Trends in Cognitive Sciences

The battle for Broca's region

State-dependency in brain stimulation studies
Perceiving personalities in faces
Event comprehension by primates

Cell PRESS

Cell
PRESS

Action comprehension in non-human primates: motor simulation or inferential reasoning?

Justin N. Wood[1] and Marc D. Hauser[2]

[1] University of Southern California, Department of Psychology, 3620 South McClintock Ave, Los Angeles, CA 90089, USA
[2] Harvard University, Department of Psychology, 33 Kirkland Street, Cambridge, MA 02138, USA

Some argue that action comprehension is intimately connected with the observer's own motor capacities, whereas others argue that action comprehension depends on non-motor inferential mechanisms. We address this debate by reviewing comparative studies that license four conclusions: monkeys and apes extract the meaning of an action (i) by going beyond the surface properties of actions, attributing goals and intentions to the agent; (ii) by using environmental information to infer when actions are rational; (iii) by making predictions about an agent's goal, and the most probable action to obtain the goal given environmental constraints; (iv) in situations in which they are physiologically incapable of producing the actions. Motor theories are, thus, insufficient to account for primate action comprehension in the absence of inferential mechanisms.

Social animals wake up every morning to a dizzying array of actions from allies and enemies, in contexts that include cooperation, coalitionary violence, inter-group territorial defense, resource competition and parental care. Some actions are intentional, motivated by either distal or proximal goals; some are accidental, but nonetheless result in similar consequences. How do individuals distinguish between intentional and accidental actions? A deep understanding of the representational systems that support this capacity requires answering two primary questions, one mechanistic and one phylogenetic. Mechanistically, what is entailed in the process of perceiving an action and subsequently, extracting the relevant details to make inferences about the actor's mental states? Phylogenetically, are the mechanisms recruited for action comprehension uniquely human or shared with other species?

Over the past decade, behavioral, psychological and neurobiological research on humans and other animals has begun to address these two questions. Studies of human infants show that core components of action comprehension are present and functional early in development [1–6], and studies of non-human primates show that some of these same components are shared with closely and distantly related animal species [7–12] (see Ref. [13] for a recent review). Furthermore, single unit physiology

studies of macaques, together with neuroimaging studies of humans, reveal that there are dedicated cortical systems that activate during action observation [14–18], including some processes that link motor production with perception [14,15], whereas other processes underpin inferential mechanisms linked to mental states [17,18].

Despite the considerable progress to date, several questions remain about the processes underlying action comprehension [19,20]. Two broad competing theoretical perspectives are currently in play. Motor theorists posit a crucial link between action comprehension and action production, such that organisms use their own motor system as a guide for understanding the actions of others [15]. Evidence for this perspective comes from studies showing that common neural substrates, in the premotor and posterior parietal cortex, activate both when an agent performs an action and when they observe, or hear, another individual perform similar actions [15,16,21]. Based on these common activation patterns between action perception and production, researchers argue that action perception consists of a direct-matching mechanism in which observed actions are mapped onto one's own premotor system. This mapping causes the motor system to 'resonate,' thereby allowing the observer to understand the goal of the observed action as if she were performing the action herself [15]. It is often claimed that the mirror neuron system is intimately connected with social cognition, such that motor representations of observed actions are necessary [22] or sufficient [15] for understanding another individual's goals.

The primary alternative to the motor simulation perspective is a class of inferential theories, which argue that action comprehension involves an interpretative inferential mechanism that analyzes various visual characteristics of an action [23]. Different kinds of inferential models have been proposed (see Ref. [24] for a discussion of the similarities and differences between existing theories]. Because of space constraints, we focus on one type of inferential model – the teleological theory – which has received considerable attention across a range of disciples in cognitive science, including developmental psychology, animal cognition, cognitive neuroscience and philosophy. Teleological models propose that an observer assigns a goal to an action by evaluating the efficiency of the action with respect to the current environment, and

Corresponding author: Wood, J.N. (justin.wood@usc.edu).

specifically, the constraints it imposes on goal attainment [23]. Although teleological theories do not deny that motor representations might have some role in action perception, they argue that motor representations of observed actions are neither necessary nor sufficient for action comprehension; some support for this position comes from behavioral studies showing that human infants properly interpret the goal of a 2D animated ball [5], even though they clearly cannot move in this way.

