Understanding Intentions : Distinct Processes for Mirroring, Representing, and Conceptualizing

Claudia Chiavarino, Ian A. Apperly and Glyn W. Humphreys

Current Directions in Psychological Science 2012 21: 284
DOI: 10.1177/0963721412452727

Current Directions in Psychological Science
21(5) 284–289

Reprints and permission:
sagepub.com/journalsPermissions.nav
DOI: 10.1177/0963721412452727
http://cdps.sagepub.com

Understanding Intentions: Distinct Processes for Mirroring, Representing, and Conceptualizing

Claudia Chiavarino[1], Ian A. Apperly[2], and Glyn W. Humphreys[3]
[1]Center for Cognitive Science, Department of Psychology, University of Turin; [2]School of Psychology, University of Birmingham; and [3]Department of Experimental Psychology, University of Oxford

Abstract

We provide converging evidence from developmental, imaging, and lesion studies that intentions can be processed at three distinct levels: a mirroring level, which infers immediate action goals on the basis of observed actions; a representational level, which is concerned with the psychological—rather than merely behavioral—representation of the mental states that underlie those actions; and a conceptual level, which allows people to reason about the semantic and logical properties of mental states. Together, the representational and conceptual levels form what is currently referred to as the *mentalizing* system. We argue that although the mirroring and mentalizing systems may work independently of each other, within the mentalizing system, the representational level subserves the conceptual level.

Keywords

intention, mirror system, mentalizing, mental representation, conceptual representation

Intentions and Behavior

Humans extract intentions from other people's behavior and perform complex mental-state reasoning quickly and apparently without effort. Let us imagine that a young boy enters the kitchen asking, "Dad, can you give me an apple?" and that, following this request, his father grasps an apple from a basket full of fruit and hands it to him. Everyone observing this scenario would immediately agree that the man's intention was to give the apple to his son. Now consider the more complex case of teasing: The father may extend his arm toward the basket, omit to grasp any fruit, and then present his empty, open hand to his son. Even in this case, a person observing the scene would likely have no doubt about the man's intention—that is, to pretend to grasp the apple. In this article, we use intentions as a case study from which we draw general lessons about the cognitive and neural basis of people's ability to understand others' mental states.

If there is agreement on the notion that intentions can be understood at different levels of complexity, from action goals to abstract mental-state reasoning (Grafton & Hamilton, 2007), there remains much to be understood about the nature of the mechanisms involved in intention understanding and how they relate to each other. In trying to understand how one "reads" other people's intentions in a given instance, one crucial issue is whether the use of behavioral cues is necessary and sufficient for performing such a skill-demanding process, or whether the representation of mental states is also required. If we do not know that the man in our example acts according to internally represented goals, can we still interpret his behavior as intentional? An additional matter, often overlooked, is that a full conceptualization of intentions entails more than the general recognition that intentions are mental representations that mediate actions in the world. A mature concept of intention involves the appreciation of its motivational, causal, and epistemic aspects (Moses, 2001). People engage in a goal-directed action because they desire something (motivational component). However, the action can be qualified as intentional only if they also believe that the desired outcome can be achieved and that they can perform the action necessary to achieve it (epistemic component), and if they actually try to bring it about (causal component). The consideration of these aspects suggests that understanding intentions is considerably more complex than is commonly supposed.

Intention Processing in the Brain

A valuable insight into the functional organization of intention processing is offered by the investigation of its neural

Corresponding Author:
Claudia Chiavarino, Department of Psychology, Center for Cognitive Science, University of Turin, Via Po 14, 10123 Turin, Italy
E-mail: claudia.chiavarino@unito.it

correlates. Most of this evidence has come from the imaging literature, which has consistently implicated two brain systems: the *mirroring* system (premotor cortex and anterior intraparietal sulcus) and the *mentalizing* system (temporoparietal junction, or TPJ, medial prefrontal cortex, and precuneus; for a review, see Van Overwalle & Baetens, 2009).The functional role of the mirroring system is reasonably well known. This system consistently responds to observation of moving body parts, irrespective of the modality of presentation of the stimuli (e.g., visual, auditory, verbal), and appears to code the immediate goal of an action (e.g., grasping an apple) as well as its physical implementation (e.g., extending one's open hand toward the apple) and its final outcome (e.g., handing the apple to someone; Rizzolatti & Sinigaglia, 2010). In other words, when a person observes another person executing an action, parts of his or her motor system become activated as if he or she were going to perform the same sequence of motor actions directed to the same final outcome. This "embodied simulation" of the observed action affords a basic understanding of the immediate and final concrete, physical goal[1] of an action, as well as of how the action is performed (Gallese, 2007; Spunt, Satpute, & Lieberman, 2010).

