



# Multisensory Integration in Body Representation

# 5

Wen Fang, Yuqi Liu, and Liping Wang

## Abstract

To be aware of and to move one's body, the brain must maintain a coherent representation of the body. While the body and the brain are connected by dense ascending and descending sensory and motor pathways, representation of the body is not hardwired. This is demonstrated by the well-known rubber hand illusion in which a visible fake hand is erroneously felt as one's own hand when it is stroked in synchrony with the viewer's unseen actual hand. Thus, body representation in the brain is not mere maps of tactile and proprioceptive inputs, but a construct resulting from the interpretation and integration of inputs across sensory modalities.

## Keywords

Causal inference · Macaques · Body representation · Self-consciousness

Accurate integration of multisensory information serves essential purposes in daily life. For example, we recognize ourselves in the mirror and distinguish our own shadow from others by matching movements we intend to generate and movements seen on the visual image. Successful motor control, such as grabbing a coffee mug, critically relies on the integration of visual and proprioceptive information on one's arm and hand position, and of visual and tactile information on where the fingers are on the handle. Therefore, accurate integration of multisensory information is crucial for distinguishing between oneself and external world as well as for interacting with the environment. As the brain is constantly flooded with sensory information both from one's own body and the environment, how it properly integrates and segregates sensory information becomes an important question for understanding the underlying mechanism of body representation. In this chapter, we first discuss behavioral work that investigates the constraints and principles underlying multisensory integration regarding the body. We then introduce a Bayesian framework that theorizes multisensory integration as inferring the source of the sensory inputs by an optimal observer. Finally, we review evidence from neuroimaging and neurophysiological studies on the neural correlates and computational principles of body representation.

---

W. Fang (✉) · Y. Liu · L. Wang  
Institute of Neuroscience, Key Laboratory of Primate Neurobiology, CAS Center for Excellence in Brain Science and Intelligence Technology, Chinese Academy of Sciences, Shanghai, China  
e-mail: [wenfang@ion.ac.cn](mailto:wenfang@ion.ac.cn); [liuyq@ion.ac.cn](mailto:liuyq@ion.ac.cn); [liping.wang@ion.ac.cn](mailto:liping.wang@ion.ac.cn)

## 5.1 Temporal and Spatial Constraints on Multisensory Integration

Multisensory integration is typically studied by creating a mismatch between inputs from different modalities. Early studies on the integration of visual and auditory information have identified spatial and temporal rules of multisensory integration, such that inputs are more likely to be combined if spatially and temporally closer (Meredith 1987; Meredith and Stein 1986). The same rules also apply to multisensory integration regarding the body. Studies have shown that the processing of bodily signals is influenced by external stimuli occurring within a limited space around the body part, also known as the peripersonal space (PPS) (Làdavvas et al. 1998; Spence et al. 2004). For example, the perceived location of tactile stimuli is more strongly biased toward a concurrent visual stimulus when the visual stimulus occurs close to the body versus far (Spence et al. 2004). These findings demonstrate the multisensory nature of body representation such that tactile perception is automatically biased by task-irrelevant visual information. Moreover, there is a spatial limit within which sensory inputs are considered bodily-related and integrated together.

Multisensory integration not only underlies the perception of bodily sensory inputs, but also how the brain represents the body itself. Because it is not always possible to separate information from one's own body, bodily illusions are often used. In these paradigms, participants' actual hand is hidden from view while seeing a fake hand. In this way, proprioceptive information and visual information are dissociated. Integration between visual and proprioceptive information is indexed by the illusory embodiment of the viewed hand, i.e., the viewed hand feels like one's own body—referred to as body ownership, and the hand is perceived closer to where the viewer sees it (Botvinick and Cohen 1998; Holmes et al. 2004). By examining the effect of spatial and temporal congruence between the viewed hand and unseen actual hand on the strength of illusion, researchers can identify factors influencing multisensory integration.

One important factor influencing multisensory integration is the spatial location of the viewed fake hand. As the distance between the viewed hand and unseen actual hand increases, hence more difficult to reconcile the discrepancy between visual and proprioceptive information, participants are less likely to feel the viewed hand as their own hand (Medina et al. 2015) or is located where the viewed hand is (Holmes et al. 2004, 2006; Holmes and Spence 2005). By gradually displacing the rubber hand further from participants' unseen actual hand, Lloyd (2007) found decreasing illusion with increasing distance. Importantly, there was an exponential non-linear decrease as the rubber hand was positioned outside of the participants' reachable space, marking the boundary of the spatial range within which an object could be embodied.

In addition to spatial location, a congruent posture of the viewed and unseen actual hand is critical. For example, with a fixed distance between the rubber hand and the unseen actual hand, subjective ownership of the rubber hand decreased with mismatch in the hand angle between the two hands (e.g., fingers pointing forward vs. pointing 30° leftward) despite synchronous visuotactile stimulation (Costantini and Haggard 2007; Ide 2013).

It has been well-established that the temporal synchrony between the viewed hand and the unseen actual hand plays an important role. In the rubber hand illusion, synchronous strokes on the unseen hidden hand and the rubber hand, such that strokes seen on the viewed hand match tactile sensation on the actual hand, elicit strong illusion. Asynchronous strokes, however, abolish the illusion (Botvinick and Cohen 1998; Tsakiris and Haggard 2005). Using a different and more powerful paradigm, researchers can manipulate whether the unseen hand and the viewed hand are performing congruent movements (mirror box illusion; Ramachandran and Rogers-Ramachandran 1996; Medina et al. 2015). Whereas participants experience strong ownership of the viewed hand when the seen movements on the viewed hand is congruent with movements performed by the unseen actual hand, the illusion is much weaker when the movements

are out of phase (Holmes and Spence 2005; Liu and Medina 2018; Medina et al. 2015). These findings provide evidence for the importance of temporal synchrony in multisensory integration.

The evidence discussed above suggests a key role of cross-modal congruence in multisensory integration and body representation. However, matching between bottom-up sensory information is not sufficient for the brain to embody an object. It makes intuitive sense that one would not perceive a dog's paw as their own hand regardless of sensory information, implying additional constraints from prior knowledge of our body independent of incoming sensory inputs. The following section discusses the influence of prior information on multisensory integration.

