



Associative account of self-cognition: extended forward model and multi-layer structure

Motoaki Sugiura^{1,2*}

¹ Institute of Development, Aging and Cancer, Tohoku University, Sendai, Japan

² International Research Institute of Disaster Science, Tohoku University, Sendai, Japan

Edited by:

Niall W. Duncan, University of Ottawa, Canada

Reviewed by:

David Rudrauf, Institut National de la Santé et de la Recherche Médicale, France

Francesca Ferri, University of Ottawa, Canada

*Correspondence:

Motoaki Sugiura, Department of Functional Brain Imaging, Institute of Development, Aging and Cancer, Tohoku University, Seiryomachi 4-1, Aoba-ku, Sendai 980-8575, Japan
e-mail: motoaki@dac.tohoku.ac.jp

The neural correlates of “self” identified by neuroimaging studies differ depending on which aspects of self are addressed. Here, three categories of self are proposed based on neuroimaging findings and an evaluation of the likely underlying cognitive processes. The physical self, representing self-agency of action, body-ownership, and bodily self-recognition, is supported by the sensory and motor association cortices located primarily in the right hemisphere. The interpersonal self, representing the attention or intentions of others directed at the self, is supported by several amodal association cortices in the dorsomedial frontal and lateral posterior cortices. The social self, representing the self as a collection of context-dependent social-values, is supported by the ventral aspect of the medial prefrontal cortex and the posterior cingulate cortex. Despite differences in the underlying cognitive processes and neural substrates, all three categories of self are likely to share the computational characteristics of the forward model, which is underpinned by internal schema or learned associations between one’s behavioral output and the consequential input. Additionally, these three categories exist within a hierarchical layer structure based on developmental processes that updates the schema through the attribution of prediction error. In this account, most of the association cortices critically contribute to some aspect of the self through associative learning while the primary regions involved shift from the lateral to the medial cortices in a sequence from the physical to the interpersonal to the social self.

Keywords: self, self-recognition, self-awareness, body-ownership, self-agency, social cognition, associative learning, neuroimaging

INTRODUCTION

A sporadic quest for the neural basis of “self” using functional neuroimaging appears to have emerged at the end of the last century. A number of researchers were interested in the cognitive processes related to physical self-awareness during action and consequential sensation (McGuire et al., 1996a; Blakemore et al., 1998; Fink et al., 1999), whereas others investigated the self-relevance of memory and knowledge (Fink et al., 1996; Craik et al., 1999; Kelley et al., 2002). An initial study evaluating stimulus-independent thought (McGuire et al., 1996b) drew attention to the relevance of spontaneous neural activity during a conscious resting state (Raichle et al., 2001) to the self-related cognitive process (Gusnard et al., 2001). The subsequent surge in studies of self-face recognition (Keenan et al., 2000; Kircher et al., 2000; Sugiura et al., 2000) rested primarily on an evolutionary or developmental perspective. Perceptions of others’ communicative intentions toward the self (Kampe et al., 2003) and perspective-taking (Vogeley et al., 2004) appear to be other independent issues. It did not take long for researchers to realize that the cortical regions supporting self-specific or self-relevant activation were far from consistent across studies.

The response of researchers to this chaotic situation has also varied considerably. Some have been pessimistic regarding the

existence of a special neural system for the self (Gillihan and Farah, 2005; Platek et al., 2008), whereas others have remained optimistic and attempted to identify such a system by sorting out how previous studies have addressed the concept of self. One such approach highlighted cortical midline structures (Northoff and Bermpohl, 2004; Northoff et al., 2006), another focused on the right lateral cortices (Keenan et al., 2000; Feinberg and Keenan, 2005), and yet another tried to reconcile these two views (Uddin et al., 2007).

This paper attempts to provide a unified framework for the neural underpinnings of the self in the context of a pessimistic stance toward the existence of a self-specific neural system. First, neuroimaging studies investigating self-related processes are reviewed, and the concepts of the self and their related brain networks are roughly categorized into three areas. Then, a unique characteristic or computational architecture that is potentially common to the processes of the three categories is proposed. Furthermore, I propose a layer structure characterized by cross-layer dynamics that operates across these three categories. Finally, the manner in which self is related to midline brain regions in this model is discussed. The proposed model is an updated version of one that was previously presented (Sugiura, 2011).

THREE CATEGORIES OF SELF AND NEUROIMAGING FINDINGS

The concepts of self or self-related processes addressed in previous neuroimaging studies may be divided into three categories: the physical self, the interpersonal self, and the social-value of the self. This categorization takes into account the presumably developmental context in which self-awareness is experienced as well as the distribution of reported self-relevant activation. The known basic functional characteristics of the relevant activated regions will be briefly discussed.

PHYSICAL SELF

One prominent category of self is the body-grounded self that dissociates one's physical existence from the external environment. Research on this category often focuses on the ability to dissociate self from non-self, such as one's own face from another's face, one's own body from another's body, one's own action from another's action, and one's own voice from another's voice. This category of self is conceptually unambiguous and experimental manipulation is often clear. The concept of self in this category may overlap with that of James' description of "physical self" (James, 1890) or other researchers' descriptions of "ecological self" (Neisser, 1988) and "minimal self" (Gallagher, 2000). This category of self is of central interest to psychologists studying animals and infants. There is a strong expectation of the existence of specialized neural regions underpinning self-recognition that stems from the fact that only a few species of animals show evidence of visual self-recognition in the mirror (Gallup, 1982; Suddendorf and Collier-Baker, 2009).

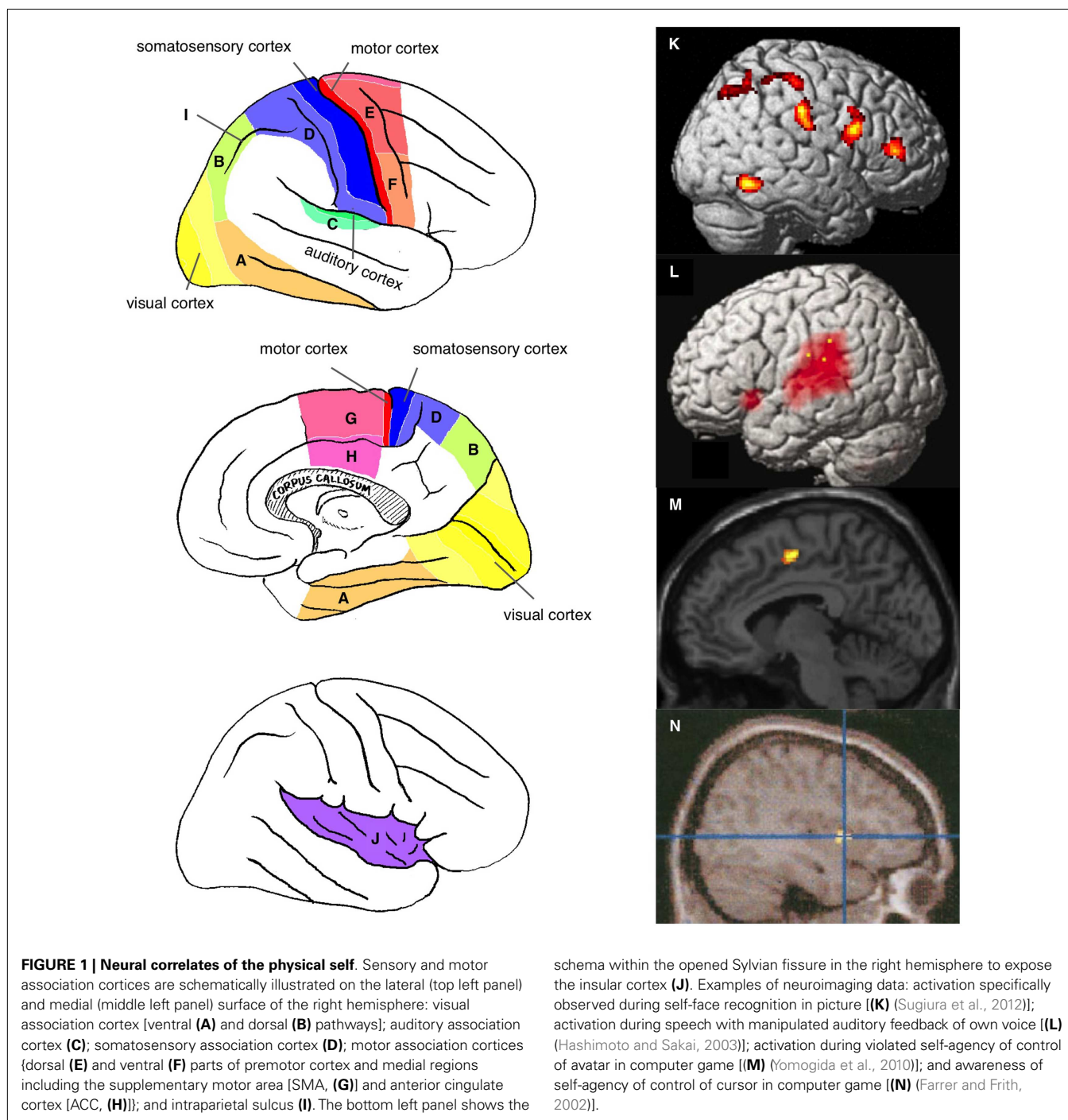
Neuroimaging studies investigating the cortical foundation of the physical self have adopted three major experimental approaches. The first approach is to contrast brain activation during the perception of a visual or auditory stimulus relevant to one's own body with activation during the perception of perceptually similar but self-irrelevant stimuli. Studies using this approach present subjects with a picture or video clip of a face or body (Kircher et al., 2000; Platek et al., 2004, 2006; Sugiura et al., 2005a, 2006, 2008, 2012; Uddin et al., 2005; Devue et al., 2007; Kaplan et al., 2008; Ferri et al., 2012; Oikawa et al., 2012), or a recorded voice (Nakamura et al., 2001); the required task is either explicit or implicit recognition (e.g., passive viewing or performance of an unrelated task). The second approach addresses the sense of body-ownership or of body-location drift, which is illusorily induced by a synchronous sensory stimulation including tactile stimuli (Ehrsson et al., 2004, 2005; Tsakiris et al., 2007; Ionta et al., 2011). Contrasting the synchronized and desynchronized conditions can isolate the cortical activation related to such a sense. The third approach deals with the sense of self-agency or self-attribution concerning one's own actions. Studies have identified neural activation in response to modulated visual feedback during hand action or manipulation of a cursor or agent on a computer (Fink et al., 1999; Farrer et al., 2003, 2008; Leube et al., 2003; David et al., 2007; Schnell et al., 2007; Corradi-Dell'Acqua et al., 2008; Spengler et al., 2009; Yomogida et al., 2010), auditory feedback during speech (McGuire et al., 1996a; Hashimoto and Sakai, 2003; Fu et al., 2006), and tactile feedback while tickling oneself (Blakemore et al., 1998). Some studies have manipulated self-agency simply by instruction (Farrer and Frith, 2002; Schnell

et al., 2007), whereas others have examined effects of the trial-by-trial fluctuation in subjective awareness in response to the same stimuli (David et al., 2007; Farrer et al., 2008).

Although the regions reportedly involved in this activation vary across studies and approaches, they include primarily the sensory and/or motor association cortices (**Figure 1**) and depend on the sensory modality of the stimulus used. Activation of the visual association cortices, including the ventral and dorsal pathways (**Figures 1A,B**, respectively), has been reported in studies using visual stimuli to address visual self-face or self-body recognition (Kircher et al., 2000; Sugiura et al., 2005a, 2006, 2008, 2012; Uddin et al., 2005; Platek et al., 2006; Kaplan et al., 2008; Ferri et al., 2012; Oikawa et al., 2012), the illusory sense of body-ownership or location (Ehrsson et al., 2004; Tsakiris et al., 2007; Ionta et al., 2011), and the violation or awareness of action-agency (Fink et al., 1999; David et al., 2007; Corradi-Dell'Acqua et al., 2008; Farrer et al., 2008; Spengler et al., 2009; Yomogida et al., 2010). Auditory association cortices (**Figure 1C**) are activated during the perception of manipulated feedback of self-voice during speaking aloud (McGuire et al., 1996a; Hashimoto and Sakai, 2003; Fu et al., 2006). Activation of somatosensory association cortices (**Figure 1D**) has been reported in studies using tactile input to manipulate the agency of self-tickling actions (Blakemore et al., 1998) or to induce an illusory sense of body-ownership or location (Ehrsson et al., 2004; Tsakiris et al., 2007; Ionta et al., 2011).

Activation of motor association cortices is frequently reported in studies in which a subject's motor action plays a critical role in self-relevance. These regions include the dorsal and ventral aspects of the premotor cortex (**Figures 1E,F**, respectively) and several medial motor association cortices, such as the supplementary motor area (SMA; **Figure 1G**) and cingulate motor area (**Figure 1H**). Examples of such studies include those in which subjects executed motor action while self-agency was manipulated (Farrer and Frith, 2002; Farrer et al., 2003, 2008; David et al., 2007; Schnell et al., 2007; Corradi-Dell'Acqua et al., 2008; Spengler et al., 2009; Yomogida et al., 2010).