However, current accounts of the function of the mentalizing system remain more speculative. The neuroimaging literature has shown that the system is mainly activated by two kinds of tasks: first, tasks in which biological-motion information is not available, in diverse conditions ranging from observations of simple two-dimensional geometric shapes moving in an apparently intentional fashion (e.g., a big triangle "chasing" a small triangle) to the making of complex abstract inferences about the intentions and beliefs of the main character of a verbal story (Castelli, Happé, Frith, & Frith, 2000; de Lange, Spronk, Willems, Toni, & Bekkering, 2008); and second, tasks that involve the observation of moving body parts and simultaneous reflection on the intention of the actor, whether prompted by task instructions or by simple observation of unusual or pretended actions (Brass, Schmitt, Spengler, & Gergely, 2007; Spunt et al., 2010). Observing that concurrent activation of the mirroring and mentalizing systems is rare, Van Overwalle and Baetens (2009) argued that these systems process human actions and intentions in qualitatively different and functionally independent ways. Consistent with this hypothesis, the activity of the mentalizing system is modulated by the degree to which mental-state inference is required by the situation, whereas the activity of the mirroring system is not (Brass et al., 2007; Spunt et al., 2010).

Beyond this growing consensus of a distinction between the mirroring and mentalizing systems, there have been some suggestions of a need for further subdivision. For example, Keysers and Gazzola (2007) have suggested that yet another system may be necessary for stimulus-independent abstract processing of mental states, whereas Mason, Magee, Kuwabara, and Nind (2010) have suggested that such social reasoning requires the mentalizing system to be supplemented by general-purpose brain circuits for high-level cognitive operations. However, although these researchers have drawn attention to the potential limitations of seeing mentalizing as a single, coherent process, they have not provided a clear alternative to it or a clear account of the neural organization of the different systems implicated in mentalizing. We turn first to the developmental literature as a potential source of such insight.

Developing Understanding of Intentions

Developmental research has suggested that some aspects of intentional understanding appear early in infancy. By 6 months of age, infants begin to understand that agents' behaviors may be associated with certain objects, and by the time infants are 1 year old, they can interpret identical behaviors as being either goal-directed or not on the basis of causal context (for a review, see Tomasello, Carpenter, Call, Behne, & Moll, 2005). Older infants show a working understanding of pretended versus "real" actions (e.g., Ma & Lillard, 2006) and of individuals' preferences (e.g., Repacholi & Gopnik, 1997), abilities that may well exceed the capacity of the mirroring system and imply the development of at least a basic mentalizing ability.[2]

Other research has suggested that children's understanding of intentions deepens through a much more protracted developmental process. Under the umbrella of research on "theory of mind," developmental researchers have often distinguished between two requirements for children's understanding of mental states: On the one hand, one must solve the particular problems of representing mental states *as such* (e.g., Perner, 1991); on the other hand, one must have a conceptual grasp of the nature of different kinds of mental states (e.g., intentions, beliefs, and knowledge) and their interrelations (e.g., Wellman, 1990). Children's well-known success on false-belief tasks at around 4 years of age (e.g., Wellman, Cross, & Watson, 2001) is typically taken as evidence of their ability to represent mental states as such—an ability that applies to intentions just as surely as it does to beliefs (Perner, 1991). However, children's conceptual understanding of intentions—of the motivational, causal, and epistemic components—begins before this age and continues afterward (Moses, 2001). By 3 years of age, children start to understand the role of beliefs in intentional action and to more accurately grasp the connections among desires, perceptions, and emotions (Tomasello et al., 2005). Nonetheless, 4- and 5-year-old children are unable to appreciate that, in order to pretend, an actor *must* have the intention to pretend (Lillard, 1998), and they struggle to distinguish between intentions and desires that relate to the same goal (Schult, 2002).