---

## 5.2 Prior Knowledge of the Body Influences Multisensory Integration

At the level of visual features, the human body has specific anatomical and structural properties that differ from other objects (e.g., “a hand has five fingers sticking out”). At the semantic level, we use labels and descriptive language to distinguish objects of different categories (e.g., a human body versus a tree). These forms of prior knowledge are learned in life experience and exist independent of online sensory information. Using the rubber hand illusion paradigm, studies found weaker illusory embodiment when individuals view an object (e.g., a wood stick) versus a realistic rubber hand (Holmes and Spence 2005; Tsakiris et al. 2008; Tsakiris and Haggard 2005). The illusion was also weaker for neutral objects whose shape deviates from a standard hand (Haans et al. 2008; Tsakiris et al. 2010). These findings indicate that individuals have prior expectations of what features constitute a hand based on stored visual body representation (Tsakiris 2010). Interestingly, visual features that are more specific to one's own body, such as size and color, do not have a dramatic effect on the embodiment of the viewed hand or body (Austen et al. 2004; Farmer et al. 2012, 2014; Peck et al.

2013; but see Pavani and Zampini 2007). Based on these findings, it was proposed that the stored visual body representation encodes general shape information of body parts instead of self-specific features (Kilteni et al. 2015; Tsakiris 2010).

Another source of prior knowledge comes from the body schema that represents online body position in space as the body moves (Head and Holmes 1911; Schwoebel and Coslett 2005). The movement of the body is limited by biomechanical constraints—resistance caused by joint and muscle structures that determines the difficulty in and the possible range of body movement (Parsons 1987, 1994). As such, biomechanical constraints are also encoded in the body schema. Importantly, biomechanical constraints not only affect physical movements but manifests in mental body representations. For example, when individuals are asked to judge the chirality of a hand image in a selected orientation, the reaction time is longer if the rotation from the individuals' own hand posture to the hand image is more biomechanically constrained, even if the rotation angle is the same (Cooper and Shepard 1975; Parsons 1987, 1994; Zapparoli et al. 2014). These findings provide evidence that the participants performed the task by mentally simulating body movements, in which biomechanical constraints are encoded.

There is evidence that biomechanical constraints between the unseen actual hand and viewed hand influences multisensory integration. With the angular difference between the viewed hand and unseen actual hand fixed, biomechanical constraints can be manipulated such that the rotation from the actual to the viewed hand is less biomechanically constrained in one condition but more constrained in another (Ide 2013; Liu and Medina 2017). Participants reported weaker illusions in the more-biomechanically-constrained condition despite matched angular differences, indicating that the amount of biomechanical constraint is also computed into the overall discrepancy between visual and proprioceptive information. These findings prove that multisensory information is influenced not only by bottom-up sensory information but also prior information stored in the body schema.

Another type of constraint regards anatomical plausibility of body configuration, i.e., whether a body posture is possible to occur. It was found that the rubber hand illusion was abolished when the viewed hand is in anatomically implausible postures (e.g., fingers pointing straight toward one's body along the sagittal axis) despite synchronous visuotactile stimulation (Ehrsson et al. 2004; Ide 2013; Tsakiris and Haggard 2005). These findings can be accounted for by the encoding of biomechanical constraints discussed above, such that the viewed hand is in an infinitely biomechanically constrained position that proprioceptive information cannot be biased toward. Alternatively, the brain may refer to a body structural description, the stored knowledge of the relative position of body parts on the body surface, for example, the arm is attached to the trunk (Buxbaum and Coslett 2010). Anatomically implausible hand postures often imply breaking of joints, hence violating the body structural description, leading to lower degrees of embodiment (Kilteni et al. 2015).

In summary, the brain uses both incoming sensory information and prior knowledge of the body in multisensory information. The interplay of both bottom-up sensory inputs and top-down knowledge is summarized in a neurocognitive model of body ownership (Tsakiris 2010) that conceptualizes body representation as a set of information comparators. What factors determine the relative importance of each source of information, and how do these factors lead to the final percept that a hand belongs to one's own? In the next section, we discuss a Bayesian framework that addresses the computational principles of multisensory integration.

### 5.3 A Bayesian Framework of Multisensory Integration

Forming a coherent body representation is a process of inferring the state of the body based on incoming sensory information. The Bayesian framework posits that upon receiving sensory inputs from multiple modalities, the brain generates hypotheses on the cause of these sensory inputs. For example, inferring the hand position gives rise to the perceived visual and propriocep-

tive information. Given the noise in both external inputs and internal sensory systems, each hypothesis is only correct at a certain probability. These probabilities are called *posterior probability* as they are conditional on a particular set of sensory inputs, written as  $P(S|x_v, x_p)$  (the probability of state  $S$ , for example, hand position, given visual ( $x_v$ ), and proprioceptive ( $x_p$ ) information). The goal of an optimal observer is to find the state  $S$  with the maximum posterior probability. Following Bayes' rule, posterior probability depends on the product of two components: the likelihood of obtaining a particular set of sensory inputs given state  $S$  ( $P(x_v, x_p|S)$ ), and the prior probability of this hypothesis based on prior knowledge ( $P(S)$ ) (see Eq. 5.1).

$$P(S|x_v, x_p) = \frac{P(x_v, x_p|S)P(S)}{P(x_v, x_p)} \quad (5.1)$$

As the denominator  $P(x_v, x_p)$  is independent of the state  $S$ , Eq. (5.1) is simplified as:

$$P(S|x_v, x_p) \propto P(x_v, x_p|S)P(S) \quad (5.2)$$

Assuming a uniform distribution  $P(S)$ , denoting equal prior probability that the hand appears anywhere in space, maximizing the posterior probability equals to maximizing the likelihood. Further assume that each sensory modality is corrupted by independent Gaussian noise, Eq. (5.2) is simplified as:

$$P(S|x_v, x_p) \propto P(x_v|S)P(x_p|S) \quad (5.3)$$

Maximum-likelihood estimation of this equation then becomes a weighted sum of visual and proprioceptive input based on their relative reliability, with reliability the inverse of the standard deviation of each input:

$$\hat{S} = \frac{\frac{1}{\sigma_v}}{\frac{1}{\sigma_v} + \frac{1}{\sigma_p}} x_v + \frac{\frac{1}{\sigma_p}}{\frac{1}{\sigma_v} + \frac{1}{\sigma_p}} x_p \quad (5.4)$$

As a result, the final percept is biased toward the more reliable unimodal estimate, and the reliability of the final percept is maximized (Ernst and Bühlhoff 2004). This model therefore

