The Social Nature of Perception and Action

Günther Knoblich and Natalie Sebanz

Rutgers University

ABSTRACT—*Humans engage in a wide range of social activities. Previous research has focused on the role of higher cognitive functions, such as mentalizing (the ability to infer others' mental states) and language processing, in social exchange. This article reviews recent studies on action perception and joint action suggesting that basic perception–action links are crucial for many social interactions. Mapping perceived actions to one's own action repertoire enables direct understanding of others' actions and supports action identification. Joint action relies on shared action representations and involves modeling of others' performance in relation to one's own. Taking the social nature of perception and action seriously not only contributes to the understanding of dedicated social processes but has the potential to create a new perspective on the individual mind and brain.*

KEYWORDS—*perception and action; action perception; joint action; social cognition; social neuroscience*

Cognitive scientists have long believed that perception and action are two servants of the mind, residing in opposite wings of the mental mansion. According to this view, perception delivers messages from the outside world to keep the mind informed, and action executes what the mind commands. Recent research in the cognitive sciences and neurosciences suggests that the mind is actually more like a kibbutz than a mansion: Perception, action, and cognition seem to form a collective community. Perception and action are intimately linked, and cognition is firmly grounded in both of them.

This new perspective has important implications for understanding the functional and brain mechanisms that support people's ability to interact with others. Individuals do not just infer intentions, emotions, and attitudes from others' behavior (Fiske, 1992); rather, researchers have postulated a more immediate way of social understanding and social interaction, based on the close link between perception and action. For example, when one observes another individual performing a particular action, this activates the representations in one's own action system that one uses to perform the observed action. Taking the social functions of perception and action seriously might help to better understand disorders of social functions, including autism and certain symptoms of schizophrenia such as delusions of control.

In this article, we discuss recent findings from two research domains that shed new light on the social nature of perception and action. Research on action perception demonstrates that individuals rely on their bodies and the action system moving their bodies to understand others' actions and to identify their own actions. Research on joint action has revealed how individuals share representations, predict each other's actions, and learn to jointly plan ahead.

BODILY AND MOTOR CONTRIBUTIONS TO ACTION PERCEPTION

Understanding

When people watch sports games like basketball or soccer, they often find their limbs twitching as though they are taking part in the game. This indicates that observing actions can directly activate the motor system. The common-coding hypothesis provides a functional principle that can explain such phenomena. According to this hypothesis, perceiving and performing an action results in the activation of the same representations—i.e., "common codes." Evidence for common coding has been found at the level of single neurons, the so-called mirror neurons. The same neurons in the premotor cortex of a macaque monkey fire both when the monkey grasps an object and when it observes an experimenter grasping the same object. The implication of this striking finding is that brain areas that were thought to be purely motor areas also support action perception. This creates a new perspective on how individuals make sense of others' actions. Rather than understanding observed actions by mapping them onto abstract concepts, people "relive" them by mapping them

---

Address correspondence to Günther Knoblich, Psychology Department, Rutgers University, Smith Hall, 101 Warren Street, Newark, NJ 07102; e-mail: knoblich@psychology.rutgers.edu.

onto their own action knowledge. This leads to an immediate recognition of the goals underlying observed actions. Numerous brain-imaging studies have demonstrated that the mirror system supports action understanding.

Identification

But how do observed actions get mapped onto action representations? According to the common-coding principle, the more similar an observed action is to the way the observer would perform the action, the higher the activation of action representations. This was recently demonstrated in a study by Calvo-Merino and colleagues (Calvo-Merino, Glaser, Grezes, Passingham, & Haggard, 2005), in which ballet dancers and capoeira dancers watched videos of ballet dancing and capoeira dancing. Activation of the mirror system was stronger when the dancers watched the type of dancing they were experts in. The mirror system should become even more strongly activated when persons perceive recordings of actions that they themselves have previously performed, such as when seeing videos of themselves dancing or hearing recordings of their own piano playing. The higher level of activation during the perception of self-produced actions might allow individuals to distinguish their own actions from those of others (Repp & Knoblich, 2004).

Several studies have shown that people are indeed able to distinguish between their own and others' actions even when there are no cues with respect to the actor's identity besides dynamic information about body movements or the effects they result in. For instance, when people see point-light displays of themselves and their best friend dancing, jumping, or boxing, they can identify themselves better than they can identify their friend (Loula, Prasad, Harber, & Shiffrar, 2005). If visual familiarity were the main factor, the opposite result should be observed, because one sees one's friend's movements much more often than one's own from a third-person perspective. In a similar vein, people can identify their own handwriting from a single moving dot (Knoblich & Flach, 2003).