However, the involvement of sensory or motor association cortices sometimes has no apparent relevance to the sensory processing of the stimulus or motor output during the task. Such involvement requires explanation in terms of the internal representation of the physical self. For example, activation of somatosensory and premotor cortices has often been reported in studies investigating self-face or self-body recognition using pictures (Uddin et al., 2005; Platek et al., 2006; Sugiura et al., 2006, 2008, 2012; Ferri et al., 2012). The visual-somatosensory association cortex in the intraparietal sulcus (**Figure 1I**), which has been implicated in the visuospatial motor control of extremities, has been found to be activated in many studies on self when the task is relevant to bodily action either directly (Fink et al., 1999; Ehrsson et al., 2004, 2005; Farrer et al., 2008) or indirectly (e.g., self-face or self-body recognition using pictures involving expressions or actions) (Sugiura et al., 2005a, 2006, 2008, 2012; Oikawa et al., 2012). Additionally, many of these activated areas overlap with regions receiving vestibular input, such as the medial temporal (MT) or medial superior temporal (MST) areas, the ventral intraparietal area (VIP), areas 2v and 3aV, and premotor regions (Smith et al., 2012; zu Eulenburg et al., 2012). The insula (**Figure 1J**), known to include primary and



association cortices for interoception (sense of the physiological conditions of the entire body) and to be involved in a wide range of subjective feelings (Craig, 2002, 2009), is also activated without the manipulation of interoceptive input. Activation of this region was observed during the recognition of self-face or self-body in a picture (Kircher et al., 2000; Devue et al., 2007; Ferri et al., 2012), the sense of action self-agency (Farrer and Frith, 2002; Farrer et al., 2003; Leube et al., 2003; David et al., 2007; Corradi-Dell'Acqua et al., 2008), and the sense of body-ownership (Ehrsson et al.,

2004; Tsakiris et al., 2007). These findings may be explained by the fact that bodily self-recognition is grounded by the experience of bodily action accompanied by visual, somatosensory, vestibular, and interoceptive feedback. These interpretations in terms of the representational role of the sensory and motor association cortices will be detailed in Section "Physical Self and Sensorimotor Schema."

It is interesting to note that some of these parietal sensory association and frontal premotor cortices coincide with the

visual–motor association system known as the mirror neuron system (MNS). This apparently contradicts the incompatible concept of “self-specific” and the MNS; which is resolved in the proposed model (See Physical Self and Sensorimotor Schema). Mirror neurons are a class of neurons that have been observed to discharge when a monkey performs a goal-directed motor act as well as when a monkey observes another individual performing the same or a similar motor act (Rizzolatti et al., 2001; Nelissen et al., 2011). In humans, the MNS has been identified as a homolog of the frontoparietal network of mirror neurons in monkeys and is considered to play a critical role in action understanding, imitation, and communication (Rizzolatti and Craighero, 2004; Iacoboni, 2005). Therefore, this system is primarily conceptualized as a mechanism involved in recognizing and interacting with others.

Moreover, these studies reported the activation of several amodal association cortices. Activation of the right lateral prefrontal cortex, specifically the inferior and middle frontal gyri, has often been reported during the recognition of self-face or self-body (Figure 1K) (Platek et al., 2004, 2006; Sugiura et al., 2005a, 2006, 2008, 2012; Uddin et al., 2005; Devue et al., 2007; Kaplan et al., 2008), voice (Nakamura et al., 2001), action-agency violation (Figure 1L) (Fink et al., 1999; Hashimoto and Sakai, 2003; David et al., 2007; Schnell et al., 2007; Farrer et al., 2008), and body-ownership (Ehrsson et al., 2004, 2005; Tsakiris et al., 2007). The manipulation of sensory-feedback for action often activates the temporoparietal junction (TPJ), the posterior part of the superior temporal sulcus (pSTS), and the medial prefrontal cortex (MPFC) (McGuire et al., 1996a; Farrer and Frith, 2002; Farrer et al., 2003, 2008; Hashimoto and Sakai, 2003; Leube et al., 2003; Fu et al., 2006; Spengler et al., 2009; Yomogida et al., 2010), which are typically considered multimodal or amodal association cortices that are implicated in conceptual rather than perceptual processes. These findings will be discussed separately from sensory or motor association cortices in Section “Multi-Layer Structure and Cross-Layer Dynamics.”

INTERPERSONAL SELF

When an individual notices that he or she is being looked at or hears his/her own name being called, he/she becomes aware that the attention or intentionality of another person is directed at him/her. This awareness is a basic mindset during social interaction. This aspect of self is obviously distinct from the physical self because it inherently requires the existence of another person. An influential inventory, the Self-Consciousness Scale (Fenigstein et al., 1975), particularly its public subscale, has been developed to measure the degree to which an individual has this type of awareness.

The activation related to this awareness is observed in several amodal association cortices in the medial frontal and lateral posterior cortices (Figure 2A). Although varying widely across studies, activation has been identified in the MPFC encompassing the adjacent anterior cingulate cortex (ACC) (Kampe et al., 2003; Schilbach et al., 2006; Steuwe et al., 2012), the TPJ/pSTS (Pelphrey et al., 2004a; Schilbach et al., 2006; Steuwe et al., 2012), the anterior temporal cortex (ATC) (Kawashima et al., 1999; Calder et al., 2002; Kampe et al., 2003; Wicker et al., 2003), the insula (Kawashima et al., 1999; Calder et al., 2002; Schilbach et al., 2006), and the

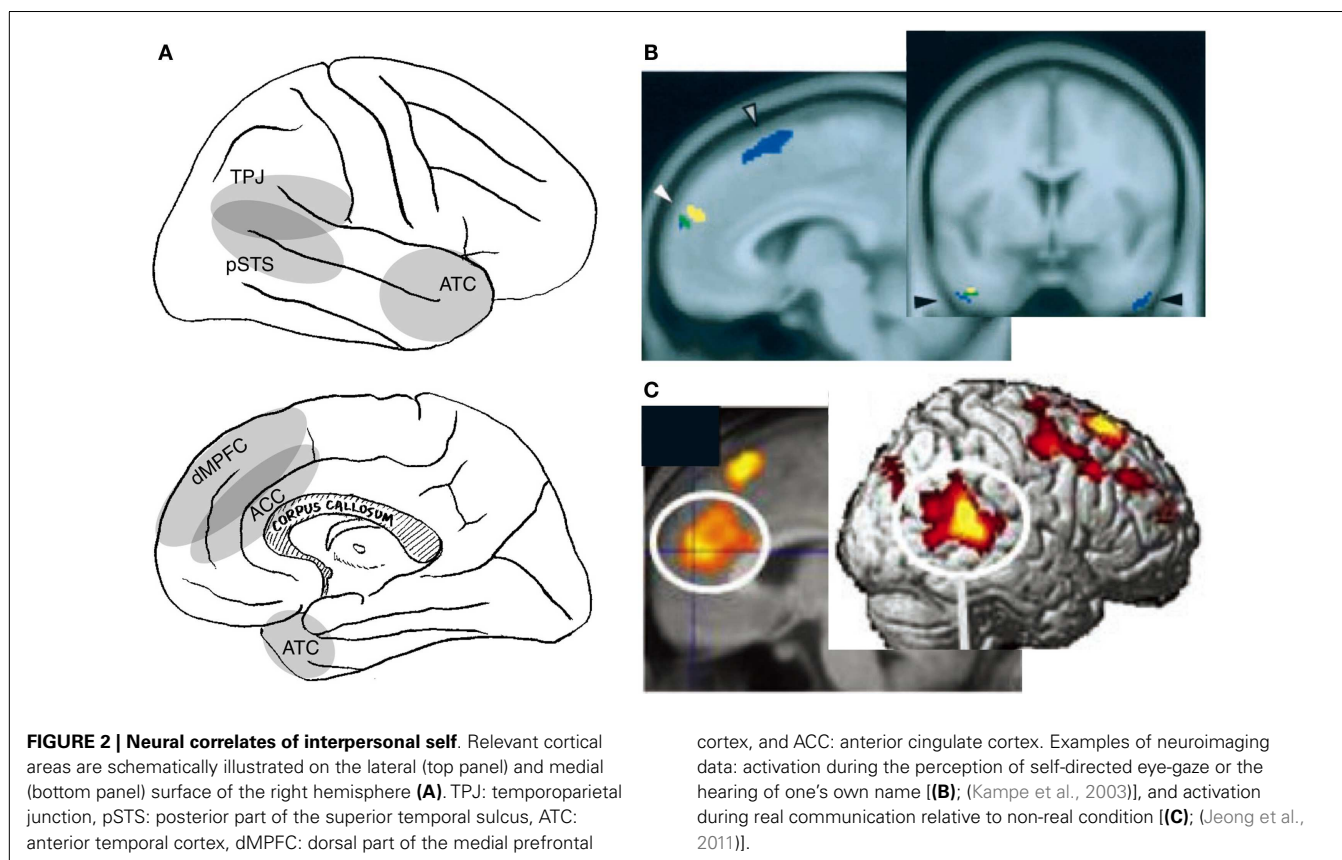
cerebellum (George et al., 2001; Wicker et al., 2003; Schilbach et al., 2006) during the perception of directed, rather than averted, eye-gaze. Activation of the MPFC/ACC, TPJ/pSTS, and ATC has also been reported in studies that compare activation during the hearing of one’s own name with the hearing of others’ names (Figure 2B) (Kampe et al., 2003; Perrin et al., 2005; Tacikowski et al., 2011). Activation is also observed in these regions when subjects believe that they are interacting with a real person rather than engaging in a similar but non-real interaction (Figure 2C) (Rilling et al., 2004; Jeong et al., 2011). Additionally, subjects who score higher on the Self-Consciousness Scale (Fenigstein et al., 1975) show a larger degree of activation in the dorsal part of the MPFC (dMPFC) during a simple sensorimotor (deviant letter detection) task (Eisenberger et al., 2005).

However, previous studies have rarely treated self-awareness as a central concept related to the interpretation of activation in these cortical regions. These regions have often been recognized as a cortical network supporting the inference of another’s mental state, namely, mentalizing or theory of mind (ToM) (Gallagher and Frith, 2003; Frith and Frith, 2006; Senju and Johnson, 2009; Spreng et al., 2009). Furthermore, it has been proposed that this network plays a role in the development of event schemata in general, including person-schema and self-schema (Krueger et al., 2009). The assumed properties of this network provide the basic reasoning for labeling this category of self the “interpersonal self,” which will be detailed in Section “Interpersonal Self and Interpersonal Schema.”

It is worth noting that some of these regions are deactivated rather than activated during self-face recognition. Activation in the TPJ/pSTS or its surrounding cortices is decreased while viewing the self-face compared with familiar or unfamiliar faces (Sugiura et al., 2005a, 2008; Uddin et al., 2005; Devue et al., 2007; Morita et al., 2008), which indicates a clear neural dissociation between physical and interpersonal selves.

SOCIAL-VALUE OF SELF

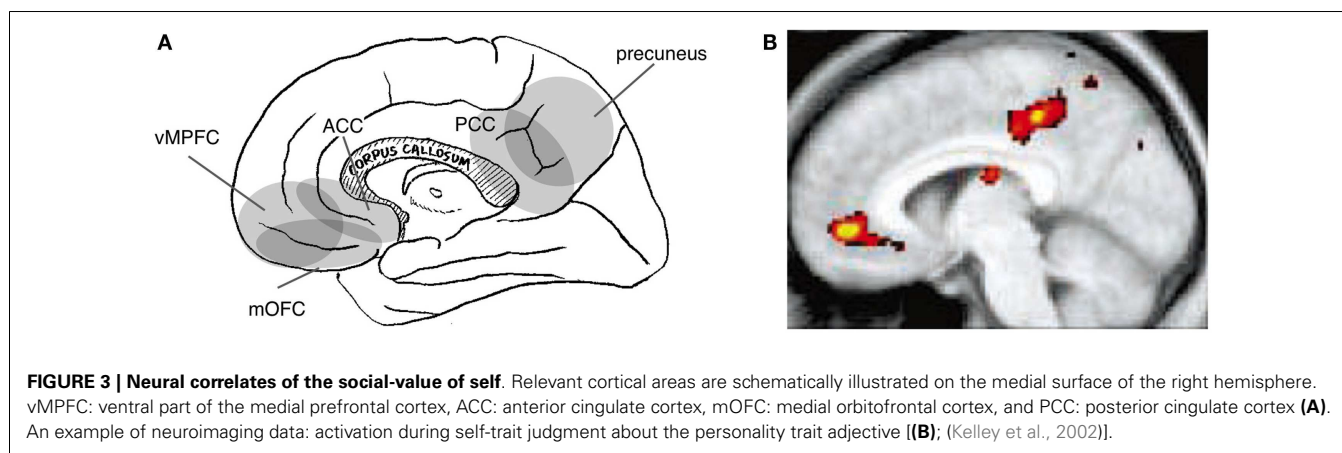
Self-reflection typically includes thoughts about one’s social-value such as “Am I good-natured?” or “Am I good-looking?” or “Am I intelligent?” or “Am I successful in my career?” Most of the attributes assigned to the self carry some social-value, and individuals are typically aware of the gap between one’s current self and one’s ideal self (Festinger, 1954; Higgins, 1987). This type of social-value is an important aspect of the “social self” according to James (1890) and is assumed to be an important determinant of human behavior [e.g., Self-Efficacy Theory (Bandura, 1982)]. To experimentally address this type of self in neuroimaging studies, self-trait (e.g., personality trait, ability) judgment tasks are typically utilized. Additionally, the perception of the evaluation of self by others, even the perception of others who have a high or low level of an attribute that is significant to the self (Gutierrez et al., 1999), is known to affect self-value. Although this type of self resembles the interpersonal self in that it is highly relevant to the existence of another person, the “person” is typically generalized to people or society rather than confined to a specific person. Furthermore, the interpersonal self does not necessarily involve social-value. It is therefore reasonable to categorize self-value separately from interpersonal self.