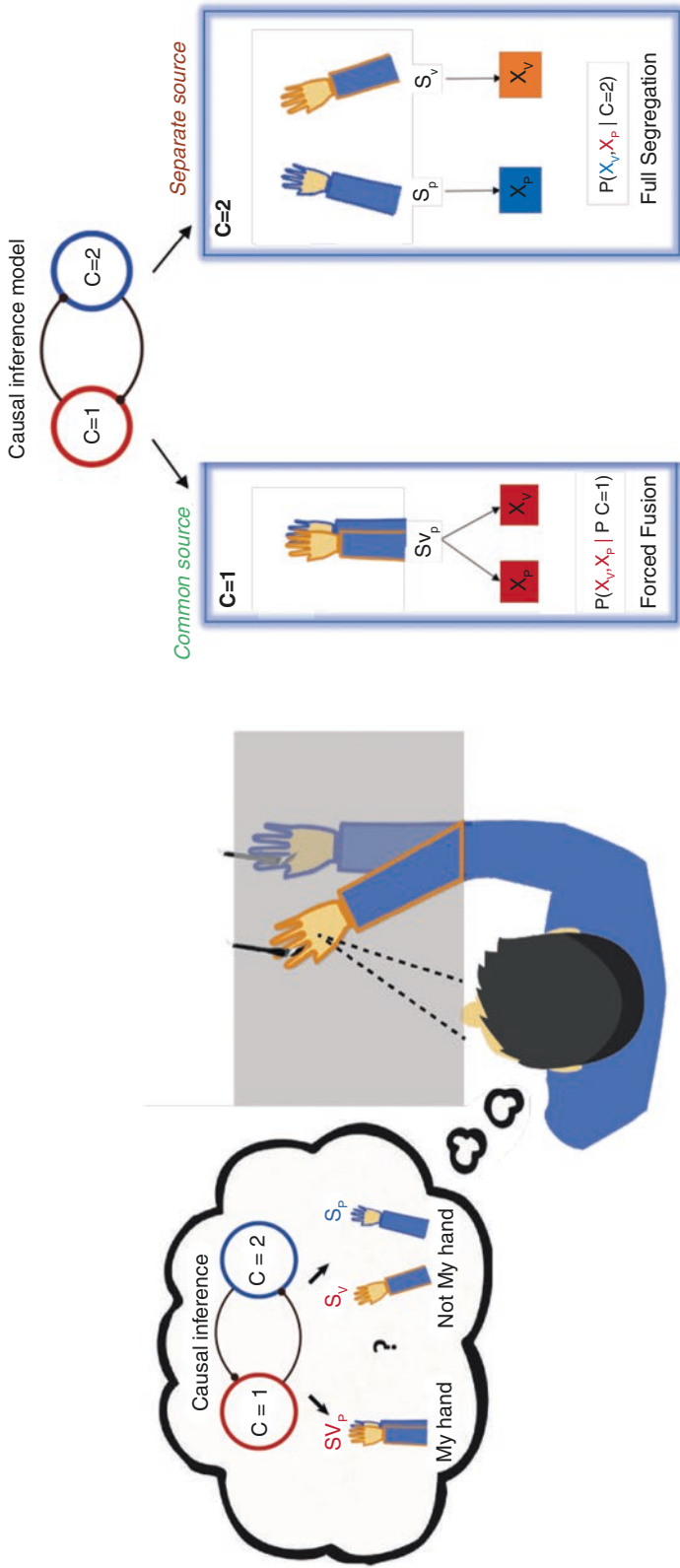
accounts for the findings that participants perceive their hand toward where the visual hand is (i.e., visual capture), presumably because visual information is typically less noisy than proprioceptive information. As an optimal weighting principle, maximum-likelihood estimation has been supported by multiple studies on various sensory modalities (Alais and Burr 2004; Ernst and Banks 2002; Van Beers et al. 1999; Witten and Knudsen 2005). For example, in one such study, participants were asked to estimate the height of a block based on haptic and visual information (Ernst and Banks 2002). Small discrepancies between visual and haptic inputs were introduced by manipulating the visual image viewed from a pair of goggles. Consistent with the model, participants' estimate was biased toward visual information, with the weight of the visual information decreasing with visual noise.

An important consequence of this model is that unimodal estimates are always integrated to a single final percept, hence the model is referred to as the forced-fusion model (Körding et al. 2007; Shams and Beierholm 2010). Although it well explains human behavior when the discrepancy between unimodal inputs is small and easy to resolve, it does not apply to all real-world situations. Organized representations require not only accurate integration of inputs belonging to the same object, but segregation of inputs that come from different sources. For example, when playing a duet, it is equally important to integrate information on one's own hand and to segregate information from the partner's hand. In experiments on bodily illusions, the viewed hand is not embodied if it is placed too far from the participant's actual hand (Lloyd 2007; Medina et al. 2015), indicating segregation of visual and proprioceptive information. The underlying problem is to infer whether inputs are emitted by the same object, or in other words, have a common cause, a process referred to as "causal inference."

Researchers have proposed a Bayesian causal inference model to account for how the brain infers the causal structure of sensory inputs (Fig. 5.1; Körding et al. 2007; Shams and

Beierholm 2010). In this model, the underlying causal structure of whether multisensory inputs come from a common cause is denoted by the posterior probability ( $P(C|x_v, x_p)$ ). Two causal hypotheses are tested: that the inputs are caused by a common source ( $C = 1$ ), which leads to complete integration of inputs from both modalities, or that the inputs are caused by independent sources ( $C = 2$ ), which leads to complete segregation of information from different modalities. In this way, the Bayesian causal inference model has a hierarchical structure in which the brain first infers the causal structure of sensory inputs, and then makes an estimate of the object state (e.g., location) under the causal structure. Following Bayes' rule, the posterior probability of each causal hypothesis depends on the likelihood of receiving the current sensory inputs given this causal hypothesis ( $P(x_v, x_a|C)$ , and the prior probability of the causal hypothesis ( $P(C)$ ). An important cue that informs the brain about the causal structure of inputs from multiple modalities is the "similarity" between the inputs: inputs that are closer in time, space, or other dimensions tend to have a higher likelihood under the common cause hypothesis, making them more likely to be integrated.