One is also able to identify the results of one's earlier performed actions in the auditory domain. It was found that nonmusicians were able to identify their own clapping (Flach, Knoblich, & Prinz, 2004). Expert piano players could distinguish a recording of a piece they had played for the first time from recordings of the same piece performed by other pianists (Repp & Knoblich, 2004). Whereas experts use subtle timing deviations from the score, known as expressive timing, to identify their playing, nonmusicians rely on tempo and salient rhythmic idiosyncrasies to identify their clapping. Thus there is converging evidence that the similarity between observed actions and the way one would perform them oneself leads to a higher activation of common codes. This, in turn, allows people to identify their own previous actions.

Simulation

In addition to action understanding and action identification, the mirror system seems to support the prediction of future action outcomes (action simulation). In particular, matching observed actions to one's own action repertoire allows one to exploit mechanisms in the motor system that are normally used to predict the outcomes of one's own actions. This solution is more parsimonious than predicting the results of others' actions based on a separate perceptual-anticipation mechanism. In support of simulation, it was found that people observing someone throwing a dart could quite accurately predict where the dart would land. Importantly, the predictions of the landing position were most accurate when participants observed videos of themselves throwing the dart, although recording session and recognition session were at least 1 week apart. This makes it very unlikely that the higher accuracy for self was due to memories for the outcome of particular throws (Knoblich & Flach, 2003). The higher degree of similarity between a perceived action and the way one would perform it oneself led to the higher prediction accuracy for oneself than for others.

There is also reason to believe that proprioceptive signals (sensing position of body parts) and tactile signals from one's own body contribute to action simulation. A recent study showed that lacking these signals impairs action understanding (Bosbach, Cole, Prinz, & Knoblich, 2005). In this study, two de-afferented individuals were tested. De-afferentation refers to the loss of body sense due to a degeneration of all nerve fibers that normally transmit sensory information to the brain. The two individuals observed videos of an actor lifting a box. Prior to lifting the box, the actor had sometimes been told the correct weight of the box and had sometimes been deceived about its weight. Both individuals had difficulties telling whether the actor lifting the box had the right or wrong expectation about its weight. In contrast, healthy participants had no problems making these judgments. They could tell from the actor's body posture and body movements whether or not the actor had been deceived, suggesting that action simulation in healthy individuals is supported by peripheral bodily signals. The lack of peripheral bodily signals in the de-afferented patients resulted in faulty simulations.

ENGAGING IN JOINT ACTION WITH OTHERS

Acknowledging the close links between perception and action also has implications for theorizing about joint action—social interactions wherein two or more individuals coordinate their actions in space and time to bring about a change in the environment. Some examples are doing the dishes together, rowing a canoe together, or playing a piano duet. Joint action involves sharing action representations and coordinating one's actions with those of others to achieve common goals (Clark, 1996).

Shared Representations

While previous research has focused on the role of language and theory of mind for successful social interaction, more immediate interpersonal links may exist in the form of a common coding