These results lend support to the possibility that the mentalizing system proposed on the basis of neuroimaging findings may be subdivided in a meaningful way, and thereby provide motivation for distinguishing between two components: a *representational* component concerned with the representation of intentions as mental states (rather than purely behavioral relations, as represented by the mirroring system); and a *conceptual* component representing the semantic and logical properties of intention, abstractly reasoning over these properties, and relating them to other mental states. In other words, the representational component of the mentalizing system makes it possible to represent intentions as mental states, and such representations are then realized to varying degrees of sophistication by the conceptual component.

A Composite Model of Intention Understanding

Combining these insights from neuroimaging and developmental research, recent studies of patients with acquired brain injuries have provided valuable evidence on the potential for dissociation between different components of mentalizing. We (Chiavarino, Apperly, & Humphreys, 2009) showed patients with frontal and parietal lesions videos of an actor performing real intentional actions (e.g., pouring juice into a glass), real accidental actions (e.g., spilling the juice from the glass), and pretended (intentional) actions (e.g., pretending to pour juice into the glass). In a first session, patients were asked to decide whether the actions depicted in the videos were real or pretended; in a second session, they were asked to decide whether the same actions were intentional or accidental.

Our first hypothesis was that the representational component of the mentalizing system would be largely sufficient to identify an act as pretense (i.e., to distinguish between real and pretended actions in the real/pretend task) and to differentiate between intentional and accidental actions (i.e., in the intentional/accidental task). The patients with parietal lesions were impaired in discriminating between real and pretended actions, whereas the patients with frontal lesions showed difficulties in differentiating between intentional and unintentional actions. Interestingly, lesions specifically encompassing the left TPJ affected both types of judgment. This suggests that the parietal and frontal lobes may both independently contribute to people's ability to infer others' mental states, but that the TPJ may play a particularly pivotal role in the representational component of the mentalizing system. This would be consistent with the finding that inferring others' short-term goals and intentions consistently activates the TPJ (Van Overwalle, 2009).

Our second hypothesis was that it may be possible to represent pretense or intentions without understanding all of their logical entailments. Evidence for this possibility came from three patients in the study who reliably distinguished between real and pretended actions and between intentional and accidental actions but nonetheless miscategorized pretended actions as "accidental" rather than "intentional." Notably, matched healthy participants were accurate in all of their judgments. Thus, it appears that an intact representational system is necessary, but not sufficient, for the conceptual system to function properly.

In a second study (Chiavarino, Apperly, & Humphreys, 2010), patients with frontal and parietal lesions viewed a series of pictures depicting a game in which a man wanted to find a target hidden in one of eight containers (his "desire," which could be satisfied or unsatisfied at the end of the game) and tried to bring about that result (his "intention," which could be independently fulfilled or unfulfilled) by hitting one of the containers from a distance. Thus, intention and desire were embedded in a means–end relation within the same action, and the satisfaction of the desire and the fulfillment of the intention did not covary. To respond correctly to questions concerning the intention and the desire of the man when the two mental states conflicted (e.g., when fulfillment of the intention did not satisfy the desire), one needs both the capacity to represent mental states as such and a conceptual grasp of the distinct semantic and logical properties of the two mental states.

All the patients performed well when the man's desire and intention were congruent. However, when the desire and intention were discrepant, the patients with left parietal damage—particularly those whose brain lesions encompassed the TPJ—were impaired in responding to questions about both the man's desire and his intention. This fits with our suggestion that the TPJ may be critically involved in the representational component of the mentalizing system, and that this component is a functional prerequisite for making complex, abstract inferences about mental states (Apperly, Samson, Chiavarino, & Humphreys, 2004). In contrast, patients with damage to left frontal regions made accurate judgments about the man's intentions, which suggests that they were able to represent intentions as mental states. However, when intentions and desires were discrepant, these patients showed impaired judgments about desires. The latter finding indicates that the frontal lobes play a role in managing the conflicts that arise when people must integrate information about more than one mental state to make a given judgment.