How is the estimate of object location contingent on the inferred causal structure? One strategy is to follow the most likely causal hypothesis (model selection, Wozny et al. 2008). If the probability of the common cause hypothesis is higher than the hypothesis of the independent cause, inputs are fully integrated by maximum-likelihood estimation. Otherwise, the brain estimates each modality independently without combining them. Another strategy is to weight the estimate under each causal hypothesis in proportion to each hypothesis's posterior probability (model averaging, see Eqs. (5.5a) and (5.5b), Wozny et al. 2010; Körding et al. 2007). By considering the causal structure of inputs from multiple modalities, the Bayesian causal inference model can account for the full range of multisensory integration from complete integration to complete segregation.



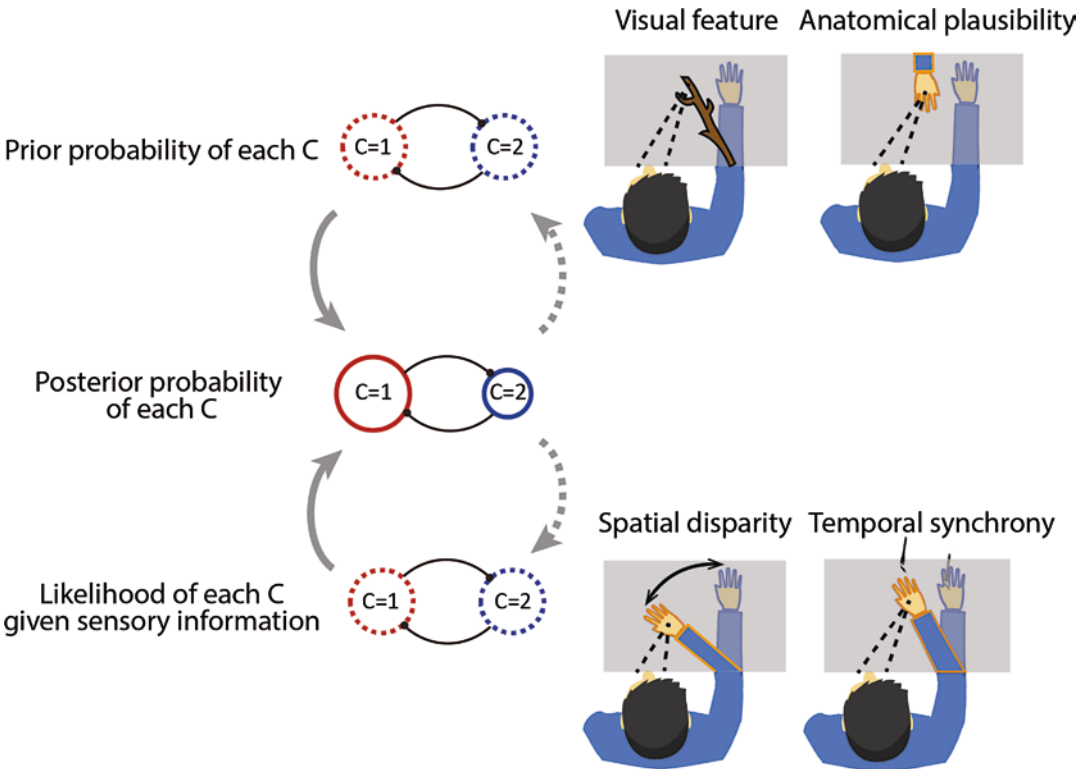
**Fig. 5.1** Causal inference problem in body representation. Upon receiving visual and proprioceptive information about the hand, the brain infers the probability that the two sources of information come from a common cause

$$\hat{S}_p = P(C = 1|x_p, x_v)_{vp,c=1} + \hat{S}P(C = 2|x_p, x_v)\hat{S}_{p,c=2} \tag{5.5a}$$

$$\hat{S}_v = P(C = 1|x_p, x_v)\hat{S}_{vp,c=1} + P(C = 2|x_p, x_v)\hat{S}_{v,c=2} \tag{5.5b}$$

The Bayesian causal inference model can account for the effect of various factors on multisensory integration regarding the body (Fig. 5.2; Fang et al. 2019; Kilteni et al. 2015; Samad et al. 2015). Provided with multisensory information about the hand including vision, proprioceptive, touch, etc., the brain is faced with a causal inference problem of whether all sources of information come from the same hand, i.e., “my hand.” The posterior probability of a common hand depends on the likelihood of receiving the current sensory information if they belong to the same hand and the prior probability of there

being one hand. The closer the visual and proprioceptive hand positions are, the more likely they define the same hand. Similarly, other spatial and temporal congruence factors discussed above contribute to the likelihood of a common cause. On the other hand, prior body knowledge constrains how strongly a viewed object can be embodied. For example, the model-fitted prior probability of a common cause was lower when participants viewed a woodblock versus a verisimilar hand (Fang et al. 2019). The higher the posterior probability of a common cause, the more strongly the viewed hand is perceived as



**Fig. 5.2** Bayesian causal inference in body representation. The posterior probability of a common cause depends on the prior probability and likelihood. Prior information refers to existent knowledge including the visual feature

and anatomical plausibility of the body. Likelihood depends on the congruence between information across modalities, such as spatial disparity and visual-tactile synchrony

one's own (Fang et al. 2019), suggesting that causal inference may be a mechanism of how the brain distinguishes oneself from the environment.

## 5.4 Neurophysiological Evidence of Body Representation

### 5.4.1 Neuronal Basis of Multisensory Integration of Body-Related Signals

The last section focuses on the neurophysiological and neuroimaging evidence related to multisensory body representation. Since numerous behavioral findings have demonstrated that the body representation depends on the integration of body-related signals from multiple modalities, most studies exploring the neural basis of body representation have focused on the multisensory neurons (Noel et al. 2018; Blanke et al. 2015). Human and non-human primate neurophysiological studies have demonstrated a high degree of overlap of brain regions in multisensory integration and body representation (Blanke et al. 2015). Using single-unit recordings in animals, numerous kinds of multisensory neurons have been identified in various brain regions, including the parietal association cortex, premotor cortex, insula, and superior colliculus (Mountcastle et al. 1975; Avillac et al. 2007; Stein and Stanford 2008). In particular, body-related multisensory neurons are predominantly located in the posterior parietal cortex, including the ventral intraparietal (VIP) area, area 5 and area 7, and premotor cortex (Graziano et al. 1994; Graziano et al. 1997; Avillac et al. 2005; Fogassi et al. 1996; Graziano et al. 1999; Leinonen 1980).