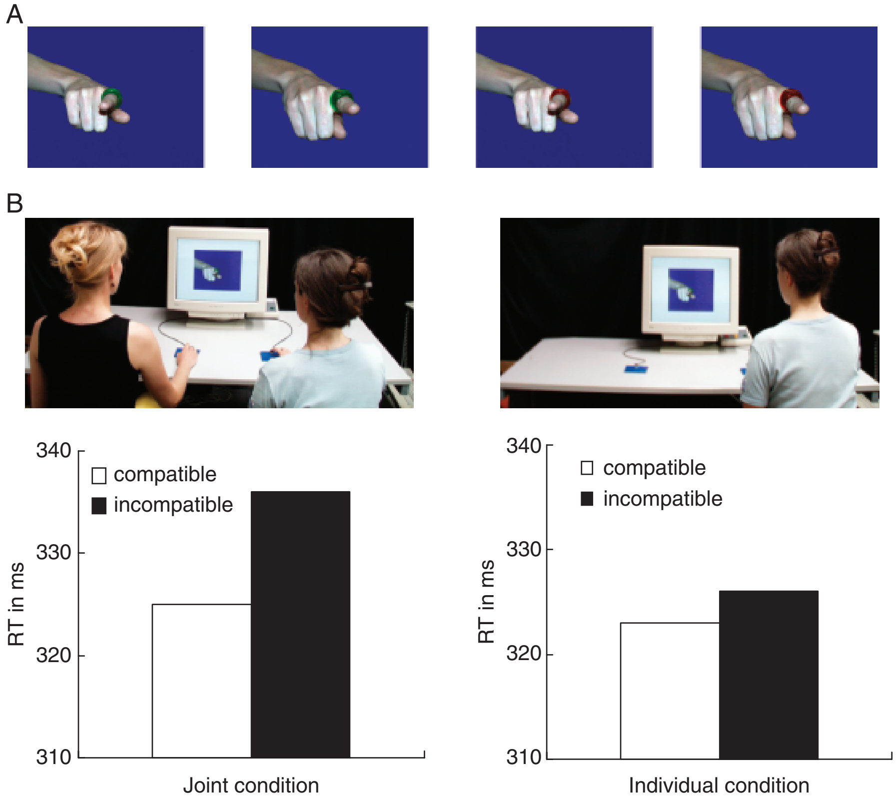

Fig. 1. Spatial-compatibility task used to test whether action representations are shared. Participants observed pictures of a hand pointing left or right (A), and were instructed to respond to the color of the ring on the index finger, ignoring the direction in which the finger was pointing. In the group condition (B, left panel), each participant responded to one color. In the individual condition (B, right panel) participants performed the same task alone, responding to one color and not responding to the other color. On compatible trials, the irrelevant pointing stimulus corresponded to the location of the required action. On incompatible trials, the pointing stimulus referred to the other's action (joint condition) or to no action (individual condition). Response times (RT) on incompatible trials in the joint condition (left graph) were slowed because the pointing stimulus activated a representation of the action at the other's command.

system for perception and action. If the actions one performs and actions one observes in others are represented in a functionally equivalent way, this would provide an optimal integration platform for performing tasks together. An implication that follows from this view is that sharing a task should be quite similar to performing it on one's own, at least when two complementary actions are distributed across two persons. In particular, the way actions are represented should not depend on whether one has all possible actions at one's own command or whether a part of the possible actions are at somebody else's command.

A simple spatial compatibility task was used to test this prediction (Sebanz, Knoblich, & Prinz, 2003). In the baseline condition, individual participants responded to a stimulus color with a left or right button press (e.g., red–left, green–right). Each stimulus also carried spatial information that had to be ignored (the stimulus, a colored ring, was on a finger that pointed either left or right, see Fig. 1A). A standard spatial-compatibility effect was observed: Participants responded slower when the irrelevant spatial dimension did not correspond to the location of the required action (e.g., red stimulus pointing right, when a left button press was required). Surprisingly, when the two action alternatives were distributed across two participants so that each participant performed a go/no-go task (responding to one stimulus and not responding to the other stimulus—e.g., Person A, red stimulus and left button press; Person B, green stimulus and right button press), the same pattern of results was observed (see Fig. 1B). However, there was no spatial-compatibility effect when individual participants performed the same go/no-go task alone (e.g., Person A, red stimulus and left button press; no Person B). Together, these results confirm that actions at another person's command are represented just as if they were at one's own command. Action representations are shared even if that leads to a decline in one's own performance. It seems that people cannot help representing what other people do.

Further evidence for the tendency to form shared action representations was obtained by measuring event-related potentials (ERPs; brain electrical activity following a stimulus) while participants performed our go/no-go task together and alone

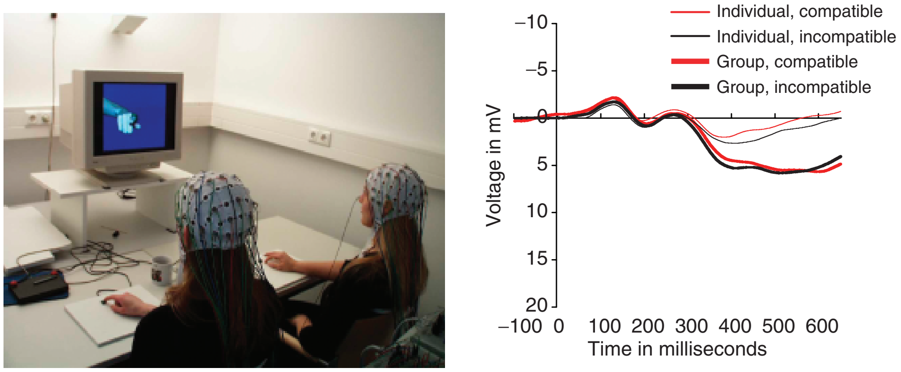

Fig. 2. Event-related potentials (brain electrical activity following a stimulus) were recorded while participants performed the spatial-compatibility task together (left panel) and alone; in both cases a participant responded to just one of the two color stimuli and not the other (no-go task). The right panel shows the amplitude of the no-go P3 component at a frontal electrode. This is a measure of how much inhibition is needed to suppress a response. The amplitude of the no-go P3 was significantly larger on no-go trials in the group condition, in which the other person acted, than it was in no-go trials in the individual condition, in which no one acted.