Indeed, the cortical regions implicated in self-value have, at least in part, a different distribution than do those implicated in the interpersonal self. Specifically, tasks that are assumed to manipulate the social-value of self typically activate the ventral part of the MPFC (vMPFC) and the posterior part of the cingulate cortex (PCC) or its adjacent medial parietal cortex (i.e., the precuneus) (**Figure 3A**). Activation of these regions has been reported during self-trait judgment, specifically when contrasted with trait-valence judgment (Craig et al., 1999; Schmitz et al., 2004) or other trait judgments (**Figure 3B**) (Craig et al., 1999; Kelley et al., 2002; Heatherton et al., 2006; D'Argembeau et al., 2007). Similarly, activation in these regions has been identified when contrasting self-descriptive and non-descriptive trait adjectives (Kircher et al., 2002; Macrae et al., 2004) and when the trait adjective is correlated with self-descriptiveness (Moran et al., 2006). Moreover, the perception of the evaluation of self by others activates these regions (Izuma et al., 2008), particularly in subjects whose self-evaluation is vulnerable to evaluation by others (Somerville et al., 2010). Interestingly, the perception of the evaluation of self by familiar others activates the dMPFC (Korn et al., 2012), which is thought to be the neural correlate of the interpersonal self, rather than vMPFC. This may be explained by the fact that this experimental manipulation affects mental representations of the self in relation to specific others rather than those related to the value of the self, illustrating the conceptual difference between the interpersonal self and the social-value of the self.

Like the physical self, the social self dissociates self and other, but it does so in a different way. The social self encompasses any people or objects that are relevant or behaviorally significant to the self, which are considered "other" in terms of the physical self. The vMPFC and PCC are activated during name and face recognition of oneself and friends relative to recognition of unfamiliar people (Sugiura et al., 2008; Tacikowski et al., 2012), and the vMPFC is correlated with the amount of self-referential thought (D'Argembeau et al., 2005). On the other hand, activation of these regions is often absent when self-trait judgment is compared with trait judgment about familiar people, such as friends and relatives (Schmitz et al., 2004; Benoit et al., 2010). Self-face recognition involves activation of the vMPFC when the number of other faces in the other trials in the task sequence is increased, probably due to the self-value processing induced by social comparison (Sugiura et al., 2012). In a similar task design using young female subjects, activation of the PCC for self-face was enhanced when the female faces in other trials were less attractive, particularly when the subject's self-esteem was high (Oikawa et al., 2012).

Again, these cortical regions are unlikely to be utilized exclusively for the processing of social-value. The vMPFC and the medial orbitofrontal cortex (mOFC), which is sometimes regarded as identical with or adjacently distinct from the vMPFC, are known to represent the value of objects in general and to play critical roles in value-based decision making (Rangel et al., 2008; Rushworth et al., 2011). This general region comprises a reward system that operates in conjunction with other deep structures, such as the



striatum and the midbrain dopamine system, which are sometimes activated during self-trait judgment (Kircher et al., 2002; Moran et al., 2006; Benoit et al., 2010) or perception of self-evaluation by others (Izuma et al., 2008; Korn et al., 2012). The relationship between the reward system and self-related processes is a matter of recent discussion (Northoff and Hayes, 2011). The PCC and the adjacent precuneus are involved in a wide range of highly integrated processes, such as visuospatial imagery, episodic-memory retrieval, and self-referential processes (Wagner et al., 2005; Cavanna and Trimble, 2006). This set of midline cortical regions is also considered to be a major component of the default mode network that is active during a conscious resting state and deactivated during the execution of attention-demanding tasks (Gusnard et al., 2001; Raichle et al., 2001).

OTHER ASPECTS OF SELF

One may consider memory, especially autobiographical memory, as a critical factor of self. Numerous functional imaging studies have investigated the neural activity specifically observed during the retrieval of an autobiographical memory. The medial prefrontal and parietal cortices, including the cingulate cortex and the lateral temporal and parietal regions with some regions lateralized to the right (Svoboda et al., 2006; Buckner and Carroll, 2007; Spreng et al., 2009) exhibit activation during such a task. These areas overlap with the cortical regions proposed to be the neural underpinnings of the three categories of self. This suggests that autobiographical memory is not merely a single essential factor but rather the “all-star” of self-related cognitive processes.

Some researchers assume perspective-taking to be a key concept in the distinction of self from other. However, the findings of neuroimaging studies investigating this issue may also be explained by the framework of the three categories of self, particularly the physical and interpersonal selves. Many neuroimaging studies compare first-person (1P) and third-person (3P) perspectives to address this issue. Cortical activation is typically more prominent in a 3P rather than a 1P perspective, but activation of associated brain regions varies widely across studies. These findings were somewhat clarified when perspective-taking was divided into visuospatial and mental perspectives, and activation was assumed to reflect the increased cognitive load related to the non-canonical nature of the 3P perspective. Greater activation for 3P visuospatial

perspective-taking relative to 1P visuospatial perspective-taking is typically reported in the visual association and premotor cortices (Vogeley et al., 2004; David et al., 2006, 2008a), which overlap with the neural correlates of the physical self. This finding may be explained by the cognitive load involved in the imaginary physical-location change (i.e., moving self-body) required to obtain a non-canonical 3P viewpoint. Regarding mental perspective-taking, a greater activation for the 3P relative to the 1P perspective is frequently reported in the pSTS/TPJ and dMPFC (Ruby and Decety, 2001, 2003, 2004; David et al., 2008a; Schnell et al., 2011; Ramsey et al., 2013), which overlap with the neural correlates of the interpersonal self. This overlap appears to be reasonable because taking the mental perspective of others (intention, emotion, belief) is synonymous with ToM. On the other hand, many of these studies have reported a greater activation in the 1P compared with the 3P perspective in the MPFC and PCC, which are the proposed neural correlates of the social self. It is often difficult to conclusively attribute this finding to self-cognition, since it is usually explained by either behavioral significance (i.e., the social self) or differential default mode activity (Gusnard et al., 2001; Raichle et al., 2001) due to differences in task difficulty (McKiernan et al., 2003).

HYPOTHESIS: THREE LAYERS OF INTERNAL SCHEMA

In the preceding section, the self was divided into three individual categories that differ according to the related supporting cognitive processes and neural substrates. The current section, however, proposes a common characteristic or computational architecture that underlies the processes of these three categories. In short, the common characteristic is a forward prediction model, which is a rather common and classical conceptualization in models of the physical self. It has been assumed that the physical self is the product of an associative learning process based on the repeated experiences of bodily motion and sensory-feedback.

Here, a novel attempt will be made to adapt the forward prediction model to the interpersonal and social selves with the intention of explaining all categories of self within the framework of associative learning. A critical component of this adaptation is the internal schema that denotes the association between the neural-representation of the output plan and the feedback input (Figures 4A,B); this schema is assumed to exist for each target of

the output and is modified depending on context. In this view, the self may be defined as a label for the capability of forward prediction (**Figure 4B**) in any system that has such characteristics. Neuroimaging findings appear to be explained by top-down and bottom-up attention to the schema that is typically driven by task requirements and prediction error, respectively.

Conceptually, the schema is the basis for all cognitive operations, including perception and behavioral control. The schema is used as an inverse model to plan output (**Figure 4C**) or even to represent each element of the external environment and may be used to simulate a schema of the mind of another; that is, to infer the internal process of others based on the observed output of that person. Given the diverse utility of the schema, the self is only a phenomenon that is occasionally experienced during its functioning, while reflecting the very basic characteristics of the schema. Additionally, a hierarchical layer structure of the three categories of self and the dynamics across the layers are also important features of the proposed model. The hierarchical structure stems from the developmental relationship between the three schemata and serves as the basis of cross-layer interaction, which may be critical for the integrity of the three self-concepts.

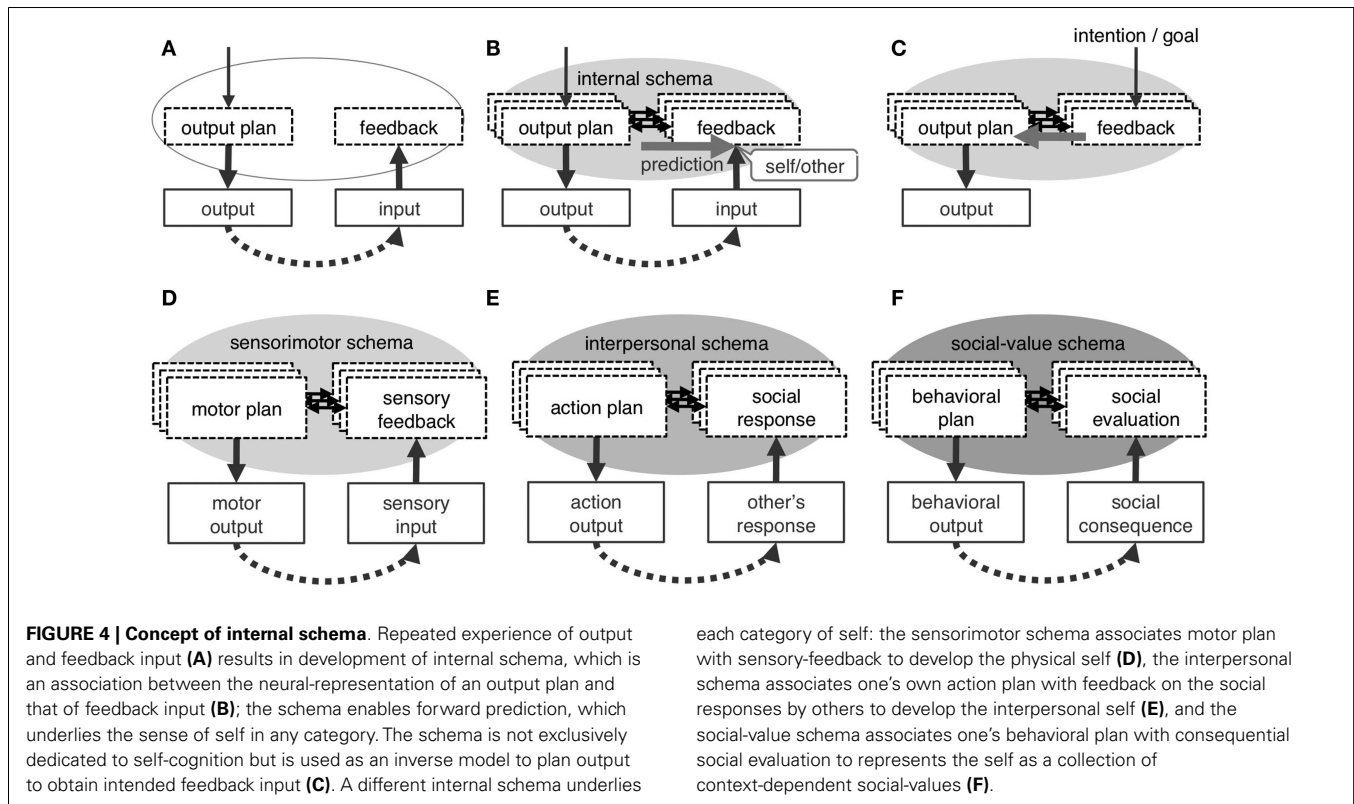
PHYSICAL SELF AND SENSORIMOTOR SCHEMA

The concept of the forward model was first applied to explain the sense of self-agency in action. The sense of self-agency, or the self-attribution of action, is widely assumed to have been derived from the consistency between the sensory input that results from action and the prediction emerging from the action intention or collateral output from the motor system (**Figure 4B**) (Wegner and Wheatley, 1999; Sato and Yasuda, 2005; David et al., 2008b). One is convinced that the observed action of one's own hand is actually performed by oneself because the action is somatically experienced (i.e., somatosensory perception) and looks (i.e., visual perception) as predicted. Frith and colleagues (Frith et al., 2000; Frith, 2005) incorporated this conceptualization into a detailed cognitive model by extending the model of feed-forward motor control (Wolpert et al., 1995) to explain the impairment in the sense of self-agency, or the delusion of control, which is a characteristic symptom of schizophrenia. In this model, the prediction of sensory input as a consequence of action is based on intended motor commands and cancels actual input. Subjectively, a successful cancellation is experienced as one's unawareness of the sensory consequences of one's own actions and is exemplified as the attenuated sensation of self-generated tickling (Weiskrantz et al., 1971; Blakemore et al., 1999). In functional neuroimaging, this cancellation is detected as an attenuation of the activation related to sensory processing when sensory input is caused by self-generated action rather than being externally produced (Blakemore et al., 1998). More frequently, in fact, this functioning is captured as an increase in activation during the violation of action self-agency due to experimental manipulation (McGuire et al., 1996a; Fink et al., 1999; Hashimoto and Sakai, 2003; Fu et al., 2006; David et al., 2007; Corradi-Dell'Acqua et al., 2008; Farrer et al., 2008; Spengler et al., 2009; Yomogida et al., 2010).