Multisensory integration of body-related signals at the single neuron level has been well studied on the bimodal neurons responding to somatosensory stimuli and visual (auditory) stimuli near the body. In most of these neurons, the visual (auditory) receptive field is localized on a given body part (Graziano et al. 1994, 1999; Avillac et al. 2005). One important multisensory property of these neurons is that the neural

response to the tactile stimulus is modulated by visual (auditory) stimuli presented within their receptive field. Similar to other multisensory neurons (e.g., visual-vestibular multisensory neurons), such multisensory modulation on neural response can be super-additive (increased firing rate) or sub-additive (decreased firing rate) compared to the arithmetic sum of the responses in unimodal conditions (Avillac et al. 2007). In analogy to peripersonal space in human psychological studies, the size of the visual (auditory) receptive field is proportional to the tactile receptive field on a given body part. The visual receptive field of monkey premotor neurons typically extends 40 cm from the upper limb (Graziano et al. 1994, 1999). It ranges from 5 cm to 1 m depending on the size of other body parts (Avillac et al. 2005; Schlack et al. 2005; Jiang et al. 2013). Furthermore, the visual (auditory) receptive field of these multisensory neurons is anchored on the corresponding body part despite the movement of the body (Fogassi et al. 1996; Graziano et al. 1997; Graziano 1999, 2000). For example, Graziano and colleagues found that the visual receptive field of monkey premotor neurons shifts to the new spatial location where the limb is placed (Graziano 1999). Such binding receptive fields can also be observed in other body-centered multisensory neurons, such as face-centered neurons in VIP (Avillac et al. 2005) and trunk-centered in area 7 (Iriki et al. 1996). Thus, these multisensory neurons with an anchored visual (auditory) receptive field encode body-related signals in body-part centered reference frames. Taken together, the visual (auditory) receptive field of these multisensory neurons is thus conceived as the neural basis of peripersonal space in human psychological studies, which constitute an interface for the body–environment interaction (Noel et al. 2018).

Human neuroimaging studies have consistently highlighted the premotor and posterior parietal cortex in integrating body-related signals (Makin et al. 2008). Similar to the single neuron responses in non-human primates, super-additive and sub-additive responses are also observed in the intra-parietal sulcus (IPS) and ventral premotor cortex when the visual stimulus is integrated



within the peripersonal space (Gentile et al. 2011). For example, human fMRI studies found enhanced BOLD response in IPS when participants were presented with a visual object near the body (Makin et al. 2007; Sereno and Huang 2006). Using the BOLD adaptation paradigm, Brozzoli and colleagues examined the neural activations when consecutive visual stimuli were presented in peripersonal space around the participants' hand. Since the adaptation paradigm (reduced neural activity in response to repeatedly presented stimuli) has been well established in electrophysiology and fMRI studies to reflect selectivity of neural responses to a specific stimulus, it is a useful method to examine the neural populations that respond to the visual stimuli within PPS. The adaptation effect was observed in IPS, inferior parietal lobe, and premotor cortex when the visual object was presented near the right hand, but not far from the hand (Brozzoli et al. 2011), indicating that these brain regions selectively respond to inputs near the body. Taken together, both human and animal neurophysiological findings suggest two key regions, the posterior parietal cortex, and premotor cortex, in multisensory processing of bodily signals. Single neuron and population activities in these areas offer a common neural basis in humans and non-human primates for multisensory integration in peripersonal space.

#### 5.4.2 Neuronal Representation of One's Own Body

Despite the rich neurophysiology studies about multisensory processing within peripersonal space in body representation, direct evidence about the neural basis of subjective recognition of one's body is obtained from body illusion studies on humans. Using synchronous visual-tactile stimulation on a viewed hand and the participant's actual hidden hand within fMRI scanner, Ehrsson and colleagues examined brain activation when the participants experienced the rubber hand illusion. They found activation in IPS, ventral premotor cortex, cerebellum (Ehrsson et al. 2004, 2005), and right posterior

insula (Tsakiris et al. 2007) was closely correlated to the change of limb ownership in RHI. These neural correlates can be further confirmed by conducting the threatening test, which is commonly used in the behavioral test in body ownership. Lloyd and colleagues found increased activity in the posterior parietal cortex and supplementary motor cortex when the threatening visual object approached the fake limb in RHI (Lloyd et al. 2006), as if the participant's own body was being threatened.

Furthermore, several studies examined the shift of peripersonal space toward the viewed hand after the rubber hand illusion has been induced. As expected, peripersonal space around the limb is remapped onto the fake limb under the RHI, and the adaptation effect in IPS and premotor cortex was observed when visual stimuli were repeatedly presented near the fake limb, but the adaptation effect is not observed if the contralateral fake limb is presented (Brozzoli et al. 2012).

The most direct evidence about the relationship between the body-related multisensory integration and body representation on a single neuron level has been examined by Graziano and colleagues in monkeys. Using the single-unit recording, the authors examined the neuronal response in monkey area 5 when the animal experienced the visual-somatosensory multisensory condition similar to the rubber hand illusion in humans (Graziano 2000). The multisensory neurons showed spatial selectivity to proprioceptive (monkey's veridical limb) information and visual (fake limb) information of limb positions. For instance, if a neuron has a higher firing rate when the monkey's veridical limb (proprioceptive input) is positioned on the left side versus the right side, its firing rate also increased when the fake visual limb is presented on the left side and decreased when it was shown on the right side. Intriguingly, such tuning modulation by visual information depended on the physiological feature of the visual limb. Non-body objects (e.g., a trunk of wood) and physically impossible limb position did not modulate the neural response. Furthermore, electrical stimulation of these body-related multisensory neurons in non-human primates results in defensive-like actions (Cooke

and Graziano 2003; Graziano and Cooke 2006). These results are comparable to the RHI results in human behavioral and imaging studies.

### 5.4.3 Electrophysiological Evidence of Causal Inference in Body Representation

In the last session, we take the Bayesian theory of multisensory integration into account to explain the neural implementation of the subjective experience of body representation and self-other discrimination. In a recent study, the authors established a moving rubber hand illusion paradigm based on reaching movements under a virtual reality system. They recorded single neuron responses in the premotor cortex in awake behaving monkeys. By introducing various disparities between a monkey's real limb and a visual fake limb, authors can examine the proprioceptive drift, which is the probe of the illusion strength under dynamic multisensory conditions. The behavioral result showed that the limb's integration of visual and proprioceptive information could be well explained by Bayesian causal inference theory, consistent with the human behavioral results (Fang et al. 2019).