(Sebanz, Knoblich, Prinz, & Wascher, in press). This method allows one to determine what happens on trials in which participants need to refrain from responding because it is the other participant's turn (no-go trials). In particular, the amplitude of the no-go P3 component (a late-appearing positive voltage potential occurring roughly 400–600 milliseconds after the stimulus appears) indicates to what extent an action tendency needs to be inhibited. Analysis of this component revealed stronger inhibition when participants were required not to act because it was the other's turn than when they were required not to act but were alone (see Fig. 2). The additional need for inhibition when with another person further supports the assumption that one represents actions at the other's command in the same way as one's own actions. Representing the other's action leads to an action tendency on no-go trials that needs to be suppressed.

Surprisingly, the tendency to form shared action representations is present even in individuals who have difficulties understanding others' mental states. When high-functioning individuals with autism performed our go/no-go task either alone or together with a healthy participant, they showed the same tendency to form shared action representations that the healthy control participants did (Sebanz, Knoblich, Stumpf, & Prinz, 2005). This contrasts with the assumption of a common-coding (mirror-system) deficit in autism and raises the possibility that high-functioning individuals with autism have a specific deficit understanding that others' beliefs can differ from their own while more basic perception–action links supporting social interaction are intact.

Coordination

How do individuals adjust their actions to those of other people to achieve common goals (Clark, 1996)? Clearly, sharing action representations is not sufficient for successful interpersonal coordination. Whereas shared action representations allow individuals to simulate and predict others' actions, successful joint action often requires choosing appropriate complementary actions at an appropriate time. Take rowing a canoe as an example: When rowing a canoe alone, one can coordinate the timing of left and right paddle strokes quite easily because both actions are at one's own command. In contrast, two people rowing a canoe together must adjust to each other, because each partner has just the left or the right paddle strokes at his or her command. Knowing whether the other performs left or right paddle strokes will not be sufficient for successful coordination. Instead, to avoid going around in circles, each rower also needs to attend to the timing of the other rower's strokes when timing his or her own. The coordination will be smoothest if each rower acquires an internal model that allows predicting the timing of the other rower's actions in relation to his or her own.

Knoblich and Jordan (2003) investigated the mechanisms underlying such anticipatory coordination, using a simple computer game that posed coordination challenges similar to rowing a canoe together. The main question was whether receiving an unambiguous signal about the timing of the other person's action would allow two individuals to achieve the same degree of coordination as a single individual playing the game alone. Although, initially, group performance was much worse than individual performance, groups who received timing feedback gradually became as effective as individuals in coordinating the two actions. Groups that were only given information about the joint outcome of their actions never reached the level of individual performance. Feedback led partners to develop a model of each other, allowing them to anticipate each other's action timing.

These results seem to suggest that one has to start from scratch when modeling the actions of unfamiliar people. However, the action-simulation account described in the last section opens up an alternative: Initial predictions about others' actions could take one's own action parameters as the default. In other words, one initially assumes that other people act just like one would act oneself if one were to perform their actions. In the course of joint action, this initial model is adjusted to match others' actual performance. This implies that coordination should be best when coordinating with people whose performance is very similar to one's own. Thus one should be one's own optimal partner, for instance when playing a piano duet. This idea was tested in a study in which expert pianists played one part of a piano duet in synchrony with a recording of the other part that they themselves had played or that had been played by another pianist (Keller, Knoblich, & Repp, in press). The temporal synchronization error between the two parts of the duet was significantly smaller when pianists synchronized with their own playing. This supports the assumption that predictions about others' action timing are initially based on one's own simulated performance.

CONCLUDING THOUGHTS

Our review of some new findings on action perception and joint action suggests that basic perception–action links are crucial building blocks for social understanding and social interaction. It seems that a comprehensive understanding of social interaction can only be achieved if we continue to investigate how "lower-level" processes related to action understanding and action coordination enable and complement "higher-level" functions involved in thinking about and communicating with others (Smith & Semin, 2004).

What else can be learned from the finding that perception and action are social in nature? The obvious conclusion is that specific perceptual, cognitive, and motor processes are dedicated to social interaction. This seems to be the currently dominant view in the new field of social cognitive neuroscience. This approach has led to important progress, because it is now possible to map social behavior to particular cognitive and brain functions. However, we suggest that a deeper understanding of the processes supporting social interaction might be achieved if one takes the more radical stance that the demands of social interaction have shaped perception, action, and cognition (Fiske, 1992) through and through (Smith & Semin, 2004).

