial dimension are $p(\chi_v|X)$ and $p(\chi_p|X)$ and the ones for the temporal dimension are $p(\tau_v|T)$ and $p(\tau_t|T)$. On a psychological level, these likelihoods represent our predictions about the sensory input given our knowledge about the state of the world. For example, $p(\chi_v|X)$ can be read as “given that my hand is at position X I expect visual input in the shape of a hand at the position χ_v , with probability p ”. Put more plainly, if I think that my hand is at a certain position on the table in front of me, then I expect to see a hand there.

Sometimes in this article it will be important to distinguish between the likelihoods under C_1 and the ones under C_2 . Recall that C_1 presumes there to be only a single position X (i.e., a single hand) and a single time point T (i.e., a single brush touching the hand). In accordance with this, we termed likelihoods under C_1 $p(\chi_v|X_{\text{hand}})$, $p(\chi_p|X_{\text{hand}})$, $p(\tau_v|T_{\text{hand}})$ and $p(\tau_t|T_{\text{hand}})$, where “hand” stands for the single hand that is assumed under C_1 (see Figure 2). C_2 presumes two separate positions X_{rub} and X_{real} and two separate time points T_{rub} and T_{real} , where “rub” refers to the rubber hand and “real” to the real hand. Accordingly, the likelihoods under C_2 are denoted as follows: $p(\chi_v|X_{\text{rub}})$, $p(\chi_p|X_{\text{real}})$, $p(\tau_v|T_{\text{rub}})$ and $p(\tau_t|T_{\text{real}})$.

We call the prior distributions of X and T the sensory priors. In psychological terms, they refer to the expected positions of one’s hand and the time points at which one expects touch events on the hand to occur. We denote the spatial sensory prior under the common cause model as $p(X_{\text{hand}}|C_1)$ and the temporal prior as $p(T_{\text{hand}}|C_1)$. We denote the spatial sensory priors under the separate causes model as $p(X_{\text{real}}|C_2)$, $p(X_{\text{rub}}|C_2)$ and the temporal sensory priors as $p(T_{\text{real}}|C_2)$, $p(T_{\text{rub}}|C_2)$.

Finally, the prior of the two models is called $p(C)$. $p(C_1) = p(C = C_1)$ denotes the prior probability of the common cause and $p(C_2) = p(C = C_2)$ the prior probability of the separate causes model. We will sometimes refer to this distribution as the model prior. The psychological interpretation of $p(C)$ is the tendency to assume that all hand-shaped objects in spatial proximity belong to one’s own body (C_1) or not (C_2).

Samad et al. [1] used Gaussians for all the distributions listed above except the model prior. This decision was probably made for both theoretical and practical reasons, since Gaussians allow for comparatively easy algebraic manipulation. For the model prior, they used a Bernoulli distribution with $p = 0.5$, meaning they assumed equal a priori probability for both hypotheses. Samad et al. [1] strove to choose “realistic values” (page 6) for all parameter values and—for the most part—succeeded in this endeavor.

All the σ values (i.e., standard deviations) for the likelihoods were based on empirical results. σ of $p(\chi_p|X)$ was set to 15 mm [20,21] and the σ s of $p(\tau_v|T)$ and $p(\tau_t|T)$ were set to 20 ms [22] respectively. The standard deviation of $p(\chi_v|X)$ was based on the visual precision of 0.36 degrees reported by van Beers et al. [21]. Samad et al.’s [1] own RHI setup had a distance of $\sim 35 - 45$ cm between the participant’s eye and the rubber hand, which in accordance with van Beers et al. [21] translates to a standard deviation of a couple of millimeters. Samad et al. [1] settled on $\sigma = 1$ mm for $p(\chi_v|X)$ and pointed out that the predictions of the model are affected very little by the exact value of this parameter.

The likelihoods inherit their μ value (i.e., mean) from their respective prior. These μ values are derived from the characteristics of the experimental setup. $p(X_{\text{rub}}|C)$, the prior distribution of the rubber hand’s position, has a mean 160 mm away from the body’s midline, which is a position commonly used in RHI experiments [1]. In a review of methodological variability in the RHI, Riemer et al. [23] reported a typical distance of 15 cm,

i.e., very close to the 16 cm of the BCIBO model. For $p(X_{\text{real}}|C)$, the mean is 320 mm, which is equivalent to the placement of the real hand typically found in RHI experiments [1]. Finally, in the synchronous condition the time points of the first brush strokes are both set to 0 ms, i.e., the brush stroking starts at the same moment as the experimental trial.

By setting the sensory priors' mean values to the actual values of the experimental setup we are using an informed prior [15]. This contrasts with Körding et al. [7] who first proposed the Bayesian causal inference model. They used an uninformed prior, meaning that they set the sensory priors' mean values to 0. They did this to implement a "bias to perceive stimuli straight ahead" (page 3 in Körding et al. [7]). In the context of the RHI this would translate to a bias to perceive stimuli close to the midline.

Schürmann et al. [15] have argued that it is more appropriate to use an informed prior, because humans constantly update their internal representations based on sensory input. From this perspective, it is likely that by the time of the brush stroke onset the participants have inferred the correct position of the hands. Since participants have no idea when the brush strokes are going to set in, this updating can only occur on the spatial, but not the temporal dimension. Hence, we use an informed prior for the spatial and an uninformed prior for the temporal dimension.

Samad et al. [1] chose a "large number" (page 6) as the standard deviation σ for all sensory priors to approximate a uniform distribution. The exact value is not mentioned in the paper, but according to private correspondence it was 10^{35} mm|ms ("the parameters I used for the spatial and temporal prior's variances were extremely large (1e35 each)", M. Samad, personal communication, 5 March 2021). We use "mm|ms" to indicate "millimeter or milliseconds".

2.5. Critique of the Model

We criticize Samad et al. [1] for their choice of the sensory priors' width, because we maintain that a model attempting to approximate the data-generating function of an aspect of human cognition should use psychologically plausible values for its parameters. 10^{35} is an unimaginably large number for humans and therefore it is implausible that such a number would be used in computations in the human mind, when body part placement is concerned. To put the magnitude of this number into perspective: On the spatial dimension of the model, 10^{35} mm is around 1000 times larger than the length of the observable universe (Bars et al. [24], page 27), and on the temporal dimension 10^{35} ms is several orders of magnitude larger than the age of the universe. On top of this, a standard deviation covers only around 68% of a normal distribution, i.e., the values we could reasonably expect with this prior are even larger.

Bayesian models have been criticized for being underconstrained. Jones and Love [25] point out that without proper constraints Bayesian models can fairly easily be fitted to empirical data. According to them often "the prior is chosen ad hoc, providing substantial unconstrained flexibility to models that are advocated as rational and assumption-free" (Jones and Love [25], page 174). Bowers and Davis [26] have also criticized Bayesian models for their flexibility, pointing out the danger of them being mere ad hoc "just so" stories without any explanatory potential.

We agree with the need for constraints to guide Bayesian modeling and pose the psychological plausibility of the model's parameter choices as one such constraint. We do not suggest that this is the only relevant criterion. For example, experiments that test hypotheses derived from the BCIBO model are crucial for its further development. Nonetheless plausibility is a relevant factor, especially because the authors of the model seemed to have adhered to it in the selection of all parameter choices except for the sensory priors' widths [1].

Given these overly wide priors, we think that the model is need of revision. The goal of our revision is to reduce the widths of the sensory priors to a plausible, human-level scale while maintaining the agreement with empirical results. Thus, in this study we are

going to present our exploratory attempts to overcome the implausibility of the sensory priors' width in the BCIBO model.