More importantly, under the Bayesian causal inference framework, the ownership of the visual fake limb is determined by the posterior probability that the sensory signals come from a common source. That is, the central neural system should integrate the signals when the visual-proprioceptive limbs are aligned and segregate the signals when the disparity is too large. To examine the neural correlates of the posterior probability, the authors conducted two control conditions to establish the ideal neuronal responses of integration and segregation. In the integration control condition, the visual and proprioceptive limb position was always perfectly aligned, while the visual limb was not presented in the segregation control condition. Thus, when the disparity was systematically changed across trials, the representation of the common source probability on a single neuron level can be approached according to how similar the neural

response is to these ideal neural responses. The single neuron analysis revealed a considerable population in the premotor cortex associated with the posterior probability of common source predicted by the Bayesian causal inference model. The dynamics of the posterior probability of common source across trials can also be decoded by the population neuronal activities. More importantly, the probability of integration at both the behavioral and neural levels was decreased when the visual feedback was replaced by a piece of wood (Fang et al. 2019).

The neural mechanism of causal inference was further extended by a recent artificial neural network study. The authors trained a neural network model to solve causal inference for motion estimation (Rideaux et al. 2021). It was suggested that the multisensory neurons with congruent Gaussian tuning may account for multisensory integration, whereas those with incongruent multisensory tuning have been considered to account for segregating (French and DeAngelis 2020). In line with this prediction, the neural network develops multisensory neurons with congruent and opposite tunings and demonstrated both congruent and opposite neurons contribute to the multisensory behavior. This simulation thus showed that to determine whether the signals should be integrated or segregated, the causal inference problem can be solved by balancing between the activities of these two populations.

---

## References

- Alais D, Burr D (2004) Ventriloquist effect results from near-optimal bimodal integration. *Curr Biol* 14(3):257–262. [https://doi.org/10.1016/S0960-9822\(04\)00043-0](https://doi.org/10.1016/S0960-9822(04)00043-0)
- Antal, Haans Wijnand A., IJsselstein Yvonne A.W., de Kort (2008) The effect of similarities in skin texture and hand shape on perceived ownership of a fake limb *Body Image* 5(4) 389–394 <https://doi.org/10.1016/j.bodyim.2008.04.003>
- Austen EL, Soto-Faraco S, Enns JT, Kingstone A (2004) Mislocalizations of touch to a fake hand. *Cogn Affect Behav Neurosci* 4(2):170–181. <https://doi.org/10.3758/CABN.4.2.170>
- Avillac M, Deneve S, Olivier E, Pouget A, Duhamel JR (2005) Reference frames for representing visual

- and tactile locations in parietal cortex. *Nat Neurosci* 8(7):941–949. <https://doi.org/10.1038/nm1480>
- Avillac M, Ben Hamed S, Duhamel JR (2007) Multisensory integration in the ventral intraparietal area of the macaque monkey. *J Neuroscience* 27(8):1922–1932. <https://doi.org/10.1523/jneurosci.2646-06.2007>
- Blanke O, Slater M, Serino A (2015) Behavioral, neural, and computational principles of bodily self-consciousness. *Neuron* 88(1):145–166. <https://doi.org/10.1016/j.neuron.2015.09.029>
- Botvinick M, Cohen J (1998) Rubber hands ‘feel’ touch that eyes see. *Nature* 391:6669–6671. <https://doi.org/10.1038/35784>
- Brozzoli C, Gentile G, Petkova VI, Ehrsson HH (2011) fMRI adaptation reveals a cortical mechanism for the coding of space near the hand. *J Neurosci* 31(24):9023–9031. <https://doi.org/10.1523/JNEUROSCI.1172-11.2011>
- Brozzoli C, Gentile G, Ehrsson HH (2012) That’s near my hand! Parietal and premotor coding of hand-centered space contributes to localization and self-attribution of the hand. *J Neurosci* 32(42):14573–14582. <https://doi.org/10.1523/jneurosci.2660-12.2012>
- Buxbaum LJ, Coslett HB (2010) Specialised structural descriptions for human body parts: evidence from autotopagnosia. *Cogn Neuropsychol* 18(4):289–306. <https://doi.org/10.1080/02643290126172>
- Cooke DF, Graziano MS (2003) Defensive movements evoked by air puff in monkeys. *J Neurophysiol* 90(5):3317–3329. <https://doi.org/10.1152/jn.00513.2003>
- Cooper LA, Shepard RN (1975) Mental transformation in the identification of left and right hands. *J Exp Psychol Hum Percept Perform* 1(1):48–56. <https://doi.org/10.1037/0096-1523.1.1.48>
- Costantini M, Haggard P (2007) The rubber hand illusion: sensitivity and reference frame for body ownership. *Conscious Cogn* 16(2):229–240. <https://doi.org/10.1016/j.concog.2007.01.001>
- David R., Wozny Ulrik R., Beierholm Ladan, Shams (2010) Probability Matching as a Computational Strategy Used in Perception *PLoS Computational Biology* 6(8):e1000871. <https://doi.org/10.1371/journal.pcbi.1000871>
- Ehrsson HH, Spence C, Passingham RE (2004) That’s my hand! Activity in premotor cortex reflects feeling of ownership of a limb. *Science* 305(5685):875–877. <https://doi.org/10.1126/SCIENCE.1097011>
- Ehrsson HH, Holmes NP, Passingham RE (2005) Touching a rubber hand: feeling of body ownership is associated with activity in multisensory brain areas. *J Neurosci* 25(45):10564–10573. <https://doi.org/10.1523/JNEUROSCI.0800-05.2005>
- Ernst MO, Banks MS (2002) Humans integrate visual and haptic information in a statistically optimal fashion. *Nature* 415(6870):429–433. <https://doi.org/10.1038/415429a>
- Ernst MO, Bühlhoff HH (2004) Merging the senses into a robust percept. *Trends Cogn Sci* 8(4):162–169. <https://doi.org/10.1016/j.tics.2004.02.002>
- Fang W, Li J, Qi G, Li S, Sigman M, Wang L (2019) Statistical inference of body representation in the macaque brain. *Proc Natl Acad Sci U S A* 116(40):20151–20157. <https://doi.org/10.1073/PNAS.1902334116/-DCSUPPLEMENTAL>
- Farmer H, Tajadura-Jiménez A, Tsakiris M (2012) Beyond the colour of my skin: how skin colour affects the sense of body-ownership. *Conscious Cogn* 21(3):1242–1256. <https://doi.org/10.1016/j.concog.2012.04.011>
- Farmer H, Maister L, Tsakiris M (2014) Change my body, change my mind: the effects of illusory ownership of an outgroup hand on implicit attitudes toward that outgroup. *Front Psychol* 4(JAN):1016. <https://doi.org/10.3389/FPSYG.2013.01016/BIBTEX>
- Fogassi L, Gallese V, Fadiga L, Luppino G, Matelli M, Rizzolatti G (1996) Coding of peripersonal space in inferior premotor cortex (area F4). *J Neurophysiol* 76(1):141–157. <https://doi.org/10.1152/jn.1996.76.1.141>
- French RL, DeAngelis GC (2020) Multisensory neural processing: from cue integration to causal inference. *Curr Opin Physiol* 16:8–13. <https://doi.org/10.1016/j.cophys.2020.04.004>
- Gentile G, Petkova VI, Ehrsson HH (2011) Integration of visual and tactile signals from the hand in the human brain: an fMRI study. *J Neurophysiol* 105(2):910–922. Retrieved from <http://jn.physiology.org/content/105/2/910.full.pdf>
- Graziano MSA (1999) Where is my arm? The relative role of vision and proprioception in the neuronal representation of limb position. *Proc Natl Acad Sci U S A* 96(18):10418–10421. <https://doi.org/10.1073/pnas.96.18.10418>
- Graziano MSA (2000) Coding the location of the arm by sight. *Science* 290(5497):1782–1786. <https://doi.org/10.1126/science.290.5497.1782>
- Graziano MS, Cooke DF (2006) Parieto-frontal interactions, personal space, and defensive behavior. *Neuropsychologia* 44(13):2621–2635. Retrieved from <https://www.ncbi.nlm.nih.gov/pubmed/17128446>
- Graziano MS, Yap GS, Gross CG (1994) Coding of visual space by premotor neurons. *Science* 266(5187):1054–1057. <https://doi.org/10.1126/science.7973661>
- Graziano MS, Hu XT, Gross CG (1997) Visuospatial properties of ventral premotor cortex. *J Neurophysiol* 77(5):2268–2292. <https://doi.org/10.1152/jn.1997.77.5.2268>
- Graziano MS, Reiss LA, Gross CG (1999) A neuronal representation of the location of nearby sounds. *Nature* 397(6718):428–430. <https://doi.org/10.1038/17115>
- Head and Holmes (1911) Sensory Disturbances From Cerebral Lesions *Brain* 34(2–3):102–254. <https://doi.org/10.1093/brain/34.2-3.102>
- Holmes NP, Spence C (2005) Visual bias of unseen hand position with a mirror: spatial and temporal factors. *Exp Brain Res* 166(3–4):489–497. <https://doi.org/10.1007/s00221-005-2389-4>
- Holmes NP, Crozier G, Spence C (2004) When mirrors lie: “visual capture” of arm position impairs reaching