There are at least two ways in which studies of non-human primate behavior can provide unique inroads into this debate, both theoretically and methodologically. First, some of the strongest evidence in favor of motor simulation comes from physiological studies with macaque monkeys, in which individual 'mirror neurons' have been discovered that activate both during action production and action perception [15]. Despite the wealth of physiological evidence from cellular recordings, there is little research examining how macaques interpret the of actions on the behavioral level [7–9,25]. Without accompanying behavioral data, it is difficult to characterize the psychological computations that mirror neurons might support under naturally occurring social interactions. For instance, the activation of mirror neurons during action perception could mean several things. On one hand, the activation might reflect processes that are necessary or sufficient for computing the intentional state of an actor [15]; if so, then macaques' ability to make inferences about the mental states of others should depend, at least to some extent, on whether the observed actions are within the observer's own motor repertoire. On the other hand, the activation of the motor system during action observation might reflect processes that are not directly related to mental state attribution [20]; if so, then macaques' ability to infer the mental states of others should be independent of whether they can perform the actions themselves. By testing how non-human primates respond to actions across various social contexts, we can ask whether their behavior accords most closely with empirical predictions made by models of motor simulation, inferential mechanisms or some combination of the two.

Second, with studies of non-human animals, it is possible to control an individual's motor experience (e.g. by teaching them how to perform new actions), use species that have restricted motor capacities (e.g. primates vary considerably in their dexterity) and modify visual experience (e.g. by restricting their visual experience of seeing others perform actions). This type of experiential control is important for distinguishing between motor and teleological theories of action comprehension because an individual that develops in a normal environment will both have motor knowledge of their own actions and visual experience seeing others act in motorically familiar ways. By systematically controlling an animal's motor or visual knowledge, it becomes possible to determine each of their respective roles in action comprehension.

We next discuss empirical work that bears directly on this debate. In the first section, we review evidence indicating that exact motor representations of observed actions are unnecessary for action comprehension. In the second section, we review evidence indicating that motor representations of observed actions are insufficient for action comprehension. We conclude that a motor simulation mechanism that operates in the absence of an inferential mechanism cannot provide a complete account of the socio-cognitive processes guiding non-human primate action comprehension.

Are exact motor representations of observed actions necessary for action comprehension?

Motor and teleological theories make different predictions about the range of actions that can be successfully interpreted by an organism. Motor theories argue that observers use their own motor representations to comprehend the meaning of the actions of another, thereby predicting that action comprehension should be limited to actions that are within the observer's motor repertoire. By contrast, teleological theories argue that action comprehension involves analyzing various visual representations of an event [23]. Thus, although action comprehension might use motor representations, it is not reliant on this kind of information.

The adult rhesus monkeys living on the island of Cayo Santiago have observed humans throw objects overhand. However, these animals never throw objects, and nor would they be capable of doing so given the lack of appropriate musculature. A recent study [26] explored whether rhesus understand and predict the consequences of throwing even though they cannot throw. An experimenter approached lone subjects and showed them that he had a rock in one hand. He then performed an overhand throw towards the subject, without releasing the rock. The dependent measure was whether the subject moved from its current location, an adaptive response and measure of avoidance in the face of a potential threat.

Eighty-five percent of the subjects moved away after observing the overhand throw, indicating that they successfully predicted the outcome of the action and interpreted it as a potential threat. Rhesus were then tested with a wide variety of different types of throws to examine whether this behavior is best explained through an inflexible learning mechanism that established an association between an observed movement pattern (the overhand throw) and a negative consequence (being hit by a rock), or through a more sophisticated mechanism that generalizes previous experience of observed throws to novel throws according to the most relevant properties of the action. When the experimenter performed a novel underhand throwing motion that rhesus had little to no experience observing, they nevertheless showed identical levels of avoidance compared to the overhand throw. By contrast, rhesus showed less avoidance when the experimenter performed a throwing action that lacked all of the kinematic components of an overhand throw (i.e. moving arm backward, rotating shoulder forward, extending forearm). Similarly, rhesus showed less avoidance when the throw was performed with an empty hand or soft food object rather than a rock, at a slower speed (i.e. overhand throw was performed at one-third of normal speed), towards a different direction from the subject (i.e. experimenter looked at subject but performed the overhand throw 90°

away from subject) or with a trajectory that could not produce sufficient torque to serve as a threat (i.e. experimenter moved arm straight towards subject). Crucially, rhesus also showed less avoidance when the experimenter performed the overhand throw but directed their eye gaze 90° in a different direction. This indicates that their avoidance behavior is based both on a biomechanical analysis of the throw and on their understanding of the actor's attention and, thus, presumably their goal [26]. Rhesus did not simply respond to learned associations between familiar throwing motions and negative consequences because many of these throws were identical in motion, but nevertheless yielded different behavioral responses. Rather, these results provide evidence that rhesus evaluate throwing actions by parsing them into at least three relevant components: the kinematics of the motor action, the observer's attention and the object held. This indicates that although rhesus monkeys lack exact motor representations of throwing actions, they are quite sophisticated in their ability to comprehend and predict the outcome of a throwing action when it is performed by a human. Thus, when non-human animals assess the mental states of others, exact motor representations of the observed actions are unnecessary for action comprehension.