We will keep the structure of the original model and change the values of the distributions' parameters. The distributions of the model can be grouped into likelihoods, sensory priors and the model prior. As we pointed out above (see page 6 in Section 2.2), Samad et al. [1] put the parameter settings of the likelihoods on firm theoretical ground. Hence, we see no justification for changing their parameter values.

Instead, we are going to discuss the effect of changing the parameter values of the sensory priors (Section 3.1) and the model prior (Section 3.3) on the predictions of the model. Next to these changes of parameter values, we are also going to consider a more drastic change, namely exchanging distributions of the model while keeping the relationship between these distributions (i.e., the structure) intact. Specifically, we will exchange the sensory priors' astronomically wide normal distributions with normal distributions truncated to more sensible bounds. We are going to explore whether this change yields empirically sound predictions in Section 3.2.

3. Results

All our results reported in this section were computed with the programming language Python [27–29], version 3.9.6. We used an open-source language to make the model more accessible to the scientific community. We released our code as open-source under the MIT license on a repository hosted by the University of Marburg at <https://doi.org/10.17192/fdr/66.2> (tagged as version 3), accessed on 12 August 2021. Included in the repository are also files for recreating the virtual environment in which the code was run. For increased transparency, we included the randomizer seed we used for the generation of all the results presented in this paper. Furthermore, we included csv files containing the exact results for all simulation runs mentioned in this section. We indexed these files in the Supplementary Materials Section as data S1, data S2, etc. and will refer to them below whenever their contents are summarized.

3.1. Change in the Sensory Priors

Samad et al. [1] ran the model for the different levels of a distance factor d . Distance refers here to the distance between the real and rubber hand. The levels of the factor were $d_i \in [160, 180, \dots, 340, 360]$ mm, i.e., the lowest was 160 mm and the distances increased by 20 mm until they reached a maximum of 360 mm.

Lloyd [30] found that an increased distance between the real and rubber hand leads to a decrease in body ownership. Samad et al. [1] computed the posterior probability of C_1 for the distance factor (s. Figure 3, left) and found results similar to Lloyd [30]. An increase in the distance factor level can be interpreted as placing the rubber hand further and further away from the real hand across different experimental conditions, similar to Lloyd's [30] setup.

We attempted to replicate Samad et al.'s [1] simulation by running the model for the same distances between the real and rubber hand. In addition to this distance factor, we also introduced a σ factor, whose levels encompass different widths for the sensory priors. We included this second factor to test whether the model can predict empirical results for σ s smaller than 10^{35} mm|ms. The levels of the factor were $\sigma_i \in [10^0, 10^5, \dots, 10^{30}, 10^{35}]$ mm|ms, i.e., we started with 10^0 (i.e., 1) mm|ms and increased the exponent in steps of 5 until we reached Samad et al.'s original value of 10^{35} mm|ms.

For each combination of factor levels, we sampled $N = 10,000$ artificial datapoints from the likelihood distributions $\chi_v \sim \mathcal{N}(320 - d_i, 1)$ mm, $\chi_p \sim \mathcal{N}(320, 15)$ mm and both $\tau_v, \tau_t \sim \mathcal{N}(0, 20)$ ms. These means and standard deviations are derived from experimental data as explained on page 6 in Section 2.2.

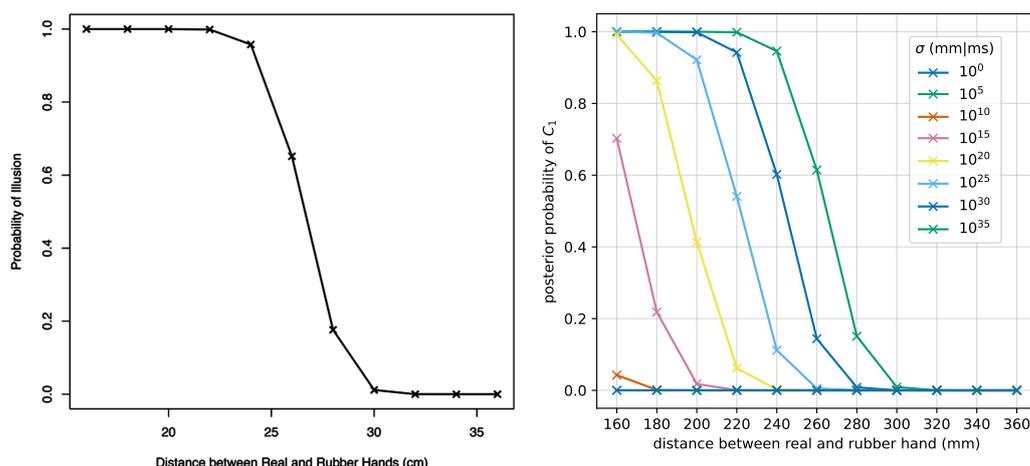


Figure 3. (Left): posterior probability of C_1 for different distances between real and rubber hand as presented by Samad et al. [1]. The image was released under a Creative Commons Attribution License. (Right): posterior probability of C_1 for different distances between real and rubber hand and across different magnitudes of σ value for the sensory priors.

As stated above, the distance factor simulated moving the rubber hand away from the real hand. Therefore, the visual input across the different levels of the distance factor was calculated by $320 - d_i$, i.e., by subtracting the distance from the position of the real hand. Explicitly, the μ values of the χ_v distribution were $\mu_i \in [160, 140, \dots, -20, -40]$ mm. In terms of the experimental setup, this means that the rubber hand moved closer and closer to the participant's body's midline and eventually crossed it, as indicated by μ_i taking on negative values.

Under a common cause (C_1), an observer would expect the visual signals of their own hand to be a reliable source of information about the actual hand position. Hence, we chose the mean of $p(X_{\text{hand}}|C_1)$ equal to the mean of the data-generating distribution of χ_v . Under separate causes (C_2), it is less clear which prior expectations one should have about the visual signals emitted by the rubber hand. Here, we chose the mean $p(X_{\text{rub}}|C_2)$ equal to the mean of the generating distribution of χ_v , too. For a more formal and concise version of the model specifications outlined above see Appendix A.1.

We treated the samples as sensory input across N trials and calculated $p(C_1|D)$, i.e., $p(C = 1|\chi_v, \chi_p, \tau_v, \tau_t)$ (see Equation (2)) for each trial. As a point estimate, we took the mean of $p(C_1|D)$ across the entire sample. The results can be seen in Figure 3 (right) and in data S1. The standard errors of the mean (SEMs) for every factor combination were all below 0.002, thus we did not draw them in the graphs.

As can be seen in Figure 3 (right), our results for the σ value used by Samad et al. [1], 10^{35} mm|ms, closely resemble their results (compare Figure 3, left), indicating a successful re-implementation of the BCIBO model.

Furthermore, Figure 3 (right) shows that the posterior probability of C_1 for all distances declines with smaller choices of the prior's σ value. To be in line with empirical results [30], a good model of body ownership should predict high chances of a BOI occurring for a 160 mm distance. However, for $\sigma = 10^{10}$ mm|ms the chance of experiencing a BOI at 160 mm is below 0.05. To illustrate the magnitude this 'small' prior, consider that 10^{10} mm is equivalent to 10,000 km, longer than the Great Wall of China, i.e., still a very implausible presupposition for the location of one's hand in space relative to one's body.