- performance. *Cogn Affect Behav Neurosci* 4(2):193–200. <https://doi.org/10.3758/CABN.4.2.193>
- Holmes NP, Snijders HJ, Spence C (2006) Reaching with alien limbs: visual exposure to prosthetic hands in a mirror biases proprioception without accompanying illusions of ownership. *Perception & Psychophysics* 2006 68:4 68(4):685–701. <https://doi.org/10.3758/BF03208768>
- Ide M (2013) The effect of “anatomical plausibility” of hand angle on the rubber-hand illusion. *Perception* 42:103–111. <https://doi.org/10.1068/p7322>
- Iriki A, Tanaka M, Iwamura Y (1996) Coding of modified body schema during tool use by macaque postcentral neurones. *Neuroreport* 7(14):2325–2330. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/8951846>
- Jiang HH, Hu YZ, Wang JH, Ma YY, Hu XT (2013) Visuospatial properties of caudal area 7b in Macaca fascicularis. *Dongwuxue Yanjiu* 34(E2):E50–E61. <https://doi.org/10.3724/SP.J.1141.2013.E02E50>
- Kilteni K, Maselli A, Kording KP, Slater M (2015) Over my fake body: body ownership illusions for studying the multisensory basis of own-body perception. *Front Hum Neurosci* 9(MAR):141. <https://doi.org/10.3389/FNHUM.2015.00141/BIBTEX>
- Körding KP, Beierholm U, Ma WJ, Quartz S, Tenenbaum JB, Shams L (2007) Causal inference in multisensory perception. *PLoS One* 2(9):e943. <https://doi.org/10.1371/journal.pone.0000943>
- Lådavas et al. (1998) Visual peripersonal space centred on the face in humans *Brain* 121(12):2317–2326. <https://doi.org/10.1093/brain/121.12.2317>
- Leinonen L (1980) Functional properties of neurones in the posterior part of area 7 in awake monkey. *Acta Physiol Scand* 108(3):301–308. <https://doi.org/10.1111/j.1748-1716.1980.tb06536.x>
- Liu Y, Medina J (2017) Influence of the body schema on multisensory integration: evidence from the Mirror box illusion. *Scientific Reports* 2017 7:1 7(1):1–11. <https://doi.org/10.1038/s41598-017-04797-0>
- Liu Y, Medina J (2018) Integrating multisensory information across external and motor-based frames of reference. *Cognition* 173:75–86. <https://doi.org/10.1016/j.cognition.2018.01.005>
- Lloyd DM (2007) Spatial limits on referred touch to an alien limb may reflect boundaries of visuo-tactile peripersonal space surrounding the hand. *Brain Cogn* 64(1):104–109. <https://doi.org/10.1016/j.bandc.2006.09.013>
- Lloyd D, Morrison I, Roberts N (2006) Role for human posterior parietal cortex in visual processing of aversive objects in peripersonal space. *J Neurophysiol* 95(1):205–214. <https://doi.org/10.1152/jn.00614.2005>
- Makin TR, Holmes NP, Zohary E (2007) Is that near my hand? Multisensory representation of peripersonal space in human intraparietal sulcus. *J Neurosci* 27(4):731–740. <https://doi.org/10.1523/JNEUROSCI.3653-06.2007>
- Makin TR, Holmes NP, Ehrsson HH (2008) On the other hand: dummy hands and peripersonal space. *Behav Brain Res* 191(1):1–10. <https://doi.org/10.1016/j.bbr.2008.02.041>
- Medina J, Khurana P, Coslett HB (2015) The influence of embodiment on multisensory integration using the mirror box illusion. *Conscious Cogn* 37:71–82. <https://doi.org/10.1016/j.concog.2015.08.011>
- Meredith MA, Nemitz JW, Stein BE (1987) Determinants of multisensory integration in superior colliculus neurons. I. Temporal factors. *J Neurosci* 7(10):3215–3229. doi: citeulike-article-id:409430
- Meredith MA, Stein BE (1986) Spatial factors determine the activity of multisensory neurons in cat superior colliculus. *Brain Res* 5:350–354. [https://doi.org/10.1016/0006-8993\(86\)91648-3](https://doi.org/10.1016/0006-8993(86)91648-3)
- Mountcastle VB, Lynch JC, Georgopoulos A, Sakata H, Acuna C (1975) Posterior parietal association cortex of the monkey: command functions for operations within extrapersonal space. *J Neurophysiol* 38(4):871–908. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/808592>
- Noel JP, Blanke O, Serino A (2018) From multisensory integration in peripersonal space to bodily self-consciousness: from statistical regularities to statistical inference. *Ann N Y Acad Sci* 1426(1):146–165. <https://doi.org/10.1111/nyas.13867>
- Parsons LM (1994) Temporal and kinematic properties of motor behavior reflected in mentally simulated action. *J Exp Psychol Hum Percept Perform* 20(4):709–730. <https://doi.org/10.1037/0096-1523.20.4.709>
- Parsons (1987) Imagined spatial transformation of one’s body. *Journal of Experimental Psychology: General* 116(2):172–191. <https://doi.org/10.1037/0096-3445.116.2.172>
- Pavani and Zampini (2007) The Role of Hand Size in the Fake-Hand Illusion Paradigm *Perception* 36(10):1547–1554. <https://doi.org/10.1068/p5853>
- Peck TC, Seinfeld S, Aglioti SM, Slater M (2013) Putting yourself in the skin of a black avatar reduces implicit racial bias. *Conscious Cogn* 22(3):779–787. <https://doi.org/10.1016/j.concog.2013.04.016>
- Ramachandran and Rogers-Ramachandran (1996) Synaesthesia in phantom limbs induced with mirrors *Proceedings of the Royal Society of London. Series B: Biological Sciences* 263(1369):377–386. <https://doi.org/10.1098/rspb.1996.0058>
- Rideaux R, Storrs KR, Maiello G, Welchman AE (2021) How multisensory neurons solve causal inference. *Proc Natl Acad Sci U S A* 118(32):e2106235118. <https://doi.org/10.1073/pnas.2106235118>
- Samad M, Chung AJ, Shams L (2015) Perception of body ownership is driven by Bayesian sensory inference. *PLoS One* 10(2):1–23. <https://doi.org/10.1371/journal.pone.0117178>
- Schlack A, Sterbing-D’Angelo SJ, Hartung K, Hoffmann KP, Bremmer F (2005) Multisensory space representations in the macaque ventral intraparietal area. *J Neurosci* 25(18):4616–4625. <https://doi.org/10.1523/JNEUROSCI.0455-05.2005>
- Schwoebel J, Coslett HB (2005) Evidence for multiple, distinct representations of the human body.