Here, the concept of a sensorimotor schema, or the learned association between one's motor plan and the feedback sensory input (**Figure 4D**), is introduced. The sensorimotor schema exists

for each effector or movement coordinated by multiple muscles and is adaptively modified depending on physical context, including posture and the external physical environment. The evidence that the schema is indeed constructed through associative learning has been experimentally provided; the repeated experience of an action and its effect on an object on a computer monitor later produce the sense of action self-agency for that virtual effector, and the related neural responses are similar to those seen in previous studies of action self-agency (Schnell et al., 2007; Spengler et al., 2009; Yomogida et al., 2010). The sensorimotor schema allows for the generalization of the forward model to different phenomena of the physical self, such as the sense of body-ownership and self-face recognition in non-contingent images (e.g., static images, prerecorded videos). It has been shown that body-ownership requires a pre-existing internal representation of the position of the limbs (Tsakiris and Haggard, 2005; Costantini and Haggard, 2007). Given that such a representation is constructed and continuously updated by matching the feed-forward prediction and re-afferent sensory input during active movement (Synofzik et al., 2006; Tsakiris et al., 2006), this representation is closely related to or identical with a sensorimotor schema. Self-face recognition ability in a non-contingent image is also likely to depend on a sensorimotor schema. Infants seem to consolidate the visual representation of one's own face into long-term memory during the experience of viewing a contingent self-face in the mirror; this idea is supported by the observation that self-face recognition first develops in a contingent and then in a non-contingent image (Bigelow, 1981). Therefore, the unique characteristic of the visual representation of the self-face seems to be realized by its association with the experience of action self-agency or body-ownership and, thus, with a sensorimotor schema.

The activation of sensory and motor association cortices related to the physical self is parsimoniously explained by attention to the sensorimotor schema. In this situation, it is advantageous to separate top-down from bottom-up attention. Top-down attention is induced by a variety of experimental manipulations in which the use of information in the schema is necessary or advantageous. Activation of sensory or motor association cortices is observed when perceived motion is explicitly required to be self-produced (Farrer and Frith, 2002) or when the task demands monitoring of one's own motor action control (Ogawa and Inui, 2007; Schnell et al., 2007). On the other hand, bottom-up attention is typically driven by prediction error, which may contribute to a recalibration of the schema. This line of interpretation most likely refers to neural activation in response to manipulated sensory-feedback, that is, a violation of self-agency in action (McGuire et al., 1996a; Blakemore et al., 1998; Fink et al., 1999; Farrer et al., 2003, 2008; Hashimoto and Sakai, 2003; Fu et al., 2006; David et al., 2007; Schnell et al., 2007; Corradi-Dell'Acqua et al., 2008; Spengler et al., 2009; Yomogida et al., 2010). Activation while experiencing a sense of body-ownership (Ehrsson et al., 2004, 2005; Tsakiris et al., 2007; Ionta et al., 2011) and bodily self-recognition in a non-contingent image (Uddin et al., 2005; Platek et al., 2006; Sugiura et al., 2006, 2008, 2012) may be attributed to either type of attention. The commonality of attention and consciousness may explain the activation of these regions in terms of top-down access to information in the schema. The activation is also attributable



to prediction error when perceptual input differs from what was expected: illusion-induced body-ownership may be imperfect, and some strange feelings may remain if the presented self-face picture is somewhat different from what one usually sees in the mirror while one remains sure that the face is one's own.

The sensorimotor schema explains not only self-cognition but also any cognitive operation related to one's physical interaction with the external environment. In fact, the forward model, as adapted to self-cognition, was originally developed for motor control (Wolpert et al., 1995), and the concept of sensorimotor schema was adopted from that model. The sensorimotor schema, or the association of the motor plan with sensory-feedback, may be used as an inverse model to calculate the motor plan to obtain the intended sensory-feedback (Figure 4C). This idea is consistent with the conceptual framework of ideomotor theory, which assumes a common coding of action and consequential perception (Prinz, 1997). Furthermore, the sensorimotor schema may play a critical role in an individual's mental representation of the physical environment. A person can have intention and a motor plan for interaction with many objects in the immediate external environment (e.g., gazing, reaching), and an essential property of the physical environment is this potential interaction, which may be represented in the sensorimotor schema. This notion is compatible with the fact that the cortical areas implicated in sensory or spatial attention overlap primarily with those supporting the sensorimotor schema (Downar et al., 2000; Corbetta and Shulman, 2002).

Furthermore, the sensorimotor schema seems to be exploited to simulate the schema of others; that is, it can be used to infer the intention or action goals of others. The schema may gain the

simulation ability by associating one's own motor output with the perceived contingent motor action of others in an interactive situation where the self and others share an intention or action goal. Such an interactive environment is common in the daily relationship between infants and their caretakers (Kaye and Fogel, 1980; Cohn and Tronick, 1988). This view is consonant with the hypothesis that the MNS is forged by sensorimotor association learning (Heyes, 2010). Further, this view implicates that the mirror neurons are a subcomponent of the simulation-capable sensorimotor schema that associates one's own motor actions not only with the same action but also with different but related actions of others. This is supported by the fact that, in the cortical areas reported to accommodate mirror neurons in primates, there are a greater number of "counter-mirror neurons," which code other's actions that are different from, but related to, one's own actions (Gallese et al., 1996). Also, in a human neuroimaging study, activation of such regions was greater during observation of other's actions that were complimentary (i.e., in joint action) to one's own actions than during observation of imitative actions (Newman-Norlund et al., 2007). Additionally, activation during the observation of another's actions is not limited to the classic human MNS (i.e., inferior parietal and frontal cortices) but has also been identified in multiple visual and motor association cortices (Caspers et al., 2010).

INTERPERSONAL SELF AND INTERPERSONAL SCHEMA

The existence of forward prediction during an individual's social interaction may be empirically or intuitively plausible. An implicit expectation about the range of possible responses usually arises

in situations in which one individual greets another. This is why people become surprised at an unexpected response or the lack of a response from the other person. The range of expected responses greatly differs depending on the identity of the responder (i.e., degree of familiarity and various demographic factors such as age, gender, cultural background, and situational context). The range of expected responses is likely to be updated after repeated experiences of prediction error with a specific familiar person or a specific type of unfamiliar persons.

In this context, it appears reasonable to assume the existence of an interpersonal schema that represents a link between one's social action (i.e., output plan) and the expected responses (i.e., feedback) of others (**Figure 4E**). The schema is constructed following repeated exploratory social interactions in daily life; that is, through associative learning involving one's own social actions toward a person and the feedback (**Figures 4A,B**). The schema exists for each familiar person or for a specific type of people and is adaptively modified depending on social contextual cues such as time, place, and occasion. These characteristics are comparable to those of the sensorimotor schema in terms of the way the schema develops, that the schema exists for each target of output, and that it is modified depending on context.

Several neuroimaging findings support the conceptualization of the interpersonal schema as the basis of the interpersonal self. Activation related to the interpersonal self appears to be explained by either top-down or bottom-up attention related to the interpersonal schema in a way that is similar to the relationship between the physical self and the sensorimotor schema. It appears reasonable to regard awareness that another's attention or intention is directed at oneself as an example of top-down attention to the interpersonal schema. In other words, activation of several medial and lateral posterior cortices during the perception of self-directed eye-gaze (Calder et al., 2002; Kampe et al., 2003; Wicker et al., 2003; Pelphrey et al., 2004a; Schilbach et al., 2006; Steuwe et al., 2012), hearing one's own name being called (Kampe et al., 2003; Perrin et al., 2005; Tacikowski et al., 2011), or real-time interaction with others (Rilling et al., 2004; Jeong et al., 2011) may reflect one's top-down attention to the representation of the other's potential response to one's own social action. Bottom-up attention is also represented by neural activation in this region in response to prediction error or the perception of an unexpected social response as feedback to one's own action. During a simple two-player strategy game, when the subject believes that the opponent is responding based on the prediction of the subject's next action, the prediction error of the perceived opponent's action induces activation in these regions (Hampton et al., 2008).

Furthermore, the functioning of the interpersonal schema is not specific to self-cognition but also relates to any cognitive operation associated with interpersonal interaction. The schema may be used as an inverse model to calculate the behavioral plan of social action toward another person to obtain an intended social response (**Figure 4C**). Accordingly, the neural correlates of the schema are more activated during speech production toward a virtual agent than during an overt description of the situation (Sassa et al., 2007). An individual can have an intention and a plan of social interaction (or of no interaction) in relation to many people in the immediate social environment. An essential property of

the immediate social environment is this potential, which is represented in the interpersonal schema. A similar notion, referred to as "social attention," is thought to be supported by the same cortical network (Nummenmaa and Calder, 2009). Furthermore, the simulation capacity of the interpersonal schema, or the inference of the intention and plan of another's social action based on perceived action, may partially overlap with ToM and may explain the overlap of their neural correlates (Gallagher and Frith, 2003; Frith and Frith, 2006; Spreng et al., 2009). However, it is important to note that ToM addresses both the social and non-social beliefs and intentions of others. According to a theory of the role of this network in the development of event schemata in general, the MPFC is assumed to support an abstract dynamic summary representation in the form of event simulators, and its interaction with posterior cortical areas is assumed to comprise knowledge of social events (Krueger et al., 2009). Activation of the implicated cortical regions has been reported in studies evaluating the detection of prediction-violating behavior or objects in the absence of self-involvement or social context (Grezes et al., 2004; Pelphrey et al., 2004b; Wakusawa et al., 2009). Thus, the mature interpersonal schema functions independent of self-cognition and comprises one aspect of a cognitive system supporting higher social and non-social processes. Nevertheless, working from the perspective that the evolution of intelligence in primates has been driven by social demand (Humphrey, 1976; Byrne and Whiten, 1988), it is tempting to assume that the initial interpersonal schema is the origin of the entire system.

SOCIAL-VALUE SCHEMA

Forward prediction is also plausible during the evaluation of one's own social-value. People are surprised when they receive an extremely high or low evaluation for a certain behavior; that is, an individual is relatively unaware of having obtained an evaluation when the evaluation is within the predicted range. It is assumed that humans have multiple self-concepts and that self-value is dependent on social role (e.g., parent, friend, worker) (Stryker and Statham, 1985; Markus and Cross, 1990; Roberts and Donahue, 1994). Thus, it appears reasonable to assume a specific range of the predicted evaluation for each contextual role, and this is updated through the feedback of prediction error.

It is assumed herein that the social-value schema represents a link between one's social behavior (i.e., output plan) and the predicted evaluation of this behavior (i.e., feedback) (**Figure 4F**). This schema is constructed for each contextual role through repeated experiences with social evaluations, which result in the learning of associations between one's own social behaviors and the evaluative feedback they elicit (**Figures 4A,B**). Again, these characteristics are comparable to those of the sensorimotor or interpersonal schemas in terms of the way the schema develops, that the schema exists for each target of output, and that it is modified depending on context. This idea is congruent with the known general roles of the neural correlates of this schema: the vMPFC (as well as the ACC and mOFC) represents values (Rangel et al., 2008; Rushworth et al., 2011), and the PCC (and precuneus) processes the different aspects of social or autobiographical contexts (Addis et al., 2004; Gilboa et al., 2004; Chiao et al., 2009) and different types of perspectives (Vollm et al., 2006; Mano et al., 2009).

Neuroimaging findings relevant to the social-value of the self are likely explained by either top-down or bottom-up attention related to the social-value schema. The activation of the vMPFC and PCC during self-trait judgment (Craig et al., 1999; Kelley et al., 2002; Schmitz et al., 2004; Heatherton et al., 2006; D'Argembeau et al., 2007), perception of self-descriptive trait adjectives (Kircher et al., 2002; Macrae et al., 2004; Moran et al., 2006), and perception of self-evaluation by others (Izuma et al., 2008; Somerville et al., 2010) may reflect top-down attention to the social-value schema. Activation of these regions in terms of bottom-up attention in response to unexpected evaluations of one's behavior was found in a study using monetary rewards for a simple estimation game involving a pair of players. These regions exhibited greater activation when the payment to the two players was unequal for the same correct performance (i.e., prediction error in evaluation) than when it was equal (Fliessbach et al., 2007).

The functioning of the social-value schema is also not specific to self-evaluation but operates for any cognitive operation related to social-value. The schema may be used as an inverse model to calculate the behavioral plan for obtaining an intended social evaluation (Figure 4C). The activation of the vMPFC and PCC during moral judgment is greater when the situation is more realistic (i.e., relevant to the real-life evaluation of self), such as when the decision is situation-based rather than rule-based (Robertson et al., 2007) or when the potential victim of the decision is humanized by mentalizing manipulation (Majdandzic et al., 2012). In daily life, we are intermittently engaged in such behavioral planning on the basis of the social-value of the self, while it is interrupted during execution of a specific attention-demanding task. This appears to be a plausible explanation for the activation of these areas during the conscious resting state (Gusnard et al., 2001; Raichle et al., 2001). Furthermore, the simulation capacity of the social-value schema (i.e., the making of inferences regarding the intentions and plans related to another's social behavior) may explain the activation of these regions during social-value judgments about others (Craig et al., 1999; Schmitz et al., 2004; Sugiura et al., 2004; Ochsner et al., 2005; Benoit et al., 2010). Given the general role of the vMPFC and PCC in value-based decision making (Rangel et al., 2008; Rushworth et al., 2011) and their specific roles in highly integrated visuospatial and memory retrieval processes (Wagner et al., 2005; Cavanna and Trimble, 2006), respectively, it seems fair to consider the social-value schema as only a subcomponent of the functioning of this neural system.