We ran the model for a σ value at a human scale, 10^4 mm|ms (see data S2, all SEMs < 0.001). On the spatial dimension this translates to $\sigma = 10$ m and on the temporal to $\sigma = 10$ s. The resulting values for $p(C_1|D)$ were tiny ($< 10^{-7}$), indicating virtually no sense of body ownership.

3.2. Truncated Model

At this point we would like to remind the reader that Samad et al. [1] used the same σ value for all the sensory priors. Since we demonstrated in the previous subsection that systematically narrowing such a “one size fits all” prior down to a psychologically plausible scale yields unsatisfactory results, a new approach seems in order. One option for reducing the widths of the sensory priors down to a psychologically plausible scale is to truncate their normal distributions. A truncated normal distribution is a normal distribution that is cut off at the two ends of an interval, such that the probability of a value outside of this interval is zero.

Truncating the sensory priors therefore means that the model will deem any sensory data-generating processes outside of this interval impossible. In psychological terms this could be understood as higher levels of cognitive processing flat out rejecting any processed sensory signals that are incongruent with its model of the world.

We now turn to the question which intervals should be chosen for the sensory priors’ truncation bounds. We argue that the “one size fits all” approach should be abandoned. Instead, we assert that the sensory priors should differ across sensory modalities. Below we are going to suggest truncation bounds for the sensory priors in the BCIBO model. All these bounds are on a human scale and well below 10^{35} mm|ms.

3.2.1. Truncation Bounds

Proprioceptive Input

Although it is admittedly difficult to define reasonable priors for some of the sensory modalities represented in the BCIBO model (see below), there is one exception: χ_p , the proprioceptive input. Under normal circumstances it is impossible for proprioceptive input to indicate a position of the hand outside of arm’s reach. Hence, the truncation boundaries for the prior on χ_p ’s likelihood should correspond to the reach of one’s arm.

It should be noted that by truncating the proprioceptive prior this variation of the model is not able to account for certain abnormal experiences of embodiment outside of the bounds of proprioception. One example for this is Kilteni et al.’s [31] very long arm illusion, in which the authors induced ownership over an elongated virtual arm in participants. However, since these kinds of experiences usually only occur in artificial situations or in atypical states of consciousness, we think this limitation is not relevant for our intended application.

We assumed arm span to be roughly equivalent to height in many humans [32]. It should be noted that this is a vast simplification for the sake of the model. In reality, this relationship depends on characteristics such as sex and ethnicity [33]. We took the average height of Germans as a proxy value. According to the Federal Statistical Office of Germany the average height in the German population was ≈ 1.7 m in 2017 [34]. In accordance with this number, we chose $[-850 \text{ mm}, 850 \text{ mm}]$ as the truncation bounds for proprioception.

Under C_1 χ_p and χ_v share the same prior, $p(X_{\text{hand}}|C_1)$, because this hypothesis assumes that there is only a single location (i.e., a single hand). This means that the prior for the visual input is also cut off at arm’s length, because expecting to see a hand outside of arm’s reach is incompatible with a healthy internal body model. However, under C_2 χ_p and χ_v are independent of one another. Therefore, we used the same proprioceptive truncation boundaries of $[-850 \text{ mm}, 850 \text{ mm}]$ for $p(X_{\text{real}}|C_2)$, but chose different boundaries for $p(X_{\text{rub}}|C_2)$ (see below).

To summarize, let a be the distribution’s lower truncation bound and b the upper truncation bound. We denote a truncated normal distribution as $\mathcal{N}(\mu, \sigma, [a, b])$. Then, under the truncated model $p(X_H|C_1) = \mathcal{N}(160 \text{ mm}, 10^{35} \text{ mm}, [-850 \text{ mm}, 850 \text{ mm}])$ and $p(X_{\text{real}}|C_2) = \mathcal{N}(160 \text{ mm}, 10^{35} \text{ mm}, [-850 \text{ mm}, 850 \text{ mm}])$.

Spatial Visual Input

Although the spatial visual prior under C_1 is coupled with the proprioceptive prior, these two are separated under C_2 . An attempt to find a reasonable σ value for $p(\chi_v|X_{\text{rub}})$ is

difficult. Perhaps most importantly, the environment in which the experiment is conducted in must be taken into account. If objects (e.g., a room's walls) block the participant's view, this sets a natural boundary for where the participant would be able to spot the rubber hand. Hence one option would be to truncate $p(\chi_v|X_{\text{rub}})$ to the distance between the participant's midline and the walls of the room.

For the purposes of the truncated model, we assumed that the participant is seated in the middle of a room and chose 2000 mm (i.e., 2 m) as the distance to the walls on either side. We realize that this choice is somewhat arbitrary. We would like to point out that the main point of truncating the sensory priors is to arrive at widths that are on a scale that humans deal with regularly. Furthermore, when trying to predict the results of a concrete experiment the spatial visual boundaries could be set to the actual distances between the participant's midline and the walls of the room.

To summarize, $p(X_{\text{rub}}|C_2) = \mathcal{N}(320 \text{ mm}, 10^{35} \text{ mm}, [-2000 \text{ mm}, 2000 \text{ mm}])$ in our truncated version of the model.

Temporal Input

The temporal prior refers to an extraordinarily abstract concept: the time the participant expects to wait until they receive the first brush stroke. If the participant had already experienced a couple of trials (e.g., as part of a training block), it would be quite easy to define a sensory prior: Its mode should be close to the average onset times in the previous trials and its precision should depend on the number of previous trials with more trials leading to higher precision. However, since we are trying to model a participant without any previous exposure to the experiment, we do not consider this approach to be a good solution.

Without presuming previous experience, it is not easy to argue for a sensible σ value for τ_v and τ_t . On the other hand, it is far easier to discard specific suggested priors for being too wide. For example, a time interval longer than an hour seems to be unlikely for a stimulus with such a low valence as a brush stroke. We therefore chose 3,600,000 ms (i.e., 1 h) as the truncation bounds.

Although for the spatial dimension the lower bound a and upper bound b were equidistant to 0, doing the same on the temporal dimension would lead to a prior that assigns non-zero probability to brush strokes in the past, which is incompatible with the trial starting at time zero. We therefore set the lower bound for the temporal sensory priors to 0. To summarize, under the truncated model:

$$\begin{aligned} p(T_H|C_1) &= \mathcal{N}(0 \text{ ms}, 10^{35} \text{ ms}, [0 \text{ ms}, 3.6 \times 10^6 \text{ ms}]) \\ p(T_{\text{rub}}|C_2) &= \mathcal{N}(0 \text{ ms}, 10^{35} \text{ ms}, [0 \text{ ms}, 3.6 \times 10^6 \text{ ms}]) \text{ and} \\ p(T_{\text{real}}|C_2) &= \mathcal{N}(0 \text{ ms}, 10^{35} \text{ ms}, [0 \text{ ms}, 3.6 \times 10^6 \text{ ms}]). \end{aligned}$$

3.2.2. Simulation Run

We ran the truncated model for the same distance \times sigma factor levels described in Section 3.1 (see data S3) and compared it with the original version of the model. For a concise description of the truncated model see Appendix A.2.

Figure 4 shows the results of the original model on the left and the ones from the truncated model on the right. All SEMs were <0.002 .

It should be noted that the distances displayed on the x axis in Figure 4 deviate from the ones displayed in Figure 3. Figure 4 displays the distance values [20, 40, ..., 180, 200] mm, while Figure 3 displays [160, 180, ..., 340, 360] mm. As can be seen in Figure 4 (right), the truncated model predicts posterior probabilities of C_1 very close to 0 (i.e., <0.001) for distances ≥ 140 mm across all considered σ values. However, questionnaire mean scores indicating an RHI have often been reported for distances of 150 mm [11,35–37]. Hence, these results show that truncating the sensory priors of the BCIBO model with intervals on a human scale strongly decreases its agreement with empirical results.