In particular, reassessing perception, action planning, and motor control in the light of their potential social roots might reveal that functions traditionally considered hallmarks of individual cognition originated through the need to interact with others. For instance, humans' ability to perform two tasks at the same time could be supported by processes that originally enabled individuals to perform one task while monitoring another individual's task performance. Along these lines, Roepstorff and Frith (2004) have speculated that the homunculus who has plagued psychology from its beginnings might be exorcized through a social exegesis. The hidden controller of our actions might be nothing more than an internalized other giving commands. Further exploration of how perception, action, and cognition are grounded in social interaction might have the potential to turn social cognitive neuroscience into a coherent framework that is more than the sum of its parts.

Recommended Reading

Barsalou, L.W., Simmons, W.K., Barbey, A., & Wilson, C.D. (2003). Grounding conceptual knowledge in modality-specific systems. *Trends in Cognitive Sciences*, *7*, 84–91.

Gallese, V., Keysers, C., & Rizzolatti, G. (2004). A unifying view of the basis of social cognition. *Trends in Cognitive Sciences*, *8*, 396–403.

Jackson, P.L., & Decety, J. (2004). Motor cognition: A new paradigm to investigate social interactions. *Current Opinion in Neurobiology*, *14*, 259–263.

Sebanz, N., Bekkering, H., & Knoblich, G. (2006). Joint action: Bodies and minds moving together. *Trends in Cognitive Sciences*, *10*, 70–76.

Wilson, M., & Knoblich, G. (2005). The case for motor involvement in perceiving conspecifics. *Psychological Bulletin*, *131*, 460–473.

Acknowledgments—This work was supported by the research initiative "Embodied Communication" at the Center of Interdisciplinary Research (ZIF) at the University of Bielefeld, Germany.

REFERENCES

Bosbach, S., Cole, J., Prinz, W., & Knoblich, G. (2005). Understanding another's expectation from action: The role of peripheral sensation. *Nature Neuroscience*, *8*, 1295–1297.

Calvo-Merino, B., Glaser, D.E., Grezes, J., Passingham, R.E., & Haggard, P. (2005). Action observation and acquired motor skills. *Cerebral Cortex*, *15*, 1243–1249.

Clark, H. (1996). *Using language*. Cambridge, England: Cambridge University Press.

Fiske, S.T. (1992). Thinking is for doing: Portraits of social cognition from daguerrotype to laserphoto. *Journal of Personality and Social Psychology*, *63*, 877–889.

Flach, R., Knoblich, G., & Prinz, W. (2004). Recognizing one's own clapping: The role of temporal cues in self-recognition. *Psychological Research*, *11*, 147–156.

Keller, P., Knoblich, G., & Repp, B.H. (in press). Pianists duet better when they play with themselves. *Consciousness & Cognition*.

Knoblich, G., & Flach, R. (2003). Action identity: Evidence from self-recognition, prediction, and coordination. *Consciousness and Cognition*, *12*, 620–632.

Knoblich, G., & Jordan, S. (2003). Action coordination in individuals and groups: Learning anticipatory control. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *29*, 1006–1016.

Loula, F., Prasad, S., Harber, K., & Shiffrar, M. (2005). Recognizing people from their movement. *Journal of Experimental Psychology: Human Perception and Performance*, *31*, 210–220.

Repp, B.H., & Knoblich, G. (2004). Perceiving action identity: How pianists recognize their own performances. *Psychological Science*, *15*, 604–609.

Roepstorff, A., & Frith, C. (2004). What's at the top in the top-down control of action? *Psychological Research*, *68*, 189–198.

Sebanz, N., Knoblich, G., & Prinz, W. (2003). Representing others' actions: Just like one's own? *Cognition*, *88*, B11–B21.

Sebanz, N., Knoblich, G., Prinz, W., & Wascher, E. (in press). Twin Peaks: An ERP study of action planning and control in co-acting individuals. *Journal of Cognitive Neuroscience*.

Sebanz, N., Knoblich, G., Stumpf, L., & Prinz, W. (2005). Far from action blind: Representation of others' actions in individuals with autism. *Cognitive Neuropsychology*, *22*, 433–454.

Smith, E.R., & Semin, G.R. (2004). Socially situated cognition: Cognition in its social context. *Advances in Experimental Social Psychology*, *36*, 53–117.