MULTI-LAYER STRUCTURE AND CROSS-LAYER DYNAMICS

It is further proposed that the three categories of self, or internal schemata, comprise a hierarchical layered structure in that the maturation of one layer, or schema, serves as the basis for the development of the next layer. Additionally, the prediction error generated in one layer may result in an updating of the schema not only in that layer but also in adjacent layers. These cross-layer dynamics may be, in part, responsible for both the integrity of the categories and the ambiguity across the three self-concepts.

The self-layers are assumed to develop in the following order: sensorimotor, interpersonal, and social-value. The development of a higher layer is dependent on the maturation of the internal schema in a lower layer; here, the maturation of the schema

denotes the acquisition of the potential to simulate the schema of others.

In terms of the sensorimotor schema, the acquisition of the potential to infer the intention or action goal of others corresponds to an infant's discovery of an "other" or an agent who has a similar mental mechanism to the self. This discovery of an other is the very basis of the development of the interpersonal schema that requires the execution of social action toward the other and the understanding of the other's social response. In fact, a similar concept to this simulation potential has been conceptualized in a recent hierarchical self-model as the Bodily Social Self (BSS), which links Phenomenal Self and Narrative Self (Farmer and Tsakiris, 2012); the former seems to correspond to self-body-dedicated (premature) physical self, and the latter to interpersonal self and social-value of self, together, in the proposed model. This simulation capacity, or the BSS, is included in the physical self in this neural-representation model because both types of self are accommodated by the sensorimotor schema. The maturation process, the acquisition of a simulation capacity by the sensorimotor schema, is assumed to develop in the first 6 months of life in infants (Kaye and Fogel, 1980; Cohn and Tronick, 1988), and its failure to develop has been proposed as responsible for the impaired development of sociality in autism (Gergely, 2001).

The next step of development is triggered by the maturation of the interpersonal schema. The acquisition of a simulation capacity, ToM or mentalizing ability, by the interpersonal schema enables one to conceive of the representation of self in another's mind. The collection of such self-representations in many others' minds allows abstraction of the value of self to construct the social-value schema. This internalization process is taken for granted in developmental theories of social self-concepts, with the process, presumably, peaking in adolescence (Cooley, 1902; Mead, 1934; Harter, 1985). Accordingly, self-dominant activation during judgment about significant social attributes is observed in the vMPFC in adults and in the dMPFC in adolescents (Pfeifer et al., 2007). This probably reflects the ongoing self-value abstraction process in the interpersonal schema.

The cross-layer functioning of the error-based updating of the schema adds tremendous complexity to one's self-related experiences as well as to the interpretation of neuroimaging findings. For example, prediction error in the sensory-feedback in response to an action by a subject may produce a sense or belief that the action is performed by another person rather than a feeling of strangeness in one's own action; that is, error-based updates do not influence the sensorimotor schema but the interpersonal schema. Indeed, the experimental manipulation of sensory-feedback during the moving of a hand by a subject or during the manipulation of an agent on a computer monitor activates the TPJ/pSTS and dMPFC (McGuire et al., 1996a; Farrer and Frith, 2002; Farrer et al., 2003, 2008; Hashimoto and Sakai, 2003; Leube et al., 2003; Fu et al., 2006; Spengler et al., 2009; Yomogida et al., 2010), which are implicated in the interpersonal schema. An abnormal functioning of this cross-layer error-attribution (i.e., attribution to other) is considered to explain several symptoms of schizophrenia, including the delusion of control (Frith et al., 2000; Frith, 2005).

Another example is the case in which prediction error in the self-value layer influences the interpersonal schema. Unexpectedly

high or low evaluation of the self by another is assumed to cause change in self-value but, alternatively, may be attributed to an idiosyncratic viewpoint or attitude of the evaluator (interpersonal schema). One example may be a case in which one thinks “his recognizing me as stupid is not because I am stupid, but because he is stupid.” In the experimental setting, feedback evaluation toward the self is typically provided by a small number of alleged evaluators, and it is highly likely that cross-layer attribution does occur. This consideration is congruent with data identifying neural responses that are positively correlated with self-value prediction error (i.e., discrepancy between a subject’s own evaluation of the self and the evaluation by others) in the major components of the interpersonal schema (Korn et al., 2012).

The implementation of hierarchical layer structure and cross-layers error-attribution, in addition to the association-based generation of the internal schema and its prediction-error-based update *per se*, makes this model conform to the Hierarchical Bayesian model based on the free-energy principle (Friston, 2010). This conformity may suggest a future potential sophistication of this model in the Bayesian framework. The proposed model may, therefore, provide an example of successful adoption of this comprehensive framework to cognitive processes that cover perception to higher-level cognition accompanying empirical data.

Several cortical areas may play a unique role in the coordination of functioning across multiple layers. Specifically, the right lateral prefrontal cortex may have a role in resolving conflicts in different layers. This region is activated during sensory-feedback manipulation when it obviously conflicts with motor control (Fink et al., 1999) or when agency-attribution judgment (i.e., self or other) is required (David et al., 2007; Schnell et al., 2007; Farrer et al., 2008). It has been proposed that this region is responsible for an impaired belief-validation process during the mirrored-self misidentification (mirror sign) due to a failure to resolve the conflict between self-face recognition and contingency detection when either process is abnormal (Coltheart, 2007, 2010). Apparently consistent with this view, activation in the right lateral prefrontal cortex is frequently reported during the recognition of self-face or self-body in non-contingent images (Platek et al., 2004, 2006; Sugiura et al., 2005a, 2006, 2008, 2012; Uddin et al., 2005; Devue et al., 2007; Kaplan et al., 2008). Moreover, this region responds to behavior that violates social norms (i.e., error in the interpersonal layer) (Wakusawa et al., 2009) or when there are discrepancies between a subject’s self-evaluation and the evaluation by others (i.e., error in the social-value layer) (Korn et al., 2012).

In summary, the concept of an internal schema with three layers operating under the assumption of cross-layer dynamics provides a relatively simple integrated conceptual framework for the self-concept. In this framework, associative learning and the hierarchical structure of the cortical network appear sufficient to explain the wide range of behavioral, developmental, and neuroimaging findings related to self-cognition.

ROLE OF MIDLINE STRUCTURES

In the proposed model, a majority of the association cortices in both lateral and medial structures critically contribute to some aspect of self. This view may be inconsistent with the notion of a special role for the midline regions in self-cognition, which might

be implied by this topic. Based on the size of areas included, however, it is possible to characterize the contribution of the midline regions in the following manner: they are most relevant to the social-value layer, less relevant to the interpersonal layer, and least relevant to the sensorimotor layer. This characterization is largely congruent with the view that has previously been discussed (Uddin et al., 2007).

Within the proposed conceptual framework, the primary interest regarding midline structures concerns the functional demarcation of the neural correlates of each schema. The border of the cortical correlates for each self-layer parallels the functional segregation of the midline structures.

In the frontal lobe, the border between the cortical correlates of the sensorimotor and interpersonal schema may be reasonably defined as the border between the premotor cortex (Brodmann area 6) and the posterior part of the prefrontal cortex (Brodmann area 8) given the motor-associated and amodal nature of these schemata, respectively. A review of neuroimaging findings evaluating the dorsal part of the medial frontal lobe suggests that the border is a few centimeters rostral to the vertical plane crossing through the anterior commissure (AC). Clusters of activation peaks related to attention to action or to sensation are located posteriorly, and those related to concepts are located anteriorly (Seitz et al., 2006). The location of this functional border appears largely congruent with the cytoarchitectonic border between area 6 and area 8 (Geyer, 2004). The border between the regions for the interpersonal and social-value schema, on the other hand, has been defined only functionally. Previous reviews have consistently demonstrated a functional inhomogeneity of the MPFC as the dorsal and ventral regions tend to be involved in cognitive and emotional processes, respectively. However, the proposed level of a horizontal plane for that border has varied across studies (i.e., from running through the AC to 20 mm above the AC) (Amodio and Frith, 2006; Van Overwalle, 2009).

In the medial parietal lobe, the extent of the cortical correlates of the social-value schema remains inconclusive but may encompass the entire precuneus and PCC. This region contains multiple functional subareas that are specialized for the processing of particular components of the implicated roles in this region, such as episodic-memory retrieval, visual imagery, and the representation of a personally familiar place (Sugiura et al., 2005b; Wagner et al., 2005; Cavanna and Trimble, 2006; Summerfield et al., 2009; Zhang and Li, 2012). Although all of these processes appear to be relevant in some way to the processing of contextual roles, a detailed account of the association of these processes with respect to the concept of the social-value schema is extremely premature.

CONCLUSION

The framework proposed herein is an attempt to rescue the integrated construct of self from the pessimistic view arguing against the existence of self-specific neural system. The concepts of self appear to be parsimoniously arranged into three categories according to the contexts of awareness and development as well as the implicated cortical regions. According to the proposed model, the internal schema, which represents the learned associations between behavioral output and feedback input, enables the system to engage in forward prediction and explains the sense of

self in all three categories. Importantly, the internal schema is not exclusively dedicated to self-cognition but is the very basis of the cognitive system underpinning interaction with a physical or social environment. Additionally, the schemata for these three categories of self comprise a hierarchical layer structure in terms of their developmental and updating processes.

The sensorimotor schema, namely, the association of a motor plan with feedback sensory input acquired through exploratory motor activity, is supported by sensory and motor association cortices and results in a sense of a physical self representing self-agency of action, body-ownership, and bodily self-recognition. As the schema matures, it becomes capable of simulating the intention or action goals of others. This allows one to explore and experience social interaction in which the interpersonal schema, or the association of one's own social action with subsequent social responses that serve as feedback, is developed in the amodal association cortices of the dMPFC and lateral posterior cortices (e.g., pSTS/TPJ and ATC). While allowing for the experience of the interpersonal self, which is the awareness of self-directed attention or the intention of others, the interpersonal schema also matures to accommodate the representation of the self in another's mind. The collection of such self-representations in many others' minds enables the development of the social-value schema, which evaluates one's social

behavior and feedback evaluation. This schema, supported by the vMPFC and PCC, enables the operation of the social self and represents the self as a collection of context-dependent social-values.

The model proposed herein explains the large variety of activated regions that have been reported by studies addressing self-related cognitive processes as well as their involvement in non-self-related processes. It also provides a unique perspective on the relationship between self-cognition and the cognitive system involved in one's interaction with the physical or social environment. In particular, the assumed layer structure provides for the development, complexity, and integrity of three categories of the self. This view understands the different, but not mutually exclusive, roles of the midline and lateral cortical regions in self-cognition in terms of the different medial-lateral distribution of the three internal schemas. With respect to midline structures, due to the different sizes of the areas that each internal schema occupies, the characteristics of the self may be ranked as follows in term of prominence: social-value self, interpersonal self, and physical self.

ACKNOWLEDGMENTS

The preparation of this manuscript was supported in part by KAKENHI (25560347) from MEXT.