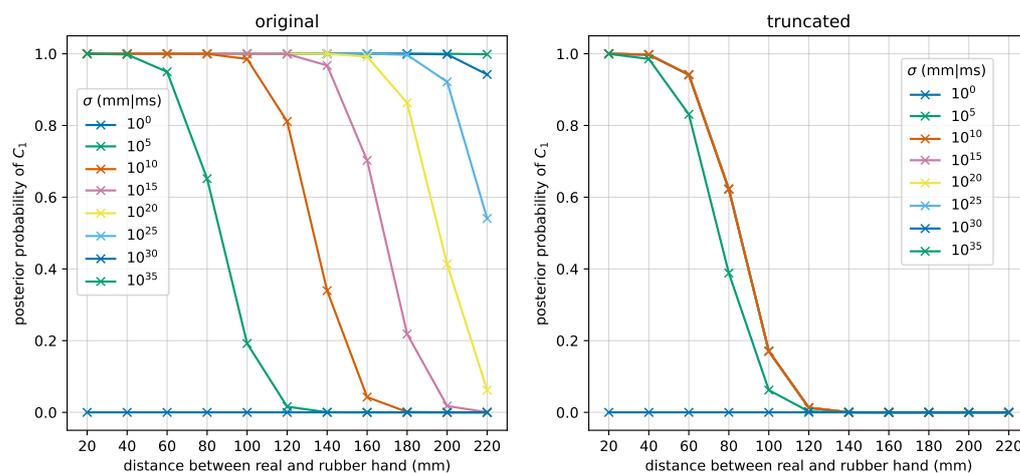


Figure 4. (Left): posterior probability of C_1 for the original model. The parameter settings of the model are equivalent to the one shown in Figure 3 (right), but the x axis has been shifted to left, showing the distances [20, 40, ..., 180, 200] mm instead of [160, 180, ..., 340, 360] mm. (Right): posterior probability of C_1 for the truncated model. For $\sigma > 10^{10}$ mm|ms, the lines are on top of each other.

In Figure 4 (right) the lines for $\sigma \geq 10^{15}$ mm|ms are hidden beneath the line for $\sigma = 10^{10}$ mm|ms, because their posterior probabilities are almost identical. The reason for this is that $\sigma = 10^{10}$ mm|ms far exceeds even the widest bounds in the truncated model, which are $[-3.6 \times 10^6 \text{ ms}, 3.6 \times 10^6 \text{ ms}]$. As a result, all the sensory priors for $\sigma \geq 10^{10}$ mm|ms in the truncated model are very close to being uniform. Under the truncated model, increases in σ beyond 10^{10} mm|ms only lead to microscopic changes in $p(C_1|D)$, which can no longer be displayed in the plot.

3.3. Change in the Model Prior

We were curious how strongly the magnitude of $p(C_1)$ influences the predictions of the model. Specifically, we wanted to determine whether increasing $p(C_1)$ would allow us to decrease the width of the astronomically wide sensory priors. Samad et al. [1] modeled the prior probability of C_1 by a Bernoulli distribution with success probability 0.5, i.e., they used an uninformed prior. In everyday life, the hands in front of us are nearly always our own hands, which results in a $p(\text{oste})\text{rior } p(C_1) \approx 1$. Therefore, we think that one can reasonably assume a value for $p(C_1)$ that is close to one. Again, for a formal description of the simulation runs discussed in this subsection see Appendix A.3.

Figure 5 shows the original uninformed prior and a prior very close to one ($p(C_1) = 0.99$) side by side (for the latter see data S4, all SEMs < 0.002). As can be seen, some of the individual values change noticeably. For example, the posterior probability of a BOI for a distance of 180 mm for $\sigma = 10^{15}$ mm|ms increases by 13 percentage points (see data S6). However, overall the increase in the posterior probability of C_1 is not enough for agreement with empirical data using plausible prior widths. This is demonstrated by the posterior probability values for $\sigma = 10^5$ mm|ms being visually indistinguishable from 0% in Figure 5 (right).

Although asymptotically increasing the value of $p(C_1)$ towards 1 increases the posterior probabilities considerably, even a value as close to 1 as $1 - 10^{-16}$ only brings the posterior probability of C_1 for $\sigma = 10^5$ mm|ms up to 37% (see data S5, all SEMs < 0.002). At this level, the model prior has reached a value nearly as unbelievable as a sensory prior width of 10^{35} mm|ms. We can therefore conclude that increasing the prior probability of C_1 is not sufficient for achieving psychological plausibility.

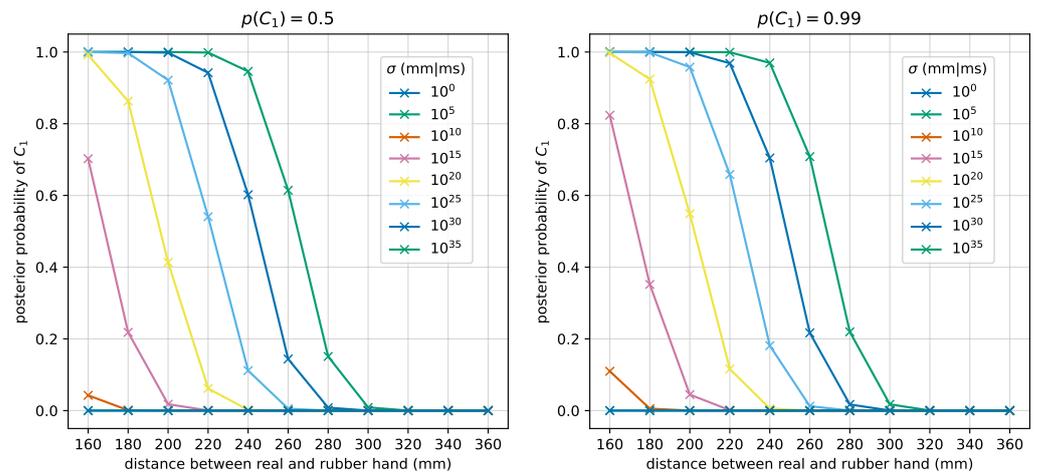


Figure 5. (Left): posterior probability of C_1 for a model prior of $p(C_1) = 0.5$. The contents of the plot are equivalent to Figure 3 (right) and are reproduced here to enable a visual comparison. (Right): posterior probability of C_1 for a model prior of $p(C_1) = 0.99$.

4. Discussion

The ways of adjusting the BCIBO model in its current form can be sorted into three categories: changing the likelihoods, changing the sensory priors or changing the model prior.

We already discussed that Samad et al. [1] set the choice of the likelihood's σ parameters on firm theoretical grounds (see page 6 in Section 2.2). In addition, the mean parameters of the sensory priors represent concrete facets of the experimental setup (e.g., the position of the rubber hand). Hence, we believe that on the level of the likelihoods the model should not be changed.

Truncating the sensory priors to reasonable widths (see Section 3.2) actually worsens the model's agreement with empirical results [30]. Finally, increasing the prior probability of C_1 (see Section 3.3) cannot fix the adverse effects of choosing sensible values for the sensory prior's σ values without introducing another implausible parameter setting in the form of a $p(C_1)$ that is unreasonably close to one.