Taken together, these results substantiate the hypothesis that the mirroring system and the mentalizing system are functionally distinct and that, within the mentalizing system, the representational level may subserve the conceptual level (see Fig. 1). They further suggest that other brain areas, likely located within the frontal lobes, may assist the mentalizing system in coordinating information from different mental-state representations or solving conflicts among them.

Conclusions and Future Directions

In this article, we have contended that the widely held distinction between a mirroring and a mentalizing system for intention understanding is too simplistic, and that two distinct but interrelated levels—namely, a representational and a conceptual level—are necessary for characterizing the mentalizing system. For the sake of clarity, our focus in this article has been on intentions, but we believe the same lessons generalize to broader discussions about theory of mind. By integrating experimental evidence from neuroimaging, neuropsychology, and developmental literature, we have argued that the same

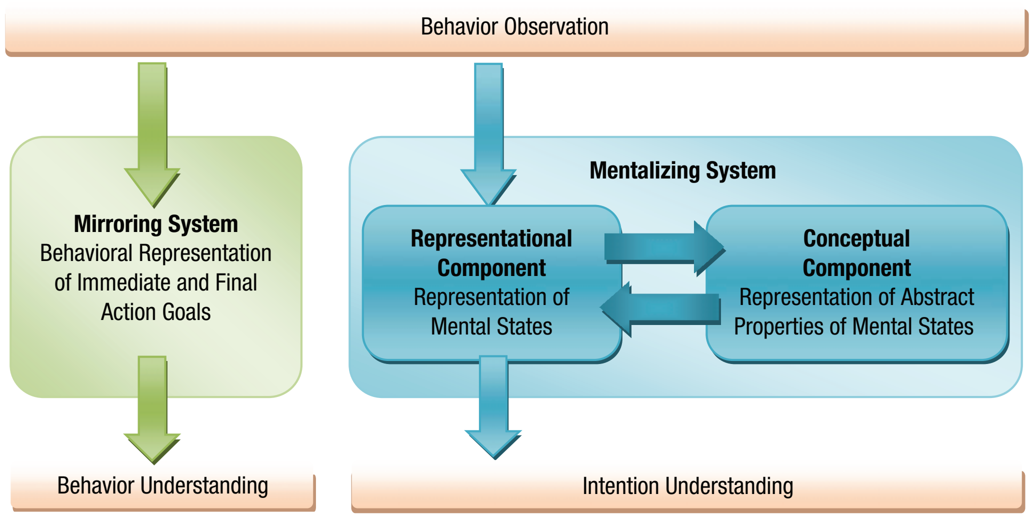

Fig. 1. The processes that lead from observing behavior to understanding intention. A first route for understanding, at a behavioral level, what another person is doing and how he or she is doing it is the mirroring system. Information from this mirroring system is not necessary for mentalizing to occur, even though the mirroring and mentalizing systems may in some cases jointly work to allow people to understand others' actions. A second, independent route, which allows understanding of intention as a mental (rather than as a merely behavioral) state, is mediated by the representational component of the mentalizing system. A candidate brain area that might implement this level of representation is the temporoparietal junction. Finally, a third component allows people to make abstract inferences about the semantic and logical properties of intentions. This route is based on the activity of the conceptual component of the mentalizing system, but it must build on an intact representational level. A frontoparietal brain network might implement this level of intention understanding.

observed behavior might be processed at different levels of complexity by distinct functional processes and differentiable neural networks within the frontal and parietal lobes. In particular, the relationship between the mirroring and mentalizing systems (globally considered) is likely to be functionally discontinuous. In contrast, within the mentalizing system, an intact representational level appears to be necessary for the conceptual level to function properly.