REFERENCES

- Addis, D. R., McIntosh, A. R., Moscovitch, M., Crawley, A. P., and McAndrews, M. P. (2004). Characterizing spatial and temporal features of autobiographical memory retrieval networks: a partial least squares approach. *Neuroimage* 23, 1460–1471. doi:10.1016/j.neuroimage.2004.08.007
- Amodio, D. M., and Frith, C. D. (2006). Meeting of minds: the medial frontal cortex and social cognition. *Nat. Rev. Neurosci.* 7, 268–277. doi:10.1038/Nrn1884
- Bandura, A. (1982). Self-efficacy mechanism in human agency. *Am. Psychol.* 37, 122–147. doi:10.1037/0003-066x.37.2.122
- Benoit, R. G., Gilbert, S. J., Volle, E., and Burgess, P. W. (2010). When I think about me and simulate you: medial rostral prefrontal cortex and self-referential processes. *Neuroimage* 50, 1340–1349. doi:10.1016/j.neuroimage.2009.12.091
- Bigelow, A. E. (1981). The correspondence between self and image movement as a cue to self-recognition for young children. *J. Genet. Psychol.* 139, 11–26. doi:10.1080/00221325.1981.10533432
- Blakemore, S. J., Frith, C. D., and Wolpert, D. M. (1999). Spatio-temporal prediction modulates the perception of self-produced stimuli. *J. Cogn. Neurosci.* 11, 551–559. doi:10.1162/089892999563607
- Blakemore, S. J., Wolpert, D. M., and Frith, C. D. (1998). Central cancellation of self-produced tickle sensation. *Nat. Neurosci.* 1, 635–640. doi:10.1038/2870
- Buckner, R. L., and Carroll, D. C. (2007). Self-projection and the brain. *Trends Cogn. Sci.* 11, 49–57. doi:10.1016/j.tics.2006.11.004
- Byrne, R. W., and Whiten, A. (eds) (1988). *Machiavellian Intelligence: Social Expertise and the Evolution of Intellect in Monkeys, Apes and Humans*. Oxford: Clarendon.
- Calder, A. J., Lawrence, A. D., Keane, J., Scott, S. K., Owen, A. M., Christofels, I., et al. (2002). Reading the mind from eye gaze. *Neuropsychologia* 40, 1129–1138. doi:10.1016/S0028-3932(02)00008-8
- Caspers, S., Zilles, K., Laird, A. R., and Eickhoff, S. B. (2010). ALE meta-analysis of action observation and imitation in the human brain. *Neuroimage* 50, 1148–1167. doi:10.1016/j.neuroimage.2009.12.112
- Cavanna, A. E., and Trimble, M. R. (2006). The precuneus: a review of its functional anatomy and behavioural correlates. *Brain* 129, 564–583. doi:10.1093/Brain/Awl004
- Chiao, J. Y., Harada, T., Komeda, H., Li, Z., Mano, Y., Saito, D., et al. (2009). Neural basis of individualistic and collectivistic views of self. *Hum. Brain Mapp.* 30, 2813–2820. doi:10.1002/Hbm.20707
- Cohn, J. F., and Tronick, E. Z. (1988). Mother infant face-to-face interaction – influence is bidirectional and unrelated to periodic cycles in either partners behavior. *Dev. Psychol.* 24, 386–392. doi:10.1037/0012-1649.24.3.386
- Coltheart, M. (2007). Cognitive neuropsychiatry and delusional belief. *Q. J. Exp. Psychol.* 60, 1041–1062. doi:10.1080/17470210701338071
- Coltheart, M. (2010). The neuropsychology of delusions. *Ann. N. Y. Acad. Sci.* 1191, 16–26. doi:10.1111/j.1749-6632.2010.05496.x
- Cooley, C. H. (1902). *Human Nature and the Social Order*. New York: Charles Scribner's Sons.
- Corbetta, M., and Shulman, G. L. (2002). Control of goal-directed and stimulus-driven attention in the brain. *Nat. Rev. Neurosci.* 3, 201–215. doi:10.1038/Nrn755
- Corradi-Dell'Acqua, C., Ueno, K., Ogawa, A., Cheng, K., Rumiati, R. I., and Iriki, A. (2008). Effects of shifting perspective of the self: an fMRI study. *Neuroimage* 40, 1902–1911. doi:10.1016/j.neuroimage.2007.12.062
- Costantini, M., and Haggard, P. (2007). The rubber hand illusion: sensitivity and reference frame for body ownership. *Conscious. Cogn.* 16, 229–240. doi:10.1016/j.concog.2007.01.001
- Craig, A. D. (2002). How do you feel? Interoception: the sense of the physiological condition of the body. *Nat. Rev. Neurosci.* 3, 655–666. doi:10.1038/Nrn894
- Craig, A. D. (2009). How do you feel – now? The anterior insula and human awareness. *Nat. Rev. Neurosci.* 10, 59–70. doi:10.1038/Nrn2555
- Craik, F. I. M., Moroz, T. M., Moscovitch, M., Stuss, D. T., Winocur, G., Tulving, E., et al. (1999). In search of the self: a positron emission tomography study. *Psychol. Sci.* 10, 26–34. doi:10.1111/1467-9280.00102
- D'Argembeau, A., Collette, F., Van der Linden, M., Laureys, S., Del Fiore, G., Degueldre, C., et al. (2005). Self-referential reflective activity and its relationship with rest: a PET study. *Neuroimage* 25, 616–624. doi:10.1016/j.neuroimage.2004.11.048
- D'Argembeau, A., Ruby, P., Collette, F., Degueldre, C., Baetens, E., Luxen, A., et al. (2007). Distinct regions of the medial prefrontal cortex are associated with self-referential processing and perspective taking. *J. Cogn. Neurosci.* 19, 935–944. doi:10.1162/jocn.2007.19.6.935
- David, N., Aumann, C., Santos, N. S., Bewernick, B. H., Eickhoff, S. B., Newen, A., et al. (2008a). Differential involvement of the posterior temporal cortex in mentalizing but not perspective taking. *Soc. Cogn. Affect. Neurosci.* 3, 279–289. doi:10.1093/Scan/Nsn023

- David, N., Newen, A., and Vogele, K. (2008b). The “sense of agency” and its underlying cognitive and neural mechanisms. *Conscious. Cogn.* 17, 523–534. doi:10.1016/J.Concog.2008.03.004
- David, N., Bewernick, B. H., Cohen, M. X., Newen, A., Lux, S., Fink, G. R., et al. (2006). Neural representations of self versus other: visual-spatial perspective taking and agency in a virtual ball-tossing game. *J. Cogn. Neurosci.* 18, 898–910. doi:10.1162/Jocn.2006.18.6.898
- David, N., Cohen, M. X., Newen, A., Bewernick, B. H., Shah, N. J., Fink, G. R., et al. (2007). The extrastriate cortex distinguishes between the consequences of one's own and others' behavior. *Neuroimage* 36, 1004–1014. doi:10.1016/J.Neuroimage.2007.03.030
- Devue, C., Collette, F., Baletau, E., Dequeldre, C., Luxen, A., Maquet, P., et al. (2007). Here I am: the cortical correlates of visual self-recognition. *Brain Res.* 1143, 169–182. doi:10.1016/J.Brainres.2007.01.055
- Downar, J., Crawley, A. P., Mikulis, D. J., and Davis, K. D. (2000). A multimodal cortical network for the detection of changes in the sensory environment. *Nat. Neurosci.* 3, 277–283. doi:10.1038/72991
- Ehrsson, H. H., Holmes, N. P., and Passingham, R. E. (2005). Touching a rubber hand: feeling of body ownership is associated with activity in multisensory brain areas. *J. Neurosci.* 25, 10564–10573. doi:10.1523/JNeurosci.0800-05.2005
- Ehrsson, H. H., Spence, C., and Passingham, R. E. (2004). That's my hand! Activity in premotor cortex reflects feeling of ownership of a limb. *Science* 305, 875–877. doi:10.1126/Science.1097011
- Eisenberger, N. I., Lieberman, M. D., and Satpute, A. B. (2005). Personality from a controlled processing perspective: an fMRI study of neuroticism, extraversion, and self-consciousness. *Cogn. Affect. Behav. Neurosci.* 5, 169–181. doi:10.3758/Cabn.5.2.169
- Farmer, H., and Tsakiris, M. (2012). The bodily social self: a link between phenomenal and narrative selfhood. *Rev. Philos. Psychol.* 3, 125–144. doi:10.1007/s13164-012-0092-5
- Farrer, C., Franck, N., Georgieff, N., Frith, C. D., Decety, J., and Jeannerod, A. (2003). Modulating the experience of agency: a positron emission tomography study. *Neuroimage* 18, 324–333. doi:10.1016/S1053-8119(02)00041-1
- Farrer, C., Frey, S. H., Van Horn, J. D., Tunik, E., Turk, D., Inati, S., et al. (2008). The angular gyrus computes action awareness representations. *Cereb. Cortex* 18, 254–261. doi:10.1093/Cercor/Bhm050
- Farrer, C., and Frith, C. D. (2002). Experiencing oneself vs another person as being the cause of an action: the neural correlates of the experience of agency. *Neuroimage* 15, 596–603. doi:10.1006/Nimg.2001.1009
- Feinberg, T. E., and Keenan, J. P. (2005). Where in the brain is the self? *Conscious. Cogn.* 14, 661–678. doi:10.1016/J.Concog.2005.01.002
- Fenigstein, A., Scheier, M. F., and Buss, A. H. (1975). Public and private self-consciousness – assessment and theory. *J. Consult. Clin. Psychol.* 43, 522–527. doi:10.1037/H0076760
- Ferri, F., Frassinetti, F., Ardizzi, M., Costantini, F., and Gallese, V. (2012). A sensorimotor network for the bodily self. *J. Cogn. Neurosci.* 24, 1584–1595. doi:10.1162/jocn_a_00230
- Festinger, L. (1954). A theory of social comparison processes. *Hum. Relat.* 7, 117–140. doi:10.1177/001872675400700202
- Fink, G. R., Markowitsch, H. J., Reinke-meier, M., Bruckbauer, T., Kessler, J., and Heiss, W. D. (1996). Cerebral representation of one's own past: neural networks involved in autobiographical memory. *J. Neurosci.* 16, 4275–4282.
- Fink, G. R., Marshall, J. C., Halligan, P. W., Frith, C. D., Driver, J., Frackowiak, R. S. J., et al. (1999). The neural consequences of conflict between intention and the senses. *Brain* 122, 497–512. doi:10.1093/Brain/122.3.497
- Fliessbach, K., Weber, B., Trautner, P., Dohmen, T., Sunde, U., Elger, C. E., et al. (2007). Social comparison affects reward-related brain activity in the human ventral striatum. *Science* 318, 1305–1308. doi:10.1126/Science.1145876
- Friston, K. (2010). The free-energy principle: a unified brain theory? *Nat. Rev. Neurosci.* 11, 127–138. doi:10.1038/nrn2787
- Frith, C. (2005). The neural basis of hallucinations and delusions. *C. R. Biol.* 328, 169–175. doi:10.1016/J.Crvi.2004.10.012
- Frith, C. D., Blakemore, S. J., and Wolpert, D. M. (2000). Abnormalities in the awareness and control of action. *Philos. Trans. R. Soc. B Biol. Sci.* 355, 1771–1788. doi:10.1098/rstb.2000.0734
- Frith, C. D., and Frith, U. (2006). The neural basis of mentalizing. *Neuron* 50, 531–534. doi:10.1016/J.Neuron.2006.05.001
- Fu, C. H. Y., Vythelingum, G. N., Brammer, M. J., Williams, S. C. R., Amaro, E., Andrew, C. M., et al. (2006). An fMRI study of verbal self-monitoring: neural correlates of auditory verbal feedback. *Cereb. Cortex* 16, 969–977. doi:10.1093/Cercor/Bhj039
- Gallagher, H. L., and Frith, C. D. (2003). Functional imaging of ‘theory of mind’. *Trends Cogn. Sci.* 7, 77–83. doi:10.1016/S1364-6613(02)00025-6
- Gallagher, S. (2000). Philosophical conceptions of the self: implications for cognitive science. *Trends Cogn. Sci.* 4, 14–21. doi:10.1016/S1364-6613(99)01417-5
- Gallese, V., Fadiga, L., Fogassi, L., and Rizzolatti, G. (1996). Action recognition in the premotor cortex. *Brain* 119, 593–609. doi:10.1093/Brain/119.2.593
- Gallup, G. G. (1982). Self-awareness and the emergence of mind in primates. *Am. J. Primatol.* 2, 237–248. doi:10.1002/Ajp.1350020302
- George, N., Driver, J., and Dolan, R. J. (2001). Seen gaze-direction modulates fusiform activity and its coupling with other brain areas during face processing. *Neuroimage* 13, 1102–1112. doi:10.1006/Nimg.2001.0769
- Gergely, G. (2001). The obscure object of desire: ‘nearly, but clearly not, like me’: contingency preference in normal children versus children with autism. *Bull. Menninger Clin.* 65, 411–426. doi:10.1521/Bumc.65.3.411.19853
- Geyer, S. (2004). The microstructural border between the motor and the cognitive domain in the human cerebral cortex. *Adv. Anat. Embryol. Cell Biol.* 174, 1–89. doi:10.1007/978-3-642-18910-4_1
- Gilboa, A., Winocur, G., Grady, C. L., Hevenor, S. J., and Moscovitch, M. (2004). Remembering our past: functional neuroanatomy of recollection of recent and very remote personal events. *Cereb. Cortex* 14, 1214–1225. doi:10.1093/Cercor/Bhh082
- Gillihan, S. J., and Farah, M. J. (2005). Is self special? A critical review of evidence from experimental psychology and cognitive neuroscience. *Psychol. Bull.* 131, 76–97. doi:10.1037/0033-2909.131.1.76
- Grezes, J., Frith, C., and Passingham, R. E. (2004). Brain mechanisms for inferring deceit in the actions of others. *J. Neurosci.* 24, 5500–5505. doi:10.1523/Jneurosci.0219-04.2004
- Gusnard, D. A., Akbudak, E., Shulman, G. L., and Raichle, M. E. (2001). Medial prefrontal cortex and self-referential mental activity: relation to a default mode of brain function. *Proc. Natl. Acad. Sci. U.S.A.* 98, 4259–4264. doi:10.1073/Pnas.071043098
- Gutierrez, S. E., Kenrick, D. T., and Partch, J. J. (1999). Beauty, dominance, and the mating game: contrast effects in self-assessment reflect gender differences in mate selection. *Pers. Soc. Psychol. Bull.* 25, 1126–1134. doi:10.1177/01461672992512006
- Hampton, A. N., Bossaerts, P., and O'Doherty, J. P. (2008). Neural correlates of mentalizing-related computations during strategic interactions in humans. *Proc. Natl. Acad. Sci. U.S.A.* 105, 6741–6746. doi:10.1073/Pnas.0711099105
- Harter, S. (1985). *The Self-Perception Profile for Children*. Denver: University of Denver.
- Hashimoto, Y., and Sakai, K. L. (2003). Brain activations during conscious self-monitoring of speech production with delayed auditory feedback: an fMRI study. *Hum. Brain Mapp.* 20, 22–28. doi:10.1002/Hbm.10119
- Heatherton, T. F., Wyland, C. L., Macrae, C. N., Demos, K. E., Denny, B. T., and Kelley, W. M. (2006). Medial prefrontal activity differentiates self from close others. *Soc. Cogn. Affect. Neurosci.* 1, 18–25. doi:10.1093/Scan/Nsl001
- Heyes, C. (2010). Where do mirror neurons come from? *Neurosci. Biobehav. Rev.* 34, 575–583. doi:10.1016/J.Neubiorev.2009.11.007
- Higgins, E. T. (1987). Self-discrepancy: a theory relating self and affect. *Psychol. Rev.* 94, 319–340. doi:10.1037/0033-295x.94.3.319
- Humphrey, N. K. (1976). “The social function of intellect,” in *Growing Points in Ethology*, eds P. P. G. Bateson, and R. A. Hinde (Cambridge: Cambridge University Press), 303–317.
- Iacoboni, M. (2005). Neural mechanisms of imitation. *Curr. Opin. Neurobiol.* 15, 632–637. doi:10.1016/J.Conb.2005.10.010
- Ionta, S., Heydrich, L., Lenggenhager, B., Mouthon, M., Fornari, E., Chapuis, D., et al. (2011). Multi-sensory mechanisms in temporoparietal cortex support self-location and first-person perspective. *Neuron* 70, 363–374. doi:10.1016/J.Neuron.2011.03.009
- Izuma, K., Saito, D. N., and Sadato, N. (2008). Processing of social and monetary rewards in the human