4.1. Limitations and Future Work

We stated in Section 3.2 that the truncated normal distributions with very high standard deviations come close to uniformity. However, technically they are not uniform distributions. Hence, strictly speaking we did not implement Samad et al.'s [1] stated goal of using uniform sensory priors.

We tried our best to come up with sensible boundaries for the sensory priors of the truncated model, but could only make truly empirically informed decisions for $p(\chi_p|X)$ (arm's length) and $p(\chi_v|X)$ (horizontal distance to the nearest visual obstacle, e.g., a wall). To be fair, trying to design experiments that could empirically inform the sensory priors $p(\tau_v|T)$ and $p(\tau_t|T)$ poses quite the challenge. Presuming we view the temporal dimension as modeling the discrepancy between the brush strokes, $p(\tau_v|T_{\text{rub}})$ and $p(\tau_t|T_{\text{real}})$ could be assessed in a round-about way: At the start of the experiment the participant could be asked a question along the lines of "We are going to each stroke the rubber hand and your real hand with a brush. How large do you expect the discrepancy between the two brush strokes to be?". In the BCIBO model either τ_v or τ_t can be set arbitrarily. What is actually relevant for the computation of the model is the difference between τ_v and τ_t . Hence, the predicted difference by the participant could be used to set a prior distribution for $p(\tau_v|T_{\text{rub}})$ and $p(\tau_t|T_{\text{real}})$. The main problem with this approach is that to our knowledge RHI participants are not typically instructed about the exact procedure of the experiment. Hence, announcing the brush strokes through asking the above question would confound the experiment.

Admittedly, the experimental design described above is quite peculiar. We discussed it to showcase the difficulties of putting some of the components of the BCIBO model such

as the sensory priors on firm empirical grounds. It seems to us that if at all, these difficulties can only be overcome by clever experimental designs that probe for these components in indirect ways.

The changes to the model considered in Section 3 are all quantitative in nature, i.e., they change the values of the model's parameters while preserving its overall structure. All these changes led to unsatisfactory results. Hence, we think that future research should explore qualitative changes to the model in the form of additional likelihoods offering new sources of sensory evidence.

Litwin [38] agrees with this assessment, but for different reasons. He points out that the BCIBO model in its current form cannot account for certain empirical observations. According to the model, having high proprioceptive precision increases the evidence for the real hand being a separate cause. As a result, participants with high proprioceptive precision should be less prone to accepting a common cause and experiencing the illusion. However, Motyka and Litwin [39] could not find evidence for this hypothesis.

Litwin [38] concludes that the BCIBO model in its current form overemphasizes the contribution of proprioception in the RHI. He suggests that by adding sensory signals to the model, the influence of proprioception could be diluted and brought in line with the findings of Motyka and Litwin [39].

We argue for the inclusion of additional sensory signals, because this could increase the sensory evidence in favor of C_1 and therefore also "overwhelm" more strongly informed priors than those with $\sigma = 10^{35}$ mm|ms. However, any such expansion of the model must be carefully considered to avoid the peril of unnecessary model complexity and overfitting.

One possibility of an additional parameter suggested by the literature is the rotation of the hand. Rotation has been shown to have a strong influence on the RHI: Kalckert and Ehrsson [40] demonstrated that a rubber hand in an anatomically implausible position (facing towards the participant) does not induce ownership. The rotation could be represented in relation to the anatomically plausible position typically used in RHI experiments.

The rotation of the rubber hand would be inferred from visual input, while the rotation of the real hand would be inferred from proprioceptive input. Under C_1 , there would be only one rotation prior for both sensory modalities, which peaks at zero. However, while the prior for the real hand under C_2 would have the same peak, the prior for the rubber hand would be wider, because it could be facing in any direction. Since values near the peak of a wide distribution are less probable than values near the peak of a narrow distribution, the rotational degrees of freedom of the hands would be less likely under C_2 than under C_1 if they are congruent. This would increase the posterior probability of C_1 , at the expense of C_2 . To summarize, we expect the addition of the hand's rotation to the model to increase $p(C_1|D)$. If this effect were strong enough it could allow for a reduction of the sensory priors' widths and therefore increase their plausibility.

After settling on a model with plausible parameters, a possible next step would be to see whether it can predict interindividual differences in empirical data. For example, one prediction of the model is that participants with higher visual acuity should have a smaller propensity to experience the illusion. VR is the research paradigm of choice for such an experiment, because it allows for accurate assessment and manipulation of both the spatial and temporal information in the model through the recording of motion capture data and its (possibly manipulated) "playback" in VR. In the case of our example, participants with equivalent visual ability could inhabit a virtual avatar and be exposed to either unmodified playback of the motion capture data or playback in which the coordinates have been shifted, therefore reducing the accuracy of the visual input.

4.2. Applications

The BCIBO model is most applicable to those VR applications that represent the user as an avatar in the virtual environment and that let them control said avatar via motion capture. We use the term "motion capture" to refer to both motion capture via sensors on clothes (e.g., data gloves) and motion-tracking controllers (such as the controllers of the

HTC Vive). In most cases, applications employing motion capture try to make the user believe that the virtual avatar is their body. In some cases, only the task-relevant body parts (e.g., the hands) are rendered (e.g., Goh et al. [41]). For the purposes of this article, we consider these virtual body parts to be partial avatars and hold that a feeling of ownership over them is also key to most applications that make use of them.

Naturally, not all applications with avatars intend to make the user feel ownership over the avatar. For example, imagine an application that tries to increase awareness of depersonalization-derealization disorder by giving healthy people a VR enabled demonstration of what it might be like to have a dissociative experience of one's body. However, such cases are the exception and not the rule. Most VR applications with avatars try to immerse the user in the experience. If this is the goal, body ownership over the avatar is key.

That being said, we would like to point out that the term "body ownership" (and with it the BCIBO model) cannot easily be applied to VR applications in which the avatar is controlled with a gamepad (e.g., Bailenson et al. [42]) instead of motion capture. A gamepad is a controller which uses buttons and/or joysticks for game input. It is more accurate to speak of self-identification instead of embodiment of the avatar in these cases. The term "self-identification" is used here to indicate that the user most likely identifies with the virtual avatar, but they probably do not "inhabit" it as they would during a BOI. The use of a gamepad instead of motion capture creates a visuomotoric mismatch: The participant sees movements of the avatar that do not match their movements on the controller. For example, the press of a certain button might cause the avatar to jump. It has been shown that visuomotoric mismatches reduce body ownership [43]. In addition to this, gamepad-controlled applications often do not co-locate the user with their virtual avatar, further weakening body ownership [44].

An example for a field in which successful embodiment is often desirable is VR psychotherapy (for a review, see Matamala-Gomez et al. [45]). For example, Keizer et al. [46] let patients with anorexia inhabit a virtual avatar with healthy body proportions. Patients tended to overestimate their body proportions before the VR treatment, but they produced more realistic estimations afterwards. Hence, inhabiting another body seemed to have adjusted their internal body model.

Body ownership has also played an important role in rehabilitation interventions: Pichiorri et al. [47] used a virtual hand to provide stroke patients with feedback about a mental task they performed. The task was to imagine opening or closing one's hand. This practice, called motor imagery, is theorized to help patients with impaired motor functions in their recovery. The stroke patients wore an electroencephalography (EEG) cap. The EEG signals were used to calculate a score that approximated the success of the motor imagery task. If they performed the task successfully, patients saw a virtual hand in front of them perform the imagined movement (either opening or closing). This embodied feedback is likely more intuitive to the patient than more abstract forms (e.g., a smiley on a screen) [48] and carries the advantage of directly demonstrating the eventual aim of the intervention. Pichiorri et al. [47] found that post intervention the treatment group outperformed a control group, who underwent a motor imagery intervention without EEG and embodied feedback, in motor functionality.