Future studies are needed to better qualify the anatomical organization of the mentalizing system. From the available data, it appears that the TPJ is a crucial node of the representational component of the system, which however seems to be assisted by a wider network of parietal and frontal areas. As for the conceptual system, it would be interesting for future research to investigate the role of two distinct frontal regions—first, the medial prefrontal cortex, which in lesion studies has been shown to be involved in the ability to relate intentions to knowledge states (Stuss, Gallup, & Alexander, 2001) and to outcomes (Young et al., 2010; but see the study by Bird, Castelli, Malik, Frith, & Husain, 2004, for contrasting results); and second, the bilateral inferior frontal gyrus, which both lesion and neuroimaging studies have indicated may be involved in solving specific situations of conflict (i.e., in which one's own perspective differs from that of the target person; e.g., Hartwright, Apperly, & Hansen, 2012; McCleery, Surtees, Graham, Richards, & Apperly, 2011; Samson, Apperly, Kathirgamanathan, & Humphreys, 2005; Saxe, Schulz, & Jiang, 2006; Vogeley et al., 2001). Existing accounts based on neuroimaging may underestimate the complexity of intention understanding in particular, and theory of mind in general, and will benefit from the consideration of data from both developmental and neuropsychological research.

Recommended Reading

Apperly, I. A., & Butterfill, S. A. (2009). Do humans have two systems to track beliefs and belief-like states? *Psychological Review, 116*, 953–970. A theoretical paper proposing that theory of mind might involve two distinct systems.

Kilner, J. M. (2011). More than one pathway to action understanding. *Trends in Cognitive Sciences, 15*, 352–357. A paper relating different brain areas with distinct levels of abstraction of goal understanding, both within and beyond the mirroring system.

Poulin-Dubois, D., Brooker, I., & Chow, V. (2009). The developmental origins of naïve psychology in infancy. *Advances in Child Development and Behavior, 37*, 55–104. A comprehensive review of the development of the understanding of intentions and other mental states.

Van Overwalle, F. (2009). (See References). Neuroimaging review discussing the role of different brain regions in social cognition.

Declaration of Conflicting Interests

The authors declared that they had no conflicts of interest with respect to their authorship or the publication of this article.

Funding

Claudia Chiavarino was supported by funding from the San Paolo Research Foundation (Neuroscience Programme: Action Representations and Their Impairment, 2010–2013).

Notes

1. There is growing agreement on the notion that all the inferences that are allowed by the mirroring system have to be viewed as low-level behavioral inferences, which do not involve representations of the mental states underlying the actions, but only anticipate actions' subsequent observable outcomes (e.g., grasp a coffee cup "to take it to the mouth"; Gallese, 2007; Grafton & Hamilton, 2007). Any time these inferences go beyond behavioral outcomes (e.g., "to stay awake" in the case of grasping a coffee cup), other brain systems involved in representing and reasoning about mental states are preferentially engaged (Spunt, Satpute, & Lieberman, 2010).
2. There is a rapidly emerging literature suggesting that infants have much more sophisticated theory-of-mind abilities than has traditionally been supposed (e.g., Baillargeon, Scott, & He, 2010). Importantly, this does not undermine the conceptual distinctions that we draw from the developmental literature in this article, given that it is widely agreed that infants' precocious mentalizing abilities cannot be explained by the mirroring system alone.

References

Apperly, I. A., Samson, D., Chiavarino, C., & Humphreys, G. W. (2004). Frontal and temporo-parietal lobe contributions to theory of mind: Neuropsychological evidence from a false-belief task with reduced language and executive demands. *Journal of Cognitive Neuroscience, 16*, 1773–1784.

Baillargeon, R., Scott, R. M., & He, Z. (2010). False-belief understanding in infants. *Trends in Cognitive Sciences, 14*, 110–118.

Bird, C. M., Castelli, F., Malik, O., Frith, U., & Husain, M. (2004). The impact of extensive medial frontal lobe damage on "theory of mind" and cognition. *Brain, 127*, 914–928.

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

Castelli, F., Happé, F., Frith, U., & Frith, C. (2000). Movement and mind: A functional imaging study of perception and interpretation of complex intentional movement patterns. *NeuroImage, 12*, 314–325.

Chiavarino, C., Apperly, I. A., & Humphreys, G. W. (2009). Frontal and parietal lobe involvement in the processing of pretence and intention. *Quarterly Journal of Experimental Psychology, 62*, 1738–1756.

Chiavarino, C., Apperly, I. A., & Humphreys, G. W. (2010). Distinguishing intentions from desires: Contributions of the frontal and parietal lobes. *Cognition, 117*, 203–216.

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.

Gallese, V. H. (2007). Before and below "theory of mind": Embodied simulation and the neural correlates of social cognition. *Philosophical Transactions of the Royal Society of London, Series B, 362*, 659–669.

