



Integrating simulation and theory of mind: from self to social cognition

Christian Keyzers and Valeria Gazzola

BCN NeuroImaging Center, University Medical Center Groningen, University of Groningen, A. Deusinglaan 2, 9713AW, The Netherlands

Investigations of brain substrates for social cognition have polarized in two camps. The simulation camp focuses on so-called shared circuits (SCs) that are involved in one's own actions, sensations and emotions and in perceiving those of others [1,2]. The theory of mind (ToM) camp emphasizes the role of midline structures in mentalizing about the states of others [3]. Scientific energy has often flown into fruitless arguments about which camp is closer to the truth [4], but the true questions for contemporary social neuroscience should be (i) why do investigators find different sets of areas to be most prominent, and (ii) how do the two sets of brain areas interact? Here we propose a highly speculative model that complements the view of Uddin *et al.* [5] to stimulate and canalize future empirical work into a direction we believe to be promising.

Social cognitions range from the intuitive examples studied by simulationists to the reflective ones used by ToM investigators. Witnessing someone drink a glass of milk with a face contracting in an expression of disgust is an example at the intuitive extreme of this continuum. In such cases, premotor and parietal areas for actions [6], the insula for emotions [7,8] and SII [9] for sensations form SCs that translate the bodily states of others into the neural language of our own states [2]. These SCs seem to implement a pre-reflective, intuitive and empathic level of representation: neural activity in these areas does not require specific instructions that encourage conscious reflections. More activity in these areas correlates with higher scores in empathy questionnaires [6–8] and lower scores in autism inventories [10]. Thinking about what gift would please a foreign colleague is an example at the more reflective extreme. In such cases, we must browse consciously through what we know about his country and culture to deduce what he might like. Such explicit knowledge about the inner life of others is the product of reflecting upon the states of others and is linked with activity in midline structures [3] and the temporoparietal junction [4]. False beliefs are prototypical examples of such reflective representations [3]. Given these differences between intuitive and reflective examples, it is not surprising that investigators find different brain areas to be involved. Social cognition, though, is often neither purely intuitive nor purely reflective, which is why knowledge of both mechanisms should be combined.

Two recent developments help us hypothesize about how these two sets of areas integrate. First, examining how our brain processes our own states has shown that a similar distinction of pre-reflective and reflective representations applies. Pre-reflective representations of visceral states of the self, for instance, seem linked to activations in the posterior and/or middle insula. By contrast, midline structures become active when subjects are asked to introspect, reflect and report these states (e.g. heartbeat) and the anterior insula seems crucial in linking the more posterior insula with these midline structures [11]. Thus, the human brain does not only represent its own bodily state but enables us to consciously introspect these states through an additional layer of reflective meta-representations that depend on midline structures and the anterior insula. Second, Mitchell and colleagues [12] asked participants to judge how much certain statements (e.g. 'to enjoy having a room-mate from a different country') applied to themselves, to another person they felt was similar to themselves and to a person they considered to be dissimilar. Their statements involved feelings, linking them to SC experiments, and verbal reports about the mental states of others, linking them to ToM experiments. Ventral aspects of the medial prefrontal cortex (mPFC), similar to those found during heartbeat introspection, were involved in judgments that related to the self and to similar others. Therefore, the ventromedial prefrontal cortex (vmPFC) might be an SC for reflective representations, regardless of whether they are about the self or about others. By contrast, a more dorsal sector was involved only when reflecting about people considered to be dissimilar.

An embryonic and speculative working hypothesis [2] for the link between classical SCs and midline structures, which is inspired by these findings but differs from that in Uddin *et al.* [5], is shown in Figure 1. While dealing with states of the self, areas of the SCs represent pre-reflective bodily states. If asked to introspect and report these states, subjects additionally activate (v)mPFC structures. When dealing with states of other individuals, activity in SCs might represent the empathic transformation of the bodily states of others into pre-reflective neural representations of similar states of the self. These simulated pre-reflective representations correlate with empathy [6–8] and might provide an intuitive understanding of what goes on in others [2]. If asked to reflect on the states of others, the pathways that are normally used to reflect on the bodily representations of the self are now used on simulated bodily states of others, leading to simulated

Corresponding author: Keyzers, C. (c.m.keyzers@rug.nl).
Available online xxxxxx.

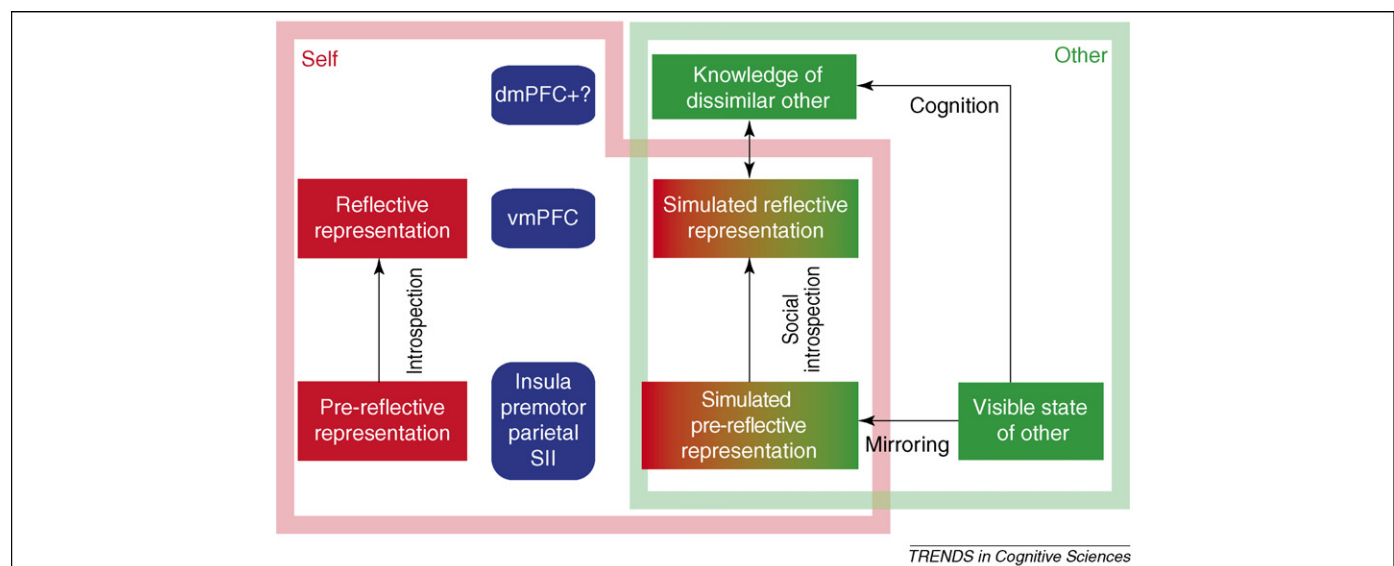


Figure 1. Illustration of the model. The self is shown in red, the other is shown in green and candidate brain areas that are thought to implement representation are shown in blue. During our own experiences, pre-reflective representations can lead, through introspection, to reflective representations (red). While witnessing the states of others, mirroring leads to activations that simulate pre-reflective representations of our own bodily states. A process of social introspection, utilizing the mechanisms of introspection, activates representations that simulate reflective representations of our own bodily states. A more cognitive route leads to more abstract knowledge about the other that escapes from the constraints of our own experiences.

reflective representations. Thus, SCs and midline structures form an integrated system that applies to cases where we perceive the other as similar enough for simulation to be useful [12]. In this view, both SCs and vmPFC reflect simulation, albeit at different levels (pre-reflective versus reflective), rather than radically different processes (SC versus ToM). This route is complemented by a more dorsal, less embodied and more cognitive route that becomes essential when simulations lead to wrong conclusions (e.g. others considered to be dissimilar in this aspect [12]) and might be the most intact route in high-functioning autism. This integrational model is highly tentative: little is known about the function of many of the areas involved. We hope this model stimulates and directs research towards a more mature integration of SCs and ToM.

In conclusion, much of the debate in social cognitions might result from choosing tasks that isolate the processes of just one route in the laboratory. However, it is essential to start designing tasks that reflect the complexity of social life to test how the social brain forms an integrated whole.

Acknowledgements

C.K. is supported by a Marie Curie Excellence Grant of the EU and a VIDI grant from the NWO. The model put forward was presented at the

Cognitive Neuroscience and Human Brain Mapping Conferences in 2006 as symposia.

References

- 1 Gallese, V. *et al.* (2004) A unifying view of the basis of social cognition. *Trends Cogn. Sci.* 8, 396–403
- 2 Keysers, C. and Gazzola, V. (2006) Towards a unifying neural theory of social cognition. *Prog. Brain Res.* 156, 379–401
- 3 Amodio, D.M. and Frith, C.D. (2006) Meeting of minds: the medial frontal cortex and social cognition. *Nat. Rev. Neurosci.* 7, 268–277
- 4 Saxe, R. (2005) Against simulation: the argument from error. *Trends Cogn. Sci.* 9, 174–179
- 5 Uddin, L.Q. *et al.* (2007) The self and social cognition: the role of cortical midline structures and mirror neurons. *Trends Cogn. Sci.* DOI: 10.1016/j.tics.2007.01.001 (www.sciencedirect.com)
- 6 Gazzola, V. *et al.* (2006) Empathy and the somatotopic auditory mirror system in humans. *Curr. Biol.* 16, 1824–1829
- 7 Jabbi, M. *et al.* (2007) Empathy for positive and negative emotions in the gustatory cortex. *Neuroimage* 34, 1744–1753
- 8 Singer, T. *et al.* (2004) Empathy for pain involves the affective but not sensory components of pain. *Science* 303, 1157–1162
- 9 Keysers, C. *et al.* (2004) A touching sight: SII/PV activation during the observation and experience of touch. *Neuron* 42, 335–346
- 10 Dapretto, M. *et al.* (2006) Understanding emotions in others: mirror neuron dysfunction in children with autism spectrum disorders. *Nat. Neurosci.* 9, 28–30
- 11 Critchley, H.D. *et al.* (2004) Neural systems supporting interoceptive awareness. *Nat. Neurosci.* 7, 189–195
- 12 Mitchell, J.P. *et al.* (2006) Dissociable medial prefrontal contributions to judgments of similar and dissimilar others. *Neuron* 50, 655–663