VR has also been employed in education and training [49]. For example, Tang et al. [50] have used VR for the training of a blood sampling procedure. The scripted nature VR provides an ideal training ground for procedures that are highly standardized, as medical procedures often are. The use of VR for the training of these procedures could free up resources among human trainers to focus on less standardized procedures and soft skill acquisition.

Participants have indicated that the use of VR increased their motivation for the training [50]. We argue that ensuring embodiment of the avatar would further increase motivation by making the training more engaging. Of course, other factors such as sense of presence and immersion [51] also play an important role in this regard.

Interest in VR as a training tool has been especially high for surgical training [52] (however, see Müns et al. [53] for an article pointing out the limits of immersive VR in this context). Among the options for surgical training simulators is the commercial software PrecisionOS (www.precisionostech.com, accessed on 12 August 2021), which offers high-fidelity motion-capture-driven VR training for orthopedic surgery. For an exemplary training procedure using PrecisionOS see Goh et al. [41].

All the exemplary interventions mentioned above rely in part on body ownership for their success. Further development of the BCIBO model promises to deepen our understanding of body ownership and therefore enable the design of more effective therapeutic interventions that rely on it. Furthermore, the BCIBO model could be used as a component in a VR user model [15]. A user model (e.g., Horvitz et al. [54]), as the name implies, tries to model relevant states of the user. An accurate body ownership user model could detect when a user's body ownership over the avatar is slipping and enact countermeasures in the virtual environment. For example, to reinforce body ownership a stimulus that encourages hand-based interaction could be presented. This would nudge the user towards looking at their virtual body which in turn should strengthen their embodiment of the avatar.

A more direct, potential application of the model is in VR-related hardware design. Here, tolerable levels of accuracy both for gathering and displaying spatial and temporal information could be predicted from the model. For example, a producer of head-mounted displays (HMDs) might have to decide between several design options all with different levels of accuracy and production costs. HMDs receive a time series of motion capture data as input and display them as a virtual environment. A well-working version of the BCIBO model would be able to predict the average user's body ownership based on the discrepancy between the actual motion capture positions and time points and the virtual positions and time points. The BCIBO model is able to quantify the trade-off between the spatial and temporal inaccuracies of the system in terms of the probability of a BOI, therefore facilitating the goal of maximizing the user's sense of body ownership.

5. Conclusions

In conclusion, while we consider the BCIBO model to be a commendable step towards a computational explanation of body ownership we think it needs revision due to its unrealistically wide prior distributions. We showed that this cannot be remedied by our proposed quantitative changes to the model and hence conclude that a qualitative revision of the model is desirable. It is our belief that a good model of body ownership will improve both our understanding of this psychological construct and the design of VR applications that rely on an embodied user experience.

Supplementary Materials: The following are available online at <https://doi.org/10.17192/fdr/66.2> (tagged as version 3), accessed on 12 August 2021: data S1, **original.csv** unmodified model in accordance with Samad et al. [1]; data S2, **humanScale.csv** all the sensory priors' σ values are set to 10^4 mm|ms; data S3, **truncated.csv** the truncated version of the model as specified in Section 3.2; data S4, **99percent.csv** $p(C_1) = 0.7$; data S5, **closeTo100percent.csv** $p(C_1) = 1 - 10^{-16}$; data S6, **diff_original-99percent.csv** difference between data S1 and data S4.

Author Contributions: Conceptualization, M.S. and D.E.; methodology, M.S. and D.E.; software, M.S.; formal analysis, D.E.; investigation, M.S.; data curation, M.S.; writing—original draft preparation, M.S.; writing—review and editing, M.S. and D.E.; visualization, M.S.; supervision, D.E.; funding acquisition, D.E. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by Deutsche Forschungsgemeinschaft (German Research Foundation) project number 220482592 (IRTG 1901, The Brain in Action).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The data analyzed in this article were drawn from probability distributions. The code and the randomizer seed used to generate said data can be found here: <https://doi.org/10.17192/fdr/66.2> (tagged as version 3), accessed on 12 August 2021.

Acknowledgments: The authors would like to thank Andreas Kalckert, Peter Scarfe and Anantha Krishna Sivasubramaniam for a stimulating conversation that informed the contents of this article.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

Abbreviations

The following abbreviations are used in this manuscript:

BCIBO	Bayesian Causal Inference of Body Ownership
BOI	body ownership illusion
EEG	electroencephalography
GUI	graphical user interface
HMD	head-mounted display
MDPI	Multidisciplinary Digital Publishing Institute
RHI	rubber hand illusion
SEM	standard error of the mean
VR	virtual reality
C_1	common cause model
C_2	separate causes model
X	inferred position of the rubber/real hand
T	inferred time point of the brush stroke
χ_v	spatial visual input
χ_p	(spatial) proprioceptive input
τ_v	temporal visual input
τ_t	(temporal) tactile input

Appendix A. Model Specifications

Appendix A.1. Original Model

$$C \sim \text{Bernoulli}(0.5)$$

$$\sigma_i \in [10^0, 10^5, 10^{10}, \dots, 10^{35}]$$

$$d_j \in [160, 180, 200, \dots, 360]$$

$$\text{if } C = 0 \left\{ \begin{array}{l} X_{\text{rub}} \sim \text{Normal}(320 - d_j, \sigma_i) \\ X_{\text{real}} \sim \text{Normal}(320, \sigma_i) \\ T_{\text{rub}}, T_{\text{real}} \sim \text{Normal}(0, \sigma_i) \\ \chi_v \sim \text{Normal}(X_{\text{rub}}, 1) \\ \chi_p \sim \text{Normal}(X_{\text{real}}, 15) \\ \tau_v \sim \text{Normal}(T_{\text{rub}}, 20) \\ \tau_t \sim \text{Normal}(T_{\text{real}}, 20) \end{array} \right. \quad (\text{A1})$$

$$\text{if } C = 1 \left\{ \begin{array}{l} X_{\text{hand}} \sim \text{Normal}(320 - d_j, \sigma_i) \\ T_{\text{hand}} \sim \text{Normal}(0, \sigma_i) \\ \chi_v \sim \text{Normal}(X_{\text{hand}}, 1) \\ \chi_p \sim \text{Normal}(X_{\text{hand}}, 15) \\ \tau_v \sim \text{Normal}(T_{\text{hand}}, 20) \\ \tau_t \sim \text{Normal}(T_{\text{hand}}, 20) \end{array} \right.$$

Appendix A.2. Truncated Model

The truncated model follows the original model (see Appendix A.1) with the following changes:

$$\begin{aligned}
 & d_j \in [20, 40, 60, \dots, 220] \\
 & \text{if } C = 0 \begin{cases} X_{\text{rub}} & \sim \text{TruncatedNormal}(320 - d_j, \sigma_i, [-2000, 2000]) \\ X_{\text{real}} & \sim \text{TruncatedNormal}(320, \sigma_i, [-850, 850]) \\ T_{\text{rub}}, T_{\text{real}} & \sim \text{TruncatedNormal}(0, \sigma_i, [0, 3.6 \times 10^6]) \end{cases} \quad (\text{A2}) \\
 & \text{if } C = 1 \begin{cases} X_{\text{hand}} & \sim \text{TruncatedNormal}(320 - d_j, \sigma_i, [-850, 850]) \\ T_{\text{hand}} & \sim \text{TruncatedNormal}(0, \sigma_i, [0, 3.6 \times 10^6]) \end{cases}
 \end{aligned}$$