- striatum. *Neuron* 58, 284–294. doi:10.1016/J.Neuron.2008.03.020
- James, W. (1890). *Principles of Psychology*. New York: Henry Holt.
- Jeong, H., Hashizume, H., Sugiura, M., Sassa, Y., Yokoyama, S., Shiozaki, S., et al. (2011). Testing second language oral proficiency in direct and semidirect settings: a social-cognitive neuroscience perspective. *Lang. Learn.* 61, 675–699. doi:10.1111/J.1467-9922.2011.00635.X
- Kampe, K. K. W., Frith, C. D., and Frith, U. (2003). “Hey John”: signals conveying communicative intention toward the self activate brain regions associated with “mentalizing,” regardless of modality. *J. Neurosci.* 23, 5258–5263.
- Kaplan, J. T., Aziz-Zadeh, L., Uddin, L. Q., and Iacoboni, M. (2008). The self across the senses: an fMRI study of self-face and self-voice recognition. *Soc. Cogn. Affect. Neurosci.* 3, 218–223. doi:10.1093/Scan/Nsn014
- Kawashima, R., Sugiura, M., Kato, T., Nakamura, A., Hatano, K., Ito, K., et al. (1999). The human amygdala plays an important role in gaze monitoring – a PET study. *Brain* 122, 779–783. doi:10.1093/Brain/122.4.779
- Kaye, K., and Fogel, A. (1980). The temporal structure of face-to-face communication between mothers and infants. *Dev. Psychol.* 16, 454–464. doi:10.1037//0012-1649.16.5.454
- Keenan, J. P., Wheeler, M. A., Gallup, G. G., and Pascual-Leone, A. (2000). Self-recognition and the right prefrontal cortex. *Trends Cogn. Sci.* 4, 338–344. doi:10.1016/S1364-6613(00)01521-7
- Kelley, W. M., Macrae, C. N., Wyland, C. L., Caglar, S., Inati, S., and Heatherton, T. F. (2002). Finding the self? An event-related fMRI study. *J. Cogn. Neurosci.* 14, 785–794. doi:10.1162/08989290260138672
- Kircher, T. T. J., Brammer, M., Bullmore, E., Simmons, A., Bartels, M., and David, A. S. (2002). The neural correlates of intentional and incidental self processing. *Neuropsychologia* 40, 683–692. doi:10.1016/S0028-3932(01)00138-5
- Kircher, T. T. J., Senior, C., Phillips, M. L., Benson, P. J., Bullmore, E. T., Brammer, M., et al. (2000). Towards a functional neuroanatomy of self processing: effects of faces and words. *Cogn. Brain Res.* 10, 133–144. doi:10.1016/S0926-6410(00)00036-7
- Korn, C. W., Prehn, K., Park, S. Q., Walter, H., and Heekeren, H. R. (2012). Positively biased processing of self-relevant social feedback. *J. Neurosci.* 32, 16832–16844. doi:10.1523/Jneurosci.3016-12.2012
- Krueger, F., Barbey, A. K., and Grafman, J. (2009). The medial prefrontal cortex mediates social event knowledge. *Trends Cogn. Sci.* 13, 103–109. doi:10.1016/J.Tics.2008.12.005
- Leube, D. T., Knoblich, G., Erb, M., Grodd, W., Bartels, M., and Kircher, T. T. J. (2003). The neural correlates of perceiving one’s own movements. *Neuroimage* 20, 2084–2090. doi:10.1016/J.Neuroimage.2003.07.033
- Macrae, C. N., Moran, J. M., Heatherton, T. F., Banfield, J. F., and Kelley, W. M. (2004). Medial prefrontal activity predicts memory for self. *Cereb. Cortex* 14, 647–654. doi:10.1093/Cercor/Bhh025
- Majdandzic, J., Bauer, H., Windischberger, C., Moser, E., Engl, E., and Lamm, C. (2012). The human factor: behavioral and neural correlates of humanized perception in moral decision making. *PLoS ONE* 7:e47698. doi:10.1371/journal.pone.0047698
- Mano, Y., Harada, T., Sugiura, M., Saito, D. N., and Sadato, N. (2009). Perspective-taking as part of narrative comprehension: a functional MRI study. *Neuropsychologia* 47, 813–824. doi:10.1016/J.Neuropsychologia.2008.12.011
- Markus, H., and Cross, S. (1990). “The interpersonal self,” in *Handbook of Personality Theory and Research*, ed. L. A. Pervin (New York: Guilford), 576–608.
- McGuire, P. K., Silbersweig, D. A., and Frith, C. D. (1996a). Functional neuroanatomy of verbal self-monitoring. *Brain* 119, 907–917. doi:10.1093/Brain/119.3.907
- McGuire, P. K., Paulesu, E., Frackowiak, R. S. J., and Frith, C. D. (1996b). Brain activity during stimulus independent thought. *Neuroreport* 7, 2095–2099.
- McKiernan, K. A., Kaufman, J. N., Kucera-Thompson, J., and Binder, J. R. (2003). A parametric manipulation of factors affecting task-induced deactivation in functional neuroimaging. *J. Cogn. Neurosci.* 15, 394–408. doi:10.1162/089892903321593117
- Mead, G. H. (1934). *Mind, Self and Society from the Standpoint of a Social Behaviorist*. Chicago: University of Chicago Press.
- Moran, J. M., Macrae, C. N., Heatherton, T. F., Wyland, C. L., and Kelley, W. M. (2006). Neuroanatomical evidence for distinct cognitive and affective components of self. *J. Cogn. Neurosci.* 18, 1586–1594. doi:10.1162/Jocn.2006.18.9.1586
- Morita, T., Itakura, S., Saito, D. N., Nakashita, S., Harada, T., Kochiyama, T., et al. (2008). The role of the right prefrontal cortex in self-evaluation of the face: a functional magnetic resonance imaging study. *J. Cogn. Neurosci.* 20, 342–355. doi:10.1162/Jocn.2008.20.2.342
- Nakamura, K., Kawashima, R., Sugiura, M., Kato, T., Nakamura, A., Hatano, K., et al. (2001). Neural substrates for recognition of familiar voices: a PET study. *Neuropsychologia* 39, 1047–1054. doi:10.1016/S0028-3932(01)00037-9
- Neisser, U. (1988). Five kinds of self-knowledge. *Philos. Psychol.* 1, 35–59. doi:10.1080/09515088808572924
- Nelissen, K., Borra, E., Gerbella, M., Rozzi, S., Luppino, G., Vanduffel, W., et al. (2011). Action observation circuits in the macaque monkey cortex. *J. Neurosci.* 31, 3743–3756. doi:10.1523/Jneurosci.4803-10.2011
- Newman-Norlund, R. D., van Schie, H. T., van Zuijlen, A. M. J., and Bekkering, H. (2007). The mirror neuron system is more active during complementary compared with imitative action. *Nat. Neurosci.* 10, 817–818. doi:10.1038/Nn1911
- Northoff, G., and Bermpohl, F. (2004). Cortical midline structures and the self. *Trends Cogn. Sci.* 8, 102–107. doi:10.1016/J.Tics.2004.01.004
- Northoff, G., and Hayes, D. J. (2011). Is our self nothing but reward? *Biol. Psychiatry* 69, 1019–1025. doi:10.1016/j.biopsych.2010.12.014
- Northoff, G., Heinzel, A., Greck, M., Bannpohl, F., Döbrowolny, H., and Panksepp, J. (2006). Self-referential processing in our brain – a meta-analysis of imaging studies on the self. *Neuroimage* 31, 440–457. doi:10.1016/J.Neuroimage.2005.12.002
- Nummenmaa, L., and Calder, A. J. (2009). Neural mechanisms of social attention. *Trends Cogn. Sci.* 13, 135–143. doi:10.1016/J.Tics.2008.12.006
- Ochsner, K. N., Beer, J. S., Robertson, E. R., Cooper, J. C., Gabrieli, J. D. E., Kihlstrom, J. F., et al. (2005). The neural correlates of direct and reflected self-knowledge. *Neuroimage* 28, 797–814. doi:10.1016/J.Neuroimage.2005.06.069
- Ogawa, K., and Inui, T. (2007). Lateralization of the posterior parietal cortex for internal monitoring of self-versus externally generated movements. *J. Cogn. Neurosci.* 19, 1827–1835. doi:10.1162/Jocn.2007.19.11.1827
- Oikawa, H., Sugiura, M., Sekiguchi, A., Tsukiura, T., Miyachi, C. M., Hashimoto, T., et al. (2012). Self-face evaluation and self-esteem in young females: an fMRI study using contrast effect. *Neuroimage* 59, 3668–3676. doi:10.1016/J.Neuroimage.2011.10.098
- Pelphrey, K. A., Viola, R. J., and McCarthy, G. (2004a). When strangers pass – processing of mutual and averted social gaze in the superior temporal sulcus. *Psychol. Sci.* 15, 598–603. doi:10.1111/J.0956-7976.2004.00726.X
- Pelphrey, K. A., Morris, J. P., and McCarthy, G. (2004b). Grasping the intentions of others: the perceived intentionality of an action influences activity in the superior temporal sulcus during social perception. *J. Cogn. Neurosci.* 16, 1706–1716. doi:10.1162/0898929042947900
- Perrin, F., Maquet, P., Peigneux, P., Ruby, P., Degueldre, C., Baetee, E., et al. (2005). Neural mechanisms involved in the detection of our first name: a combined ERPs and PET study. *Neuropsychologia* 43, 12–19. doi:10.1016/J.Neuropsychologia.2004.07.002
- Pfeifer, J. H., Lieberman, M. D., and Dapretto, M. (2007). “I know you are but what am I?!”: neural bases of self- and social knowledge retrieval in children and adults. *J. Cogn. Neurosci.* 19, 1323–1337. doi:10.1162/Jocn.2007.19.8.1323
- Platek, S. M., Keenan, J. P., Gallup, G. G., and Mohamed, F. B. (2004). Where am I? The neurological correlates of self and other. *Cogn. Brain Res.* 19, 114–122. doi:10.1016/J.Cogbrainres.2003.11.014
- Platek, S. M., Loughhead, J. W., Gur, R. C., Busch, S., Ruparel, K., Phend, N., et al. (2006). Neural substrates for functionally discriminating self-face from personally familiar faces. *Hum. Brain Mapp.* 27, 91–98. doi:10.1002/Hbm.20168
- Platek, S. M., Wathne, K., Tierney, N. G., and Thomson, J. W. (2008). Neural correlates of self-face recognition: an effect-location meta-analysis. *Brain Res.* 1232, 173–184. doi:10.1016/J.Brainres.2008.07.010
- Prinz, W. (1997). Perception and action planning. *Eur. J. Cogn. Psychol.* 9, 129–154. doi:10.1080/713752551
- Raichle, M. E., MacLeod, A. M., Snyder, A. Z., Powers, W. J., Gusnard, D. A., and Shulman, G. L. (2001). A default mode of brain function. *Proc. Natl. Acad. Sci. U.S.A.* 98, 676–682. doi:10.1073/Pnas.98.2.676
- Ramsey, R., Hansen, P., Apperly, I., and Samson, D. (2013). Seeing it my way or your way: frontoparietal brain areas sustain