- J Cogn Neurosci 17(4):543–553. <https://doi.org/10.1162/0898929053467587>
- Sereno MI, Huang RS (2006) A human parietal face area contains aligned head-centered visual and tactile maps. *Nat Neurosci* 9(10):1337–1343. <https://doi.org/10.1038/nn1777>
- Shams L, Beierholm UR (2010) Causal inference in perception. *Trends Cogn Sci* 14(9):425–432. <https://doi.org/10.1016/j.tics.2010.07.001>
- Spence C, Pavani F, Driver J (2004) Spatial constraints on visual-tactile cross-modal distractor congruency effects. *Cogn Affect Behav Neurosci* 4(2):148–169. <https://doi.org/10.3758/CABN.4.2.148>
- Stein BE, Stanford TR (2008) Multisensory integration: current issues from the perspective of the single neuron. *Nat Rev Neurosci* 9(4):255–266. <https://doi.org/10.1038/nrn2331>
- Tsakiris M (2010) My body in the brain: a neurocognitive model of body-ownership. *Neuropsychologia* 48(3):703–712. <https://doi.org/10.1016/j.neuropsychologia.2009.09.034>
- Tsakiris M, Haggard P (2005) The rubber hand illusion revisited: Visuotactile integration and self-attribution. *J Exp Psychol Hum Percept Perform* 31(1):80–91. <https://doi.org/10.1037/0096-1523.31.1.80>
- Tsakiris M, Hesse MD, Boy C, Haggard P, Fink GR (2007) Neural signatures of body ownership: a sensory network for bodily self-consciousness. *Cereb Cortex* 17(10):2235–2244. <https://doi.org/10.1093/cercor/bhl131>
- Tsakiris M, Costantini M, Haggard P (2008) The role of the right temporo-parietal junction in maintaining a coherent sense of one's body. *Neuropsychologia* 46(12):3014–3018. <https://doi.org/10.1016/j.neuropsychologia.2008.06.004>
- Tsakiris M, Carpenter L, James D, Fotopoulou A (2010) Hands only illusion: multisensory integration elicits sense of ownership for body parts but not for non-corporeal objects. *Exp Brain Res* 204(3):343–352. <https://doi.org/10.1007/s00221-009-2039-3>
- Van Beers RJ, Sittig AC, Denier Van Der Gon JJ (1999) Integration of proprioceptive and visual position-information: an experimentally supported model. *J Neurophysiol* 81(3):1355–1364. <https://www.ncbi.nlm.nih.gov/pubmed/10085361>
- Witten IB, Knudsen EI (2005) Why seeing is believing: merging auditory and visual worlds. *Neuron* 48(3):489–496. <https://doi.org/10.1016/j.neuron.2005.10.020>
- Wozny DR, Beierholm UR, Shams L (2008) Human trimodal perception follows optimal statistical inference. *J Vis* 8(3):24. <https://doi.org/10.1167/8.3.24>
- Zapparoli L, Invernizzi P, Gandola M, Berlinger M, De Santis A, Zerbi A et al (2014) Like the back of the (right) hand? A new fMRI look on the hand laterality task. *Exp Brain Res* 232(12):3873–3895