Appendix A.3. Changes in the Model Prior

Simulation runs that changed the model prior (see Section 3.3) follow the original model (see Appendix A.1) with the following changes:

$$\begin{aligned}
 & C \sim \text{Bernoulli}(\theta) \\
 & \theta \in [0.99, 1 - 10^{-16}] \quad (\text{A3})
 \end{aligned}$$

References

- Samad, M.; Chung, A.J.; Shams, L. Perception of Body Ownership Is Driven by Bayesian Sensory Inference. *PLoS ONE* **2015**, *10*, e0117178. [CrossRef] [PubMed]
- Schubert, M.; Endres, D. The Bayesian Causal Inference of Body Ownership Model: Use in VR and Plausible Parameter Choices. In Proceedings of the 2021 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW), Lisbon, Portugal, 27 March–1 April 2021; pp. 67–70. [CrossRef]
- Botvinick, M.; Cohen, J. Rubber Hands ‘Feel’ Touch That Eyes See. *Nature* **1998**, *391*, 756. [CrossRef] [PubMed]
- Neustadter, E.S.; Fineberg, S.K.; Leavitt, J.; Carr, M.M.; Corlett, P.R. Induced Illusory Body Ownership in Borderline Personality Disorder. *Neurosci. Conscious.* **2019**, *5*, niz017. [CrossRef] [PubMed]
- Tsakiris, M. The Sense of Body Ownership. In *The Oxford Handbook of the Self*; Oxford University Press: Oxford, UK, 2011; pp. 180–203.
- Lewis, E.; Lloyd, D.M. Embodied Experience: A First-Person Investigation of the Rubber Hand Illusion. *Phenomenol. Cogn. Sci.* **2010**, *9*, 317–339. [CrossRef]
- Körding, K.P.; Beierholm, U.; Ma, W.J.; Quartz, S.; Tenenbaum, J.B.; Shams, L. Causal Inference in Multisensory Perception. *PLoS ONE* **2007**, *2*, e943. [CrossRef]
- Hohwy, J. *The Predictive Mind*, 1st ed.; Oxford University Press: Oxford, UK, 2013.
- Magnotti, J.F.; Ma, W.J.; Beauchamp, M.S. Causal Inference of Asynchronous Audiovisual Speech. *Front. Psychol.* **2013**, *4*, 798. [CrossRef]
- Geisler, W.S.; Kersten, D. Illusions, Perception and Bayes. *Nat. Neurosci.* **2002**, *5*, 508–510. [CrossRef]
- Riemer, M.; Bublatzky, F.; Trojan, J.; Alpers, G.W. Defensive Activation during the Rubber Hand Illusion: Ownership versus Proprioceptive Drift. *Biol. Psychol.* **2015**, *109*, 86–92. [CrossRef]
- Makin, T.R.; Holmes, N.P.; Ehrsson, H.H. On the Other Hand: Dummy Hands and Peripersonal Space. *Behav. Brain Res.* **2008**, *191*, 1–10. [CrossRef]
- Armel, K.C.; Ramachandran, V.S. Projecting Sensations to External Objects: Evidence from Skin Conductance Response. *Proc. R. Soc. Biol. Sci.* **2003**, *270*, 1499–1506. [CrossRef]
- Reader, A.T.; Trifonova, V.S.; Ehrsson, H.H. The Rubber Hand Illusion Does Not Influence Basic Movement. 2021. Available online: <https://doi.org/10.31219/osf.io/6dyzq> (accessed on 12 August 2021). [CrossRef]
- Schürmann, T.; Vogt, J.; Christ, O.; Beckerle, P. The Bayesian Causal Inference Model Benefits from an Informed Prior to Predict Proprioceptive Drift in the Rubber Foot Illusion. *Cogn. Process.* **2019**, *20*, 447–457. [CrossRef]
- Chancel, M.; Ehrsson, H.H.; Ma, W.J. Uncertainty-Based Inference of a Common Cause for Body Ownership. 2021. Available online: <https://doi.org/10.31219/osf.io/yh2z7> (accessed on 12 August 2021). [CrossRef]
- Crea, S.; D’Alonzo, M.; Vitiello, N.; Cipriani, C. The Rubber Foot Illusion. *J. Neuroeng. Rehabil.* **2015**, *12*, 77. [CrossRef] [PubMed]
- Flögel, M.; Kalveram, K.T.; Christ, O.; Vogt, J. Application of the Rubber Hand Illusion Paradigm: Comparison between Upper and Lower Limbs. *Psychol. Res.* **2015**, *80*, 298–306. [CrossRef] [PubMed]
- Christ, O.; Elger, A.; Schneider, K.; Rapp, A.; Beckerle, P. Identification of Haptic Paths with Different Resolution and Their Effect on Body Scheme Illusion in Lower Limbs. *Tech. Assist. Rehabil.* **2013**, 1–4. Available online: https://www.ige.tu-berlin.de/fileadmin/fg176/IGE_Printreihe/TAR_2013/paper/Session-10-Event-1-Christ.pdf (accessed on 12 August 2021).