It is important to note that this study cannot definitively rule out the possibility that motor representations have some role in the comprehension of throwing, or other motor actions that fall outside their repertoire. Mirror neurons in the premotor cortex are both 'strictly congruent' (i.e. they activate to observed actions that are the same as the corresponding motor action) and 'broadly congruent' (i.e. they can activate to observed actions that are different from the corresponding motor action provided that the action accomplishes the same goal), raising the possibility that broadly congruent neurons activate when rhesus observe a human throw. Indeed, after extensive observation, some mirror neurons activate for observed actions that the monkey cannot perform themselves [27]. Future studies might, therefore, explore whether broadly congruent motor representations are causally related to action comprehension in situations in which organisms lack exact motor representations of the actions.

Are motor representations of actions sufficient for action comprehension?

As part of the suite of inferential theories, a teleological mechanism assigns a goal to an action by evaluating the efficiency of the agent's action with respect to environmental constraints on goal attainment [23]. Thus, under this model, an analysis of an action will involve information about a potential goal, the observed action as an optimal means of obtaining that goal, and the environmental constraints facing the agent at the moment of action.

Recent studies of free-ranging rhesus monkeys explored whether these kinds of processes underlie action comprehension [7–9]. An experimenter presented two potential food sources (overturned coconut shells) to a subject, acted on one, and then walked away, allowing the subject to selectively approach. Although coconuts are native to the island on which these animals live, rhesus cannot open the hard

outer shells themselves and, therefore, only obtain the desired inner fruit when the coconuts have been opened and discarded by a human. It, thus, logically follows that if subjects perceive the experimenter's action as goal-directed, then they should selectively approach the coconut contacted as this maximizes the odds of obtaining food.

Rhesus selectively approached the targeted coconut when the experimenter grasped the coconut with his hand, foot or with a precision grip involving the pointer finger and thumb; by contrast, they approached the two coconuts at chance levels when the experimenter flopped the back of his hand on the coconut, touched or grasped the coconut with a tool or grasped the container with his hand for balance while standing up. These results rule out low-level association accounts, and show that when assessing the meaning of actions, rhesus are highly sensitive to the action means used to achieve a goal – for example, perceiving a hand grasp action as goal-directed but a hand flop action as accidental, despite the fact that the experimenter's body position, eye gaze and duration of contact with the coconut were identical across the two conditions.

Subsequent studies [9] explored whether rhesus assess the meaning of actions by evaluating whether an action is 'rational' given the constraints of the situation. We use the notion of rational here in the same way that it has been discussed by Gergely and Csibra [23,28] in their developmental studies of human infants, specifically, to describe an interpretation of actions as attentive to the agent's goals, and the means by which these goals can be achieved. Using the two-option approach measure discussed earlier, subjects observed an experimenter perform the same elbow touch action under two contrasting environmental circumstances: in one condition, the experimenter's acting hand was occupied by holding an object, and in the second condition, the experimenter's acting hand was free.

Rhesus used the elbow touch as a cue to find the hidden food when the experimenter's acting hand was occupied, but not when it was free. These results again rule out low-level association accounts: if rhesus restricted their attention to the locus of contact or attention, then they should have shown the same behavior across conditions because the experimenter attended to and made contact with a coconut in both conditions. Instead, rhesus took into account current environmental constraints on goal-directed action, perceiving the elbow touch as a rational alternative when the acting hand was occupied, but not when it was free. An identical pattern of results was found with cotton-top tamarins and chimpanzees [9], and other laboratories, using both rhesus monkeys and other mammals, have obtained convergent evidence in several species using a variety of methods, ranging from violation of expectancy looking-time measures [12] to selective imitation [29,30]. Together, these studies show that non-human animals infer the meaning of an action by evaluating action means in relation to the environmental constraints imposed on the agent in relation to a potential goal state. Action comprehension can, therefore, not be based solely on a direct-matching motor simulation mechanism in which the observed action means (initially 'devoid of meaning' with regards to the goal of the action [15]) is sufficient as the sole input.

Interpreting the primate data in a broader comparative context

A wealth of studies show that the pre-motor cortex activates during action observation. How do the results discussed here fit within this broader literature? We make two general points. First, our findings converge with cellular recordings of mirror neurons: as Csibra [20] points out, the pattern of mirror neuron activation might be better understood by appealing to a teleological mechanism as opposed to a direct-matching mechanism. Teleological theories predict that observers judge an action to be goal-directed only if the action is an optimal means to obtain the goal within the context of the situation. Likewise, mirror neuron activation is heavily influenced by the broader context of the situation: they fail to activate to intransitive gestures [14], they activate to a hand grasp gesture only if there is a visible (or hidden) goal object [31] and they activate differently to the same action depending on the situational context [32]. In these studies, the context and the action means (as opposed to solely the action means) carried information about the actor's goal. Thus, although these results are often cited as evidence that mirror neurons compute the goal of an action, they also fit perfectly with the idea that some other process – a teleological mechanism – computes whether an action is goal-directed, with mirror neuron activation reflecting this process. Mirror neuron activation seems to be conditional on action understanding and not the other way around [20].

Second, action comprehension occurs on different levels of analysis: 'what' action was performed (e.g. 'grasping an apple'), 'how' the action was performed (e.g. 'with a whole hand prehension'), and 'why' the action was performed (e.g. 'because the actor wanted the apple'), and motor representations might have greater computational roles for some levels than others [19]. There is some neurological support for this: when individuals reflect on why an individual performed an action, non-motor regions of the brain involved in interpretative inferential processes – the medial prefrontal cortex, posterior cingulate cortex and right posterior superior temporal sulcus – show enhanced activation. By contrast, although the activation patterns of the motor system indicate that it represents information about what action was performed and how the action was performed, it does not seem to represent information about inferred mental states of agents, such as why the action was performed [17,18]. Thus, when debating whether a particular process is involved in action comprehension, it is important to consider not only the type of process but also the level of action analysis.

Conclusions

An important question in the cognitive sciences concerns the cognitive and neural processes that are used to understand the intentions of others in our social world. Studies with non-human animals can help resolve several ongoing controversies by examining which components of a general capacity are necessary and which sufficient, and by revealing how different mechanisms (e.g. the capacity to throw and the psychological mechanisms needed to understand and predict the consequences of throwing actions) can dissociate over the course of evolution. Furthermore, studies with

---

Box 1. Outstanding questions

- Although exact motor representations appear unnecessary for action comprehension, they might nevertheless have some role in action comprehension. Is action comprehension 'enhanced' in some way when an observer has an exact motor representation of an observed action?
- Humans and non-human animals infer an agent's goal by evaluating action means in relation to environmental constraints on goal attainment. But what part of the action means, specifically, is evaluated during this process? Does the analysis focus solely on the acting body part? Or does it evaluate a wide range of information, such as non-acting body parts, and the differential capacities of particular agents (e.g. skilled versus unskilled individuals)?
- What kind of the information about the broader context is evaluated during action comprehension? Does the analysis focus on what the actor could do in the future? Or is it restricted to the current and immediate constraints on the agent?

---

animals enable precise control over various types of knowledge, allowing for systematic investigations into the role of various processes in social cognition (Box 1). Along these lines, the studies reviewed here provide evidence for four conclusions: monkeys and apes make inferences about the mental states of others (i) by going beyond the surface appearance of actions, attributing goals and intentions to the agent; (ii) by using details about the environment to infer when an action is rational or irrational; (iii) by making predictions about an agent's goals, and the most probable action to obtain a goal, within the constraints of the situation and, (iv) in situations in which they lack accompanying motor knowledge of the actions. This indicates that a motor simulation mechanism that operates in the absence of a teleological mechanism cannot provide a complete account of the socio-cognitive processes guiding non-human primate action comprehension.

References

1 Woodward, A.L. (1998) Infants selectively encode the goal object of an actor's reach. *Cognition* 69, 1–34
2 Woodward, A.L. (1999) Infants' ability to distinguish between purposeful and non-purposeful behaviors. *Infant Behav. Dev.* 17, 515–521
3 Woodward, A.L. and Sommerville, J.A. (2000) Twelve-month-old infants interpret action in context. *Psychol. Sci.* 11, 73–77
4 Spaepen, E. and Spelke, E. (2007) Will any doll do? 12-month-olds' reasoning about goal objects. *Cognit. Psychol.* 54, 133–154
5 Gergely, G. *et al.* (1995) Taking the intentional stance at 12 months of age. *Cognition* 56, 165–193
6 Call, J. *et al.* (2004) Unwilling or unable? Chimpanzees' understanding of intentional action. *Dev. Sci.* 7, 488–498
7 Hauser, M.D. *et al.* (2007) Wild, untrained and non-enculturated rhesus monkeys correctly read the goal-relevant gestures of a human agent. *Proc. Biol. Sci.* 274, 1913–1918
8 Wood, J.N. *et al.* (2008) Rhesus monkeys' understanding of actions and goals. *Soc. Neurosci.* 3, 60–68
9 Wood, J.N. *et al.* (2007) The perception of rational, goal-directed action in non-human primates. *Science* 317, 1402–1405
10 Premack, D. and Woodruff, G. (1978) Does the chimpanzee have a theory of mind? *Behav. Brain Sci.* 1, 515–526
11 Byrne, R.W. and Whiten, A. (1990) Tactical deception in primates: the 1990 database. *Primate Rep.* 27, 1–101
12 Rochat, M.J. *et al.* (2008) The evolution of social cognition: goal familiarity shapes monkeys' action understanding. *Curr. Biol.* 18, 227–232
13 Call, J. and Tomasello, M. (2008) Does the chimpanzee have a theory of mind? 30 years later. *Trends Cogn. Sci.* 12, 187–192

14 Gallese, V. *et al.* (1996) Action recognition in the premotor cortex. *Brain* 119, 593–609

15 Rizzolatti, G. *et al.* (2001) Neurophysiological mechanisms underlying the understanding and imitation of action. *Nat. Rev. Neurosci.* 2, 661–670

16 Perrett, D.I. *et al.* (1990) Understanding the visual appearance and consequences of hand actions. In *Vision and Action: The Control of Grasping* (Goodale, M.A., ed.), pp. 163–342, Ablex

17 de Lange, F.P. *et al.* (2008) Complementary systems for understanding action intentions. *Curr. Biol.* 18, 454–457

18 Brass, M. *et al.* (2007) Investigating action understanding: inferential processes versus action simulation. *Curr. Biol.* 17, 2117–2121

19 Thioux, M. *et al.* (2008) Action understanding how, what, and why. *Curr. Biol.* 18, R431–R434

20 Csibra, G. (2007) Action mirroring and action interpretation: an alternative account. In *Sensorimotor Foundations of Higher, Cognition. Attention and Performance XXII* (Haggard, P. *et al.*, eds), pp. 435–459, Oxford University Press

21 Kohler, E. *et al.* (2002) Hearing sounds, understanding actions: action representation in mirror neurons. *Science* 297, 846–848

22 Nelissen, K. *et al.* (2005) Observing others: multiple action representation in the frontal lobe. *Science* 310, 332–336

23 Gergely, G. and Csibra, G. (2003) Teleological reasoning in infancy: the naïve theory of rational action. *Trends Cogn. Sci.* 7, 287–292

24 Csibra, G. and Gergely, G. (2007) 'Obsessed with goals': functions and mechanisms of teleological interpretation of actions in humans. *Acta Psychol. (Amst.)* 124, 60–78

25 Lyons, D.E. *et al.* (2006) Reflections of other minds: how primate social cognition can inform the function of mirror neurons. *Curr. Opin. Neurobiol.* 16, 230–234

26 Wood, J.N. *et al.* (2007) The uniquely human capacity to throw evolved from a non-throwing primate: an evolutionary dissociation between action and perception. *Biol. Lett.* 3, 360–364

27 Ferrari, P.F. *et al.* (2005) Mirror neurons responding to the observation of actions made with tools in the monkey ventral premotor cortex. *J. Cogn. Neurosci.* 17, 212–226

28 Gergely, G. *et al.* (2002) Rational imitation in preverbal infants. *Nature* 415, 755

29 Buttelmann, D. *et al.* (2007) Enculturated chimpanzees imitate rationally. *Dev. Sci.* 10, F31–F38

30 Range, F. *et al.* (2007) Selective imitation in domestic dogs. *Curr. Biol.* 17, 868–872

31 Umiltà, M.A. *et al.* (2001) I know what you are doing: a neurophysiological study. *Neuron* 31, 155–165

32 Fogassi, L. *et al.* (2005) Parietal lobe: from action organization to intention understanding. *Science* 308, 662–667