aps ASSOCIATION FOR PSYCHOLOGICAL SCIENCE

Current Directions in Psychological Science
20(3) 197-200

DOI: 10.1177/0963721411409007
http://cdps.sagepub.com
SAGE

Two Mechanisms for Simulating Other Minds: Dissociations Between Mirroring and Self-Projection

Adam Waytz and Jason P. Mitchell
Harvard University

Abstract

People often attempt to understand other minds by using their own thoughts and experiences as a proxy for those of others, a process known broadly as simulation. Recent research in cognitive neuroscience has identified the neural bases of two forms of simulation: mirroring and self-projection. Mirroring involves a vicarious response in which a perceiver experiences the same current mental state as that of another person, and has been linked recently to brain regions that "mirror" the experiential states of others. In contrast, self-projection involves imagining oneself in the same situation as another person, predicting one's thoughts and feelings in that hypothetical scenario and assuming that the other would think and feel the same way. This form of simulation has been linked to a set of regions known collectively as the default network and includes the medial prefrontal cortex, precuneus and posterior cingulate, and lateral parietal cortex. Although most discussions of simulation have conflated these two processes, here we describe the conceptual and empirical reasons to distinguish between self-projection and mirroring and suggest the unique role each plays in understanding others.

Keywords

mentalizing, theory of mind, simulation, self-projection, mirror system

Despite never directly perceiving others' thoughts and feelings, humans can use an important source of understanding the internal experience of others: their own minds. This ability—called simulation—involves using one's own thoughts, feelings, and intentions as a guide to what others are thinking, feeling, or intending. Although considerable research has demonstrated that people use themselves to understand others' mental states, theoretical discussions have tended to conflate two distinct types of simulation: mirroring and self-projection. Perceivers can understand others by mirroring their experience, vicariously experiencing another person's mental states. At the same time, perceivers can also project themselves into another person's situation, consider the thoughts and feelings engendered by that situation, and then infer that the other would think and feel similarly. Here we distinguish between these two forms of simulation by reviewing their different functions and dissociable neural bases.

The notion that humans can adopt the mental states of others as their own dates at least as far back as David Hume (1739/1958, xi), who described emotional resonance with others as the basis for sympathy, noting, "A cheerful countenance infuses a sensible complacency and serenity into my mind; as an angry or sorrowful one throws a damp upon me . . . we never remark any passion or principle in others, of which . . . we may not find a parallel in ourselves." In the later part of the 20th century, philosophers of mind attempted to define simulation more formally. For example, Gordon (1986) described simulation essentially an identical process to predicting and generating one's own actions—"*deciding what to do*, but extended to people of 'minds' different from one's own" (p. 162, italics in original).

These definitions vary in terms of the immediacy with which perceivers interact with other minds. Hume described the ability to vicariously experience another person's state—feeling as that person feels—at the very same moment as the other. In contrast, Gordon described a process involving at least two steps—imagining one's own experience and then projecting it onto another person. These two extremes capture the primary distinction between mirroring and self-projection. Whereas the former involves a near-simultaneous resonance with the other person's experience, the latter involves first imagining "off-line" what one's own thoughts and feelings would be in the other person's

Corresponding Author:
Jason Mitchell or Adam Waytz, Harvard University, Northwest Science Building, 52 Oxford Street, Cambridge, MA 02138
E-mail: jason_mitchell@harvard.edu or waytz@wjh.harvard.edu

situation and then projecting them onto that person. Recent research in cognitive neuroscience has supported this conceptual distinction by identifying dissociable neural mechanisms underlying both forms of simulation.

Neural Bases of Mirroring

Considerable interest in the ability to mirror others' experience followed the discovery that some neurons in the macaque premotor cortex respond when a monkey either observes or executes an action (di Pellegrino, Fadiga, Fogassi, Gallese, & Rizzolatti, 1992). Since these initial observations, researchers have reported dozens of experiments in both macaques and humans that consistently identify a set of regions in the inferior frontal cortex and superior parietal lobe (i.e., the parieto-frontal circuit) that are involved both in the production of goal-directed motor action and observation of others' goal-directed actions (for a recent review, see Rizzolatti & Sinigaglia, 2010; see Hickok, 2009, for an alternative view).

Moreover, researchers have repeatedly observed similar mirroring processes across a variety of mental states other than goals and intentions. Observing another person in pain activates many of the same regions involved in the subjective experience of pain, including the rostral anterior cingulate cortex and cerebellum (Singer et al., 2004). Observing others receiving monetary gains activates regions involved in the subjective experience of reward, such as the ventral striatum (Mobbs et al., 2009). Observing others experience disgust activates regions involved in the subjective feeling of disgust, such as the insula (Wicker et al., 2003). And observing others experience fear activates regions such as the amygdala (Whalen et al., 2001), a region also involved in the subjective experience of fear and anxiety. Humans appear to have a general tendency to adopt the pattern of brain activity associated with experiencing a particular kind of mental state when merely watching another person having that same type of experience—that is, to mirror others' mental states in themselves.

Neural Bases of Self-Projection

Over the past few years, researchers have demonstrated that a network of brain regions subserves the ability to imagine experiences unrelated to one's current perceptual environment. This set of regions—known collectively as the default network—comprises the medial prefrontal cortex, precuneus and posterior cingulate, and lateral parietal cortex (Raichle et al., 2001) and has been linked repeatedly to tasks in which people consider their own mental states and traits, imagine themselves experiencing fictitious events, consider experiencing possible events in the future, or recall their experiences from the past (Buckner & Carroll, 2007). Interestingly, involvement of these same regions has consistently been reported in studies in which participants explicitly consider the mental states of other people (Frith & Frith, 2003). That is, the same neural regions that subserve the human capacity to contemplate oneself experiencing events outside of the here-and-now also subserve the capacity to imagine the mental states of other people. This overlap suggests that the cognitive mechanisms underlying people's ability to mentally conjure up fictitious experiences are also those that allow people to mentalize about the internal states of others, consistent with the notion that perceivers engage in similar types of self-projection regardless of whether the counterfactual experience they are considering involves thoughts about a hypothetical time and place or the possible goings-on of another mind (Tamir & Mitchell, 2011).

Dissociable Functions of Mirroring and Self-Projection

Existing data suggest that mirroring and self-projection operate in distinct contexts, depending on the kind of information available to perceivers about another person's mental states. Mirroring typically occurs when perceivers can make use of observable cues about what other persons are experiencing. Perceivers mirror when they see or hear another person's physical actions, when they witness a clearly painful situation such as a needle penetrating a person's hand, or when they observe another person's emotional expression. In other words, perceptible cues to another person's experience seem to evoke a similar experience in oneself, as reflected in vicarious neural responses to others' actions that reflect their mental states (de Lange, Spronk, Willems, Toni, & Bekkering, 2008).

However, many instances in which people consider the mental states of another person occur when that person is not immediately present. People anticipate their boss's mood before asking for a raise, wonder how their relatives in distant geographical locations are feeling, consider their opponent's intentions before entering a business negotiation, and ruminate on the effect of their words on others long after speaking them. In other words, perceivers can readily contemplate minds that exist at a significant distance from the here-and-now in the absence of external cues such as facial expression, vocal tone, or bodily movement that could provide observable information about others' internal states.

Consistent with their putative role in self-projection of this kind, activation in regions of the default network has typically been observed when perceivers consider the minds of individuals that they cannot physically perceive. For example, in some of the earliest studies to suggest the role of these regions in mentalizing, participants were asked about the potential mental states of Christopher Columbus (Goel, Grafman, Sadato, & Hallett, 1995) or about fictitious characters in cartoons and stories (Gallagher et al., 2000). Likewise, activation in this network of regions is routinely observed when participants read narrative stories that are, by definition, devoid of observable cues to others' mental states (Saxe & Kanwisher, 2003). Even when perceivers can see the person whose mind they are to consider, regions of the default network contribute to mentalizing when physical cues are irrelevant to the specific inferences to be made—for example, when judging another person's stable personality characteristics or dispositions (Mitchell, Macrae, & Banaji, 2004).

One of the few studies to simultaneously examine the neural bases of self-projection and mirroring demonstrated the distinct role of each process during mentalizing (Zaki, Hennigan, Weber, & Ochsner, 2010; see also Brass, Schmitt, Spengler, & Gergely, 2007; Spunt, Falk, & Lieberman, 2010). In this study, participants inferred a target's emotional state under three conditions. During perceptually cued trials, participants watched a silent video of the target describing an emotionally evocative autobiographical event (unaccompanied by any contextual description of the event). During context-only trials, participants made their inferences on the basis of a sentence describing the event that the person had experienced (unaccompanied by video). And during combined trials, participants received both perceptual and contextual information about the emotions of the other person. Consistent with the proposed division of labor between two systems for mentalizing, perceptual cues to the target's mental states elicited stronger activation in brain regions involved in mirroring (i.e., the fronto-parietal circuit), whereas contextual cues engaged the default network. The presence of both sources of information was accompanied by activation in both sets of regions, suggesting that a complete understanding of others' inner experiences relies on brain regions that integrate both perceptible external cues to and imagined simulations of others' mental states.

The Heterogeneity of Social Cognition

In suggesting that the human mind uses at least two forms of simulation, these findings support a view of social cognition as a collection of mental processes, each specialized for making sense of others under specific circumstances. Just as humans make use of several different perceptual senses (sight, smell, taste, touch, hearing) to represent the physical world around them, humans use several different social-cognitive processes to construct a useful representation of the social world around them. These distinct processes provide different information about other minds, handle different content, and are more or less useful depending on the goals of the perceiver and the specifics of the social environment. Here, we suggest that two of these processes rely on using one's own mental states as a proxy for those of others, a strategy broadly known as simulation. Recent findings from cognitive neuroscience suggest that these distinct forms of simulation—mirroring and self-projection—may operate under different situational constraints, such as whether relevant perceptual cues to another person's mental states are available.

These distinctions serve to remind us that, rather than a single capacity, human social cognition comprises a rich set of mechanisms for understanding other minds. In addition to relying on different forms of simulation, perceivers make inferences about others by using categorical information (i.e., stereotypes), unique information about specific individuals, and theories about what other people are generally like. Although initial attempts to reduce social inference to a single cognitive process produced spirited debates over whether humans mentalize only in one or another manner, research has now demonstrated that social cognition relies on a suite of cognitive processes that, far from being mutually exclusive, interact in important ways. Armed with new methods for understanding the neural basis of social cognition, researchers continue to make considerable progress in understanding the contexts in which perceivers avail themselves of these various strategies for making sense of others.

Recommended Reading

Brass, M., Schmitt, R.M., Spengler, S., & Gergeley. (2007). (See References). A study that dissociates the role of brain regions involved in mirroring and brain regions involved in more inferential mentalizing.

Keysers, C., & Gazzola, V. (2007). Integrating simulation and theory of mind. *Trends in Cognitive Sciences*, 11, 194–196. An additional framework for integrating the mechanisms of mirroring and more reflective forms of mentalizing such as self–projection.

Rizzolatti, G., & Craighero, L. (2004). The mirror-neuron system. *Annual Review of Neuroscience*, 27, 169–192. A comprehensive review of the initial evidence in support of the existence of the mirror-neuron system and its function.

Zaki, J., Hennigan, K., Weber, J., Ochsner, K. (2010). (See References). A clear empirical demonstration of the dissociation between self-projection and mirroring.

Acknowledgments

We thank Jamil Zaki for comments on a previous draft of this manuscript.

Declaration of Conflicting Interests

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

References

Brass, M., Schmitt, R.M., Spengler, S., & Gergely, G. (2007). Investigating action understanding: Inferential processes versus action simulation. *Current Biology*, *17*, 2117–2121.

Buckner, R.L., & Carroll, D.C. (2007). Self-projection and the brain. *Trends in Cognitive Sciences*, *11*, 49–57.

de Lange, F.P., Spronk, M., Willems, R.M., Toni, I., & Bekkering, H. (2008). Complementary systems for understanding action intentions. *Current Biology*, *18*, 454–457.

di Pellegrino, G., Fadiga, L., Fogassi, L., Gallese, V., & Rizzolatti, G. (1992). Understanding motor events: A neurophysiological study. *Experimental Brain Research*, *91*, 176–180.

Frith, U., & Frith, C.D. (2003). Development and neurophysiology of mentalizing. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, *459*, 358.

Gallagher, H.L., Happe, F., Brunswick, N., Fletcher, P.C., Frith, U., & Frith, C.D. (2000). Reading the mind in cartoons and stories: An fMRI study of 'theory of mind' in verbal and nonverbal tasks. *Neuropsychologia*, *38*, 11–21.

Goel, V., Grafman, J., Sadato, N., & Hallett, M. (1995). Modeling other minds. *Neuroreport*, *6*, 1741–1746.

Gordon, R. (1986). Folk psychology as simulation. *Mind & Language*, *1*, 158–171.

Hickok, G. (2009). Eight problems for the mirror neuron theory of action understanding in monkeys and humans. *Journal of Cognitive Neuroscience*, *21*, 1229–1243.

Hume, D. (1958). *A treatise of human nature*. Oxford, England: Oxford University Press. Original work published 1739.

Mitchell, J.P., Macrae, C.N., & Banaji, M.R. (2004). Encoding specific effects of social cognition on the neural correlates of subsequent memory. *Journal of Neuroscience*, *24*, 4912–4917.

Mobbs, D., Yu, R., Meyer, M., Passamonti, L., Seymour, B., Calder, A.J., et al. (2009). A key role for similarity in vicarious reward. *Science*, *324*, 900.

Raichle, M.E., MacLeod, A.M., Snyder, A.Z., Powers, W.J., Gusnard, D.A., & Shulman, G.L. (2001). A default mode of brain function. *Proceedings of the National Academy of Sciences USA*, *98*, 676–682.

Rizzolatti, G., & Sinigaglia, C. (2010). The functional role of the parieto-frontal mirror circuit: Interpretations and misinterpretations. *Nature Reviews Neuroscience*, *11*, 264–274.

Saxe, R., & Kanwisher, N. (2003). People thinking about thinking people: fMRI investigations of theory of mind. *NeuroImage*, *19*, 1835–1842.

Singer, T., Seymour, B., O'Doherty, J., Kaube, H., Dolan, R.J., & Frith, C.D. (2004). Empathy for pain involves the affective but not sensory components of pain. *Science*, *303*, 1157–1162.

Spunt, R.P., Falk, E.B., & Lieberman, M.D. (2010). Dissociable neural systems support retrieval of *how* and *why* action knowledge. *Psychological Science*, *21*, 1593–1598.

Tamir, D.I., & Mitchell, J.P. (in press). The default network distinguishes construals of proximal versus distal events. *Journal of Cognitive Neuroscience.*

Whalen, P.J., Shin, L.M., McInerney, S.C., Fischer, H., Wright, C.I., & Rauch, S.L. (2001). A functional MRI study of human amygdala responses to facial expressions of fear versus anger. *Emotion*, *1*, 70–83.

Wicker, B., Keysers, C., Plailly, J., Royet, J.P., Gallese, V., & Rizzolatti, G. (2003). Both of us disgusted in my insula: The common neural basis of seeing and feeling disgust. *Neuron*, *40*, 655–664.

Zaki, J., Hennigan, K., Weber, J., & Ochsner, K.N. (2010). Social cognitive conflict resolution: Contributions of domain-general and domain-specific neural systems. *Journal of Neuroscience*, *30*, 8481–8488.