20. Jones, S.A.H.; Cressman, E.K.; Henriques, D.Y.P. Proprioceptive Localization of the Left and Right Hands. *Exp. Brain Res.* **2010**, *204*, 373–383. [[CrossRef](#)] [[PubMed](#)]
21. van Beers, R.J.; Sittig, A.C.; Denier van der Gon, J.J. The Precision of Proprioceptive Position Sense. *Exp. Brain Res.* **1998**, *122*, 367–377. [[CrossRef](#)]
22. Hirsh, I.J.; Sherrick, C.E., Jr. Perceived Order in Different Sense Modalities. *J. Exp. Psychol.* **1961**, *62*, 423–432. [[CrossRef](#)] [[PubMed](#)]
23. Riemer, M.; Trojan, J.; Beauchamp, M.; Fuchs, X. The Rubber Hand Universe: On the Impact of Methodological Differences in the Rubber Hand Illusion. *Neurosci. Biobehav. Rev.* **2019**, *104*, 268–280. [[CrossRef](#)]
24. Bars, I.; Terning, J.; Nekoogar, F. *Extra Dimensions in Space and Time; Multiversal Journeys*; Springer: New York, NY, USA, 2010.
25. Jones, M.; Love, B.C. Bayesian Fundamentalism or Enlightenment? On the Explanatory Status and Theoretical Contributions of Bayesian Models of Cognition. *Behav. Brain Sci.* **2011**, *34*, 169. [[CrossRef](#)] [[PubMed](#)]
26. Bowers, J.S.; Davis, C.J. Bayesian Just-so Stories in Psychology and Neuroscience. *Psychol. Bull.* **2012**, *138*, 389–414. [[CrossRef](#)] [[PubMed](#)]
27. Hunter, J.D. Matplotlib: A 2D Graphics Environment. *Comput. Sci. Eng.* **2007**, *9*, 90–95. [[CrossRef](#)]
28. Harris, C.R.; Millman, K.J.; van der Walt, S.J.; Gommers, R.; Virtanen, P.; Cournapeau, D.; Wieser, E.; Taylor, J.; Berg, S.; Smith, N.J.; et al. Array Programming with NumPy. *Nature* **2020**, *585*, 357–362. [[CrossRef](#)]
29. Virtanen, P.; Gommers, R.; Oliphant, T.E.; Haberland, M.; Reddy, T.; Cournapeau, D.; Burovski, E.; Peterson, P.; Weckesser, W.; Bright, J. SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nat. Methods* **2020**, *17*, 261–272. [[CrossRef](#)]
30. 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.* **2007**, *64*, 104–109. [[CrossRef](#)]
31. Kilteni, K.; Normand, J.M.; Sanchez-Vives, M.V.; Slater, M. Extending Body Space in Immersive Virtual Reality: A Very Long Arm Illusion. *PLoS ONE* **2012**, *7*, e40867. [[CrossRef](#)] [[PubMed](#)]
32. Miller, M.R.; Crapo, R.; Hankinson, J.; Brusasco, V.; Burgos, F.; Casaburi, R.; Coates, A.; Enright, P.; van der Grinten, C.P.M.; Gustafsson, P.; et al. General Considerations for Lung Function Testing. *Eur. Respir. J.* **2005**, *26*, 153–161. [[CrossRef](#)] [[PubMed](#)]
33. Reeves, S.L.; Varakamin, C.; Henry, C.J. The Relationship between Arm-Span Measurement and Height with Special Reference to Gender and Ethnicity. *Eur. J. Clin. Nutr.* **1996**, *50*, 398–400. [[PubMed](#)]
34. Federal Statistical Office of Germany. Körpermaße nach Altersgruppen und Geschlecht-Statistisches Bundesamt. Available online: <https://web.archive.org/web/20210514201250/https://www.destatis.de/DE/Themen/Gesellschaft-Umwelt/Gesundheit/Gesundheitszustand-Relevantes-Verhalten/Tabellen/liste-koerpermasse.html> (accessed on 12 August 2021).
35. Kammers, M.P.M.; de Vignemont, F.; Verhagen, L.; Dijkerman, H.C. The Rubber Hand Illusion in Action. *Neuropsychologia* **2009**, *47*, 204–211. [[CrossRef](#)]
36. Durgin, F.H.; Evans, L.; Dunphy, N.; Klostermann, S.; Simmons, K. Rubber Hands Feel the Touch of Light. *Psychol. Sci.* **2007**, *18*, 152–157. [[CrossRef](#)]
37. Abdulkarim, Z.; Hayatou, Z.; Ehrsson, H.H. Sustained Rubber Hand Illusion after the End of Visuotactile Stimulation with a Similar Time Course for the Reduction of Subjective Ownership and Proprioceptive Drift. 2021. Available online: <https://doi.org/10.31234/osf.io/wt82m> (accessed on 12 August 2021). [[CrossRef](#)]
38. Litwin, P. Extending Bayesian Models of the Rubber Hand Illusion. *Multisens. Res.* **2020**, *33*, 127–160. [[CrossRef](#)] [[PubMed](#)]
39. Motyka, P.; Litwin, P. Proprioceptive Precision and Degree of Visuo-Proprioceptive Discrepancy Do Not Influence the Strength of the Rubber Hand Illusion. *Perception* **2019**, *48*, 882–891. [[CrossRef](#)] [[PubMed](#)]
40. Kalckert, A.; Ehrsson, H.H. Moving a Rubber Hand That Feels Like Your Own: A Dissociation of Ownership and Agency. *Front. Hum. Neurosci.* **2012**, *6*, 40. [[CrossRef](#)] [[PubMed](#)]
41. Goh, G.S.; Lohre, R.; Parvizi, J.; Goel, D.P. Virtual and Augmented Reality for Surgical Training and Simulation in Knee Arthroplasty. *Arch. Orthop. Trauma Surg.* **2021**. [[CrossRef](#)]
42. Bailenson, J.N.; Yee, N.; Blascovich, J.; Beall, A.C.; Lundblad, N.; Jin, M. The Use of Immersive Virtual Reality in the Learning Sciences: Digital Transformations of Teachers, Students, and Social Context. *J. Learn. Sci.* **2008**, *17*, 102–141. [[CrossRef](#)]
43. Kokkinara, E.; Slater, M. Measuring the Effects through Time of the Influence of Visuomotor and Visuotactile Synchronous Stimulation on a Virtual Body Ownership Illusion. *Perception* **2014**, *43*, 43–58. [[CrossRef](#)]
44. Kilteni, K.; Groten, R.; Slater, M. The Sense of Embodiment in Virtual Reality. *Presence Teleoper. Virtual Environ.* **2012**, *21*, 373–387. [[CrossRef](#)]
45. Matamala-Gomez, M.; Maselli, A.; Malighetti, C.; Realdon, O.; Mantovani, F.; Riva, G. Virtual Body Ownership Illusions for Mental Health: A Narrative Review. *J. Clin. Med.* **2021**, *10*, 139. [[CrossRef](#)]
46. Keizer, A.; van Elburg, A.; Helms, R.; Dijkerman, H.C. A Virtual Reality Full Body Illusion Improves Body Image Disturbance in Anorexia Nervosa. *PLoS ONE* **2016**, *11*, e0163921. [[CrossRef](#)] [[PubMed](#)]
47. Pichiorri, F.; Morone, G.; Petti, M.; Toppi, J.; Pisotta, I.; Molinari, M.; Paolucci, S.; Inghilleri, M.; Astolfi, L.; Cincotti, F.; et al. Brain-Computer Interface Boosts Motor Imagery Practice during Stroke Recovery. *Ann. Neurol.* **2015**, *77*, 851–865. [[CrossRef](#)]
48. Braun, N.; Debener, S.; Spychala, N.; Bongartz, E.; Sörös, P.; Müller, H.H.O.; Philippsen, A. The Senses of Agency and Ownership: A Review. *Front. Psychol.* **2018**, *9*, 535. [[CrossRef](#)] [[PubMed](#)]

49. Freina, L.; Ott, M. A Literature Review on Immersive Virtual Reality in Education: State of the Art and Perspectives. In *The International Scientific Conference eLearning and Software for Education*; Carol I National Defence University Publishing House: Bucharest, Romania, 2015; Volume 11, pp. 133–141.
50. Tang, Y.M.; Ng, G.W.Y.; Chia, N.H.; So, E.H.K.; Wu, C.H.; Ip, W.H. Application of Virtual Reality (VR) Technology for Medical Practitioners in Type and Screen (T&S) Training. *J. Comput. Assist. Learn.* **2021**, *37*, 359–369. [[CrossRef](#)]
51. Skarbez, R.; Brooks, F.P., Jr.; Whitton, M.C. A Survey of Presence and Related Concepts. *ACM Comput. Surv.* **2017**, *50*, 96:1–96:39. [[CrossRef](#)]
52. Slater, M.; Sanchez-Vives, M.V. Enhancing Our Lives with Immersive Virtual Reality. *Front. Robot. AI* **2016**, *3*, 74. [[CrossRef](#)]
53. Müns, A.; Meixensberger, J.; Lindner, D. Evaluation of a Novel Phantom-Based Neurosurgical Training System. *Surg. Neurol. Int.* **2014**, *5*, 173. [[CrossRef](#)] [[PubMed](#)]
54. Horvitz, E.; Breese, J.; Heckerman, D.; Hovel, D.; Rommelse, K. The Lumière Project: Bayesian User Modeling for Inferring the Goals and Needs of Software Users. In *Proceedings of the Fourteenth Conference on Uncertainty in Artificial Intelligence, UAI'98*, Madison, Wisconsin, 24–26 July 1998; Morgan Kaufmann Publishers Inc.: San Francisco, CA, USA, 1998; pp. 256–265.