Grafton, S. T., & Hamilton, A. F. (2007). Evidence for a distributed hierarchy of action representation in the brain. *Human Movement Science, 26*, 590–616.

Hartwright, C., Apperly, I. A., & Hansen, P. C. (2012). Multiple roles for executive control in belief-desire reasoning: Distinct neural networks are recruited for self perspective inhibition and complexity of reasoning. *NeuroImage, 61*, 921–930.

Keysers, C., & Gazzola, V. (2007). Integrating simulation and theory of mind: From self to social cognition. *Trends in Cognitive Sciences, 11*, 194–196.

Lillard, A. S. (1998). Wanting to be it: Children's understanding of intentions underlying pretense. *Child Development, 69*, 981–993.

Ma, L., & Lillard, A. S. (2006). Where is the real cheese? Young children's ability to discriminate between real and pretend acts. *Child Development, 77*, 1762–1777.

Mason, M. F., Magee, J. C., Kuwabara, K., & Nind, L. (2010). Specialization in relational reasoning: The efficiency, accuracy, and neural substrates of social versus non-social inferences. *Social Psychological and Personality Science, 1*, 318–326.

McCleery, J. P., Surtees, A., Graham, K. A., Richards, J., & Apperly, I. A. (2011). The neural and cognitive time-course of theory of mind. *Journal of Neuroscience, 31*, 12849–12854.

Moses, L. J. (2001). Some thoughts on ascribing complex intentional concepts to young children. In B. F. Malle, L. J. Moses, & D. A. Baldwin (Eds.), *Intentions and intentionality: Foundations of social cognition* (pp. 69–83). Cambridge, MA: MIT Press.

Perner, J. (1991). *Understanding the representational mind*. Cambridge, MA: MIT Press.

Repacholi, B. M., & Gopnik, A. (1997). Early reasoning about desires: Evidence from 14- and 18-month olds. *Developmental Psychology, 33*, 12–21.

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

Samson, D., Apperly, I. A., Kathirgamanathan, U., & Humphreys, G. W. (2005). Seeing it my way: A case of selective deficit in inhibiting self-perspective. *Brain, 128*, 1102–1111.

Saxe, R., Schulz, L., & Jiang, Y. (2006). Reading minds versus following rules: Dissociating theory of mind and EC in the brain. *Social Neuroscience, 1*, 284–298.

Schult, C. A. (2002). Children's understanding of the distinction between intentions and desires. *Child Development, 73*, 1727–1747.

Spunt, R. P., Satpute, A. B., & Lieberman, M. D. (2010). Identifying the what, why and how of an observed action: An fMRI study of mentalizing and mechanizing during action observation. *Journal of Cognitive Neuroscience, 23*, 63–74.

Stuss, D. T., Gallup, G. G., & Alexander, M. P. (2001). The frontal lobes are necessary for "theory of mind." *Brain, 124*, 279–286.

Tomasello, M., Carpenter, M., Call, J., Behne, T., & Moll, H. (2005). Understanding and sharing intentions: The origins of cultural cognition. *Behavioral & Brain Sciences, 28*, 675–735.

Van Overwalle, F. (2009). Social cognition and the brain: A meta-analysis. *Human Brain Mapping, 30*, 829–858.

Van Overwalle, F., & Baetens, K. (2009). Understanding others' actions and goals by mirror and mentalizing systems: A meta-analysis. *NeuroImage, 48*, 564–584.

Vogeley, K., Bussfeld, P., Newen, A., Hermann, S., Happé, F., Falkai, P., . . . Zilles, K. (2001). Mind reading: Neural mechanisms of theory of mind and self-perspective. *NeuroImage, 14*, 170–181.

Wellman, H. M. (1990). *The child's theory of mind*. Cambridge, MA: MIT Press.

Wellman, H. M., Cross, D., & Watson, J. (2001). Meta-analysis of theory of mind development: The truth about false belief. *Child Development, 72*, 655–684.

Young, L., Bechara, A., Tranel, D., Damasio, H., Hauser, M., & Damasio, A. (2010). Damage to ventromedial prefrontal cortex impairs judgment of harmful intent. *Neuron, 65*, 845–851.