- viewpoint-independent perspective selection processes. *J. Cogn. Neurosci.* 25, 670–684. doi:10.1162/jocn_a_00345 Epub 2012/12/20.
- Rangel, A., Camerer, C., and Montague, P. R. (2008). A framework for studying the neurobiology of value-based decision making. *Nat. Rev. Neurosci.* 9, 545–556. doi:10.1038/Nrn2357
- Rilling, J. K., Sanfey, A. G., Aronson, J. A., Nystrom, L. E., and Cohen, J. D. (2004). The neural correlates of theory of mind within interpersonal interactions. *Neuroimage* 22, 1694–1703. doi:10.1016/J.Neuroimage.2004.04.015
- Rizzolatti, G., and Craighero, L. (2004). The mirror-neuron system. *Annu. Rev. Neurosci.* 27, 169–192. doi:10.1146/Annurev.Neuro.27.070203.144230
- Rizzolatti, G., Fogassi, L., and Gallese, V. (2001). Neurophysiological mechanisms underlying the understanding and imitation of action. *Nat. Rev. Neurosci.* 2, 661–670. doi:10.1038/35090060
- Roberts, B. W., and Donahue, E. M. (1994). One personality, multiple selves – integrating personality and social roles. *J. Pers.* 62, 199–218. doi:10.1111/J.1467-6494.1994.tb00291.x
- Robertson, D., Snarey, J., Ousley, O., Harenski, K., Bowman, E. D., Gilkey, R., et al. (2007). The neural processing of moral sensitivity to issues of justice and care. *Neuropsychologia* 45, 755–766. doi:10.1016/J.Neuropsychologia.2006.08.014
- Ruby, P., and Decety, J. (2001). Effect of subjective perspective taking during simulation of action: a PET investigation of agency. *Nat. Neurosci.* 4, 546–550.
- Ruby, P., and Decety, J. (2003). What you believe versus what you think they believe: a neuroimaging study of conceptual perspective-taking. *Eur. J. Neurosci.* 17, 2475–2480. doi:10.1046/J.1460-9568.2003.02673.x
- Ruby, P., and Decety, J. (2004). How would you feel versus how do you think she would feel? A neuroimaging study of perspective-taking with social emotions. *J. Cogn. Neurosci.* 16, 988–999. doi:10.1162/0898929041502661
- Rushworth, M. F. S., Noonan, M. P., Boorman, E. D., Walton, M. E., and Behrens, T. E. (2011). Frontal cortex and reward-guided learning and decision-making. *Neuron* 70, 1054–1069. doi:10.1016/J.Neuron.2011.05.014
- Sassa, Y., Sugiura, M., Jeong, H., Horie, K., Sato, S., and Kawashima, R. (2007). Cortical mechanism of communicative speech production. *Neuroimage* 37, 985–992. doi:10.1016/J.Neuroimage.2007.05.059
- Sato, A., and Yasuda, A. (2005). Illusion of sense of self-agency: discrepancy between the predicted and actual sensory consequences of actions modulates the sense of self-agency, but not the sense of self-ownership. *Cognition* 94, 241–255. doi:10.1016/J.Cognition.2004.04.003
- Schilbach, L., Wohlschlaeger, A. M., Kraemer, N. C., Newen, A., Shah, N. J., Fink, G. R., et al. (2006). Being with virtual others: neural correlates of social interaction. *Neuropsychologia* 44, 718–730. doi:10.1016/J.Neuropsychologia.2005.07.017
- Schmitz, T. W., Kawahara-Baccus, T. N., and Johnson, S. C. (2004). Metacognitive evaluation, self-relevance, and the right prefrontal cortex. *Neuroimage* 22, 941–947. doi:10.1016/J.Neuroimage.2004.02.018
- Schnell, K., Bluschke, S., Konrad, B., and Walter, H. (2011). Functional relations of empathy and mentalizing: an fMRI study on the neural basis of cognitive empathy. *Neuroimage* 54, 1743–1754. doi:10.1016/J.Neuroimage.2010.08.024
- Schnell, K., Heekeren, K., Schnitker, R., Daumann, J., Weber, J., Hesselmann, V., et al. (2007). An fMRI approach to particularize the frontoparietal network for visuomotor action monitoring: detection of incongruence between test subjects' actions and resulting perceptions. *Neuroimage* 34, 332–341. doi:10.1016/J.Neuroimage.2006.08.027
- Seitz, R. J., Nickel, J., and Azari, N. P. (2006). Functional modularity of the medial prefrontal cortex: involvement in human empathy. *Neuropsychology* 20, 743–751. doi:10.1037/0894-4105.20.6.743
- Senju, A., and Johnson, M. H. (2009). The eye contact effect: mechanisms and development. *Trends Cogn. Sci.* 13, 127–134. doi:10.1016/J.Tics.2008.11.009
- Smith, A. T., Wall, M. B., and Thilo, K. V. (2012). Vestibular inputs to human motion-sensitive visual cortex. *Cereb. Cortex* 22, 1068–1077. doi:10.1093/Cercor/Bhr179
- Somerville, L. H., Kelley, W. M., and Heatherton, T. F. (2010). Self-esteem modulates medial prefrontal cortical responses to evaluative social feedback. *Cereb. Cortex* 20, 3005–3013. doi:10.1093/Cercor/Bhq049
- Spengler, S., von Cramon, D. Y., and Brass, M. (2009). Was it me or was it you? How the sense of agency originates from ideomotor learning revealed by fMRI. *Neuroimage* 46, 290–298. doi:10.1016/J.Neuroimage.2009.01.047
- Spreng, R. N., Mar, R. A., and Kim, A. S. N. (2009). The common neural basis of autobiographical memory, prospection, navigation, theory of mind, and the default mode: a quantitative meta-analysis. *J. Cogn. Neurosci.* 21, 489–510. doi:10.1162/Jocn.2008.21029
- Steuwe, C., Daniels, J. K., Frewen, P. A., Densmore, M., Pannasch, S., Beblo, T., et al. (2012). Effect of direct eye contact in PTSD related to interpersonal trauma: an fMRI study of activation of an innate alarm system. *Soc. Cogn. Affect. Neurosci.* doi:10.1093/scan/nss105 Epub 2012/09/15.
- Stryker, S., and Statham, A. (1985). “Symbolic interaction and role theory,” in *Handbook of Social Psychology*, eds G. Lindzey, and E. Aronson (New York: Random House), 311–378.
- Suddendorf, T., and Collier-Baker, E. (2009). The evolution of primate visual self-recognition: evidence of absence in lesser apes. *Proc. R. Soc. Lond. B Biol. Sci.* 276, 1671–1677. doi:10.1098/Rspb.2008.1754
- Sugiura, M. (2011). “The multi-layered model of self: a social neuroscience perspective,” in *New Frontiers in Social Cognitive Neuroscience*, eds R. Kawashima, M. Sugiura, and T. Tsukiura (Sendai: Tohoku University Press), 111–135.
- Sugiura, M., Gotoh, R., Okada, K., Yamaguchi, K., Itoh, M., Fukuda, H., et al. (2004). Target dependency of brain mechanism involved in dispositional inference: a PET study. *Neuroimage* 21, 1377–1386. doi:10.1016/J.Neuroimage.2003.11.021
- Sugiura, M., Kawashima, R., Nakamura, K., Okada, E., Kato, T., Nakamura, A., et al. (2000). Passive and active recognition of one's own face. *Neuroimage* 11, 36–48. doi:10.1006/Nimg.1999.0519
- Sugiura, M., Sassa, Y., Jeong, H., Horie, K., Sato, S., and Kawashima, R. (2008). Face-specific and domain-general characteristics of cortical responses during self-recognition. *Neuroimage* 42, 414–422. doi:10.1016/J.Neuroimage.2008.03.054
- Sugiura, M., Sassa, Y., Jeong, H., Miura, N., Akitsuki, Y., Horie, K., et al. (2006). Multiple brain networks for visual self-recognition with different sensitivity for motion and body part. *Neuroimage* 32, 1905–1917. doi:10.1016/J.Neuroimage.2006.05.026
- Sugiura, M., Sassa, Y., Jeong, H., Wakusawa, K., Horie, K., Sato, S., et al. (2012). Self-face recognition in social context. *Hum. Brain Mapp.* 33, 1364–1374. doi:10.1002/Hbm.21290
- Sugiura, M., Watanabe, J., Maeda, Y., Matsue, Y., Fukuda, H., and Kawashima, R. (2005a). Cortical mechanisms of visual self-recognition. *Neuroimage* 24, 143–149. doi:10.1016/J.Neuroimage.2004.07.063
- Sugiura, M., Shah, N. J., Zilles, K., and Fink, G. R. (2005b). Cortical representations of personally familiar objects and places: functional organization of the human posterior cingulate cortex. *J. Cogn. Neurosci.* 17, 183–198. doi:10.1162/0898929053124956
- Summerfield, J. J., Hassabis, D., and Maguire, E. A. (2009). Cortical midline involvement in autobiographical memory. *Neuroimage* 44, 1188–1200. doi:10.1016/J.Neuroimage.2008.09.033
- Svoboda, E., McKinnon, M. C., and Levine, B. (2006). The functional neuroanatomy of autobiographical memory: a meta-analysis. *Neuropsychologia* 44, 2189–2208. doi:10.1016/J.Neuropsychologia.2006.05.023
- Synofzik, M., Thier, P., and Lindner, A. (2006). Internalizing agency of self-action: perception of one's own hand movements depends on an adaptable prediction about the sensory action outcome. *J. Neurophysiol.* 96, 1592–1601. doi:10.1152/Jn.00104.2006
- Tacikowski, P., Brechmann, A., Marchewka, A., Jednorog, K., Dobrowolny, M., and Nowicka, A. (2011). Is it about the self or the significance? An fMRI study of self-name recognition. *Soc. Neurosci.* 6, 98–107. doi:10.1080/17470919.2010.490665
- Tacikowski, P., Brechmann, A., and Nowicka, A. (2012). Cross-modal pattern of brain activations associated with the processing of self- and significant other's name. *Hum. Brain Mapp.* 34, 2069–2077. doi:10.1002/hbm.22048
- Tsakiris, M., and Haggard, P. (2005). The rubber hand illusion revisited: visuotactile integration and self-attribution. *J. Exp. Psychol. Hum. Percept. Perform.* 31, 80–91. doi:10.1037/0096-1523.31.1.80
- Tsakiris, M., Hesse, M. D., Boy, C., Haggard, P., and Fink, G. R. (2007). Neural signatures of body ownership: a sensory network for bodily self-consciousness. *Cereb. Cortex* 17, 2235–2244. doi:10.1093/Cercor/Bhl131
- Tsakiris, M., Prabhu, G., and Haggard, P. (2006). Having a body versus moving your body: how

- agency structures body-ownership. *Conscious. Cogn.* 15, 423–432. doi:10.1016/J.Concog.2005.09.004
- Uddin, L. Q., Iacoboni, M., Lange, C., and Keenan, J. P. (2007). The self and social cognition: the role of cortical midline structures and mirror neurons. *Trends Cogn. Sci.* 11, 153–157. doi:10.1016/J.Tics.2007.01.001
- Uddin, L. Q., Kaplan, J. T., Molnar-Szakacs, I., Zaidel, E., and Iacoboni, M. (2005). Self-face recognition activates a frontoparietal “mirror” network in the right hemisphere: an event-related fMRI study. *Neuroimage* 25, 926–935. doi:10.1016/J.Neuroimage.2004.12.018
- Van Overwalle, F. (2009). Social cognition and the brain: a meta-analysis. *Hum. Brain Mapp.* 30, 829–858. doi:10.1002/Hbm.20547
- Vogeley, K., May, M., Ritzl, A., Falkai, P., Zilles, K., and Fink, G. R. (2004). Neural correlates of first-person perspective as one constituent of human self-consciousness. *J. Cogn. Neurosci.* 16, 817–827. doi:10.1162/089892904970799
- Vollm, B. A., Taylor, A. N. W., Richardson, P., Corcoran, R., Stirling, J., McKie, S., et al. (2006). Neuronal correlates of theory of mind and empathy: a functional magnetic resonance imaging study in a nonverbal task. *Neuroimage* 29, 90–98. doi:10.1016/J.Neuroimage.2005.07.022
- Wagner, A. D., Shannon, B. J., Kahn, I., and Buckner, R. L. (2005). Parietal lobe contributions to episodic memory retrieval. *Trends Cogn. Sci.* 9, 445–453. doi:10.1016/J.Tics.2005.07.001
- Wakusawa, K., Sugiura, M., Sassa, Y., Jeong, H., Horie, K., Sato, S., et al. (2009). Neural correlates of processing situational relationships between a part and the whole: an fMRI study. *Neuroimage* 48, 486–496. doi:10.1016/J.Neuroimage.2009.06.024
- Wegner, D. M., and Wheatley, T. (1999). Apparent mental causation – sources of the experience of will. *Am. Psychol.* 54, 480–492. doi:10.1037/0003-066x.54.7.480
- Weiskrantz, L., Elliott, J., and Darlington, C. (1971). Preliminary observations on tickling oneself. *Nature* 230, 598–599. doi:10.1038/230598a0 Epub 1971/04/30.
- Wicker, B., Perrett, D. I., Baron-Cohen, S., and Decety, J. (2003). Being the target of another’s emotion: a PET study. *Neuropsychologia* 41, 139–146. doi:10.1016/S0028-3932(02)00144-6
- Wolpert, D. M., Ghahramani, Z., and Jordan, M. I. (1995). An Internal Model for Sensorimotor Integration. *Science* 269, 1880–1882. doi:10.1126/Science.7569931
- Yomogida, Y., Sugiura, M., Sassa, Y., Wakusawa, K., Sekiguchi, A., Fukushima, A., et al. (2010). The neural basis of agency: an fMRI study. *Neuroimage* 50, 198–207. doi:10.1016/J.Neuroimage.2009.12.054
- Zhang, S., and Li, C. S. R. (2012). Functional connectivity mapping of the human precuneus by resting state fMRI. *Neuroimage* 59, 3548–3562. doi:10.1016/J.Neuroimage.2011.11.023
- zu Eulenburg, P. Z., Caspers, S., Roski, C., and Eickhoff, S. B. (2012). Meta-analytical definition and functional connectivity of the human vestibular cortex. *Neuroimage* 60, 162–169. doi:10.1016/J.Neuroimage.2011.12.032

Conflict of Interest Statement: The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Received: 04 June 2013; paper pending published: 03 July 2013; accepted: 16 August 2013; published online: 30 August 2013.

Citation: Sugiura M (2013) Associative account of self-cognition: extended forward model and multi-layer structure. *Front. Hum. Neurosci.* 7:535. doi:10.3389/fnhum.2013.00535

This article was submitted to the journal *Frontiers in Human Neuroscience*.

Copyright © 2013 Sugiura. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.