

Editorial Note:

The following article by Susan Hurley, “The Shared Circuits Model: How Control, Mirroring, and Simulation Can Enable Imitation, Deliberation, and Mindreading,” with its commentaries and response was produced under unusual and sad circumstances. Susan Hurley passed away in August 2007 following a long struggle with cancer after her target article had been completed, and the list of those invited to comment had been assembled. Because she had foreseen the need for help in producing her response to the commentaries, she enlisted Andy Clark, Professor at Edinburgh University, for this purpose, with BBS’s full encouragement. Julian Kiverstein, another colleague at the University of Edinburgh with particular interest in the shared circuits model, volunteered to help as well in the composition of the response to commentators.

Commentators were specifically enjoined from writing eulogies and asked to produce the lively intellectual dialogue that Susan Hurley certainly had sought in sending her work to BBS. Kiverstein and Clark undertook not to emulate a response from Susan Hurley, but rather to clarify misunderstandings, organize the commentaries thematically, and show where the research might lead. We are grateful to all commentators, and particularly Kiverstein and Clark, for their graceful execution of what even in the normal case is a challenging task.

The shared circuits model (SCM): How control, mirroring, and simulation can enable imitation, deliberation, and mindreading

Susan Hurley

Department of Philosophy, University of Bristol, Bristol BS8 1TB, and All Souls College, Oxford University, Oxford OX1 4AL, United Kingdom

Susan.hurley@bristol.ac.uk

<http://eis.bris.ac.uk/~plsh/>

Abstract: Imitation, deliberation, and mindreading are characteristically human sociocognitive skills. Research on imitation and its role in social cognition is flourishing across various disciplines. Imitation is surveyed in this target article under headings of behavior, subpersonal mechanisms, and functions of imitation. A model is then advanced within which many of the developments surveyed can be located and explained. The shared circuits model (SCM) explains how imitation, deliberation, and mindreading can be enabled by subpersonal mechanisms of control, mirroring, and simulation. It is cast at a middle, functional level of description, that is, between the level of neural implementation and the level of conscious perceptions and intentional actions. The SCM connects shared informational dynamics for perception and action with shared informational dynamics for self and other, while also showing how the action/perception, self/other, and actual/possible distinctions can be overlaid on these shared informational dynamics. It avoids the common conception of perception and action as separate and peripheral to central cognition. Rather, it contributes to the situated cognition movement by showing how mechanisms for perceiving action can be built on those for active perception.

The SCM is developed heuristically, in five layers that can be combined in various ways to frame specific ontogenetic or phylogenetic hypotheses. The starting point is dynamic online motor control, whereby an organism is closely attuned to its embedding environment through sensorimotor feedback. Onto this are layered functions of prediction and simulation of feedback, mirroring, simulation of mirroring, monitored inhibition of motor output, and monitored simulation of input. Finally, monitored simulation of input specifying possible actions plus inhibited mirroring of such possible actions can generate information about the possible as opposed to actual instrumental actions of others, and the possible causes and effects of such possible actions, thereby enabling strategic social deliberation. Multiple instances of such shared circuits structures could be linked into a network permitting decomposition and recombination of elements, enabling flexible control, imitative learning, understanding of other agents, and instrumental and strategic deliberation. While more advanced forms of social cognition, which require tracking multiple others and their multiple possible actions, may depend on interpretative theorizing or language, the SCM shows how layered mechanisms of control, mirroring, and simulation can enable distinctively human cognitive capacities for imitation, deliberation, and mindreading.

Keywords: action; active perception; control; embodied cognition; imitation; instrumental deliberation; isomorphism; mindreading; mirroring; mirror neurons; shared circuits; simulation; social cognition

1. Introduction: From active perception to social cognition and beyond

Like many today, I view perception as inherently active (Hurley 1998; Hurley & Noë 2003, O'Regan & Noë

SUSAN HURLEY, who passed away in August 2007, had been Professor and Chair in Philosophy at the University of Bristol since August 2006 and was also a Fellow of All Souls College in Oxford. She had served as Professor and Politics and International Studies affiliate at the University of Warwick for the previous twelve years. Her most recent research had been in philosophy of psychology and neuroscience, focusing on consciousness, social cognition (imitation and mindreading), and action (rationality, control, responsibility). She had also worked in political philosophy and related areas, with a particular interest in bringing the cognitive and social sciences into constructive contact. Hurley's books include *Natural Reasons: Personality and Polity* (Oxford University Press, 1989), *Consciousness in Action* (Harvard University Press, 1998), *Justice, Luck and Knowledge* (Harvard University Press, 2003), and edited volumes on the foundations of decision theory, on imitation, on rationality in animals, and on human rights. Hurley did her undergraduate work in philosophy at Princeton University and her graduate work in philosophy (a B.Phil. and a doctorate) at Oxford University. She also earned a law degree at Harvard University. After four years (1981–1984) as a Junior Research Fellow at All Souls College, Oxford, Hurley spent ten years as a Tutorial Fellow in Philosophy at Oxford, before moving to Warwick. At the time of her passing, Hurley was one of the Principal Investigators on a large multicentre project studying the role of the natural and social environment in shaping consciousness.

ANDY CLARK is Professor of Philosophy in the School of Philosophy, Psychology, and Language Sciences, at Edinburgh University in Scotland. He was a close friend to Susan Hurley, whose ideas concerning mind and dynamics have had a large influence on his own work which concerns the nature of mind, and the cognitive role of bodily and environmental structures and processes. He is the author of several books including *Being There: Putting Brain, Body And World Together Again* (MIT Press, 1997), *Natural-Born Cyborgs: Minds, Technologies And The Future Of Human Intelligence* (Oxford University Press, 2003) and *Supersizing the Mind: Embodiment, Action and Cognitive Extension* (forthcoming with Oxford University Press).

JULIAN KIVERSTEIN is a Postdoctoral Research Fellow in the Department of Philosophy at the University Edinburgh working as part of the collaborative research project CONTACT – Consciousness in Interaction. CONTACT is a part of the EUROCORES program, Consciousness in the Natural and Cultural Context, which will run from 2006 to 2009. Susan Hurley was a Principle Investigator of the CONTACT group based at Bristol University. Kiverstein and Hurley were just beginning to collaborate on a project investigating the relationship between social cognition and consciousness using the shared circuits model, one of the many applications of shared circuits Hurley had planned to explore. Kiverstein plans to complete this work as part of postdoctoral research.

2001a; 2001b; Noë 2004) and cognition as embodied and situated. How does cognition relate to active perception? This article shows how subpersonal resources for social cognition can be built on those for active perception. Its central issue is the following: How is it possible to perceive events as instrumentally structured intentional actions, and to learn new instrumental actions by means of such perceptions of actions? The article shows how the structures and mechanisms for perceiving action and for situated social cognition can be built on those for active perception. It extends my previous contributions to a view of perception as inherently active and of cognition as embodied and situated (Hurley 1998; Hurley & Noë 2003).

The *classical sandwich* conception of the mind – widespread across philosophy and empirical sciences of the mind – regards perception as input from world to mind, action as output from mind to world, and cognition as sandwiched between. I have argued that the mind isn't *necessarily* structured in this vertically modular way (Brooks 1999; Hurley 1998). Moreover, there is growing evidence that it is not *actually* so structured in specific domains, where perception and action share dynamic information-processing resources as embodied agents interact with their environments, rather than functioning as separate buffers around domain-general central cognition.

Rather, cognitive resources and structure can emerge, layer by layer, from informational dynamics, enabling both perception and action. Such a horizontally modular structure can do significant parts (I don't claim *all*) of the work the classical sandwich conception assigned to central cognition. Here I show how this promise can be fulfilled for the perception of action and associated social cognition, as embodied agents interact with their social environments.

I first review recent work on social cognition, focusing on imitation (Hurley 2005b; Hurley & Chater 2005a; 2005b). Imitation is still popularly regarded as cognitively undemanding. However, Thorndike (1898) showed that many animals can learn through individual trial and error but not imitatively; scientists regard the later as more cognitively demanding. Imitative ability is rare across animal species and linked to characteristically human capacities: for language, culture, and understanding other minds (Arbib et al. (2000); Arbib 2005; Arbib & Rizzolatti 1997; Barkley 2001; Frith & Wolpert 2004; Gallese 2000; 2001; 2005; Gallese & Goldman 1998; Gallese et al. 2004; Gordon 1995b; Iacoboni 2005; Meltzoff 2005; Rizzolatti & Arbib 1998; 1999; Stamenov & Gallese 2002; Tomasello 1999; Whiten et al. 2005b; Williams et al. 2001). Imitation is important in adult human sociality, as well as human development, in ways we're just beginning to understand.

Part 1 of this article reviews recent research on imitation, under the headings of behavior, subpersonal mechanisms, and functions. Part 2 presents a functional architecture that shows how subpersonal mechanisms of control, mirroring, and simulation can enable distinctively human skills of imitation, deliberation, and action understanding. The *shared circuits model* (SCM) draws together many threads of work from Part 1. It includes elements suggested by various researchers, contributes further elements, and unifies these in a distinctive framework. SCM aims to show how the following are possible:

1. Building subpersonal informational resources for situated social cognition on those for active perception, while
2. Uniting a large body of evidence and theorizing in a common framework;
3. Avoiding the “classical sandwich”; and
4. Respecting the personal/subpersonal distinction and avoiding interlevel isomorphism assumptions.

Philosophers distinguish descriptions of contentful actions and mental states of persons from subpersonal (informational or neural) descriptions (Bermudez 2000; 2003; Dennett 1969; 1991; Elton 2000; Hornsby 2000; McDowell 1994). At the subpersonal level of description, information is processed and the cycling of causes and effects knits actively embodied nervous systems into environments they interact with.¹ But these processes are not correctly attributed to persons. Persons see trees, make friends, look through microscopes, vote, want to be millionaires. Subpersonal informational and causal theories explain how personal-level phenomena become possible – are *enabled* – but need not share structure with personal-level descriptions of processes as rational or conscious.² I distinguish three levels of description: the personal level, the informational/functional subpersonal level, and the neural subpersonal level.

Two questions arise about personal/subpersonal relations: (1) How are specific personal-level capacities *actually* enabled by subpersonal processes? (2) What kinds of subpersonal processes could *possibly* do the enabling work? For example, must there be isomorphism between levels? Views about question 2 can influence answers to question 1.

SCM addresses question 2 for social cognition by using subpersonal resources from an active perception approach. SCM is cast at the subpersonal functional level, not the personal or the neural levels, though it aims both to show how certain personal-level capacities can be informationally enabled and to raise empirical questions about neural implementation. Since SCM addresses the “how possibly?” rather than the “how actually?” question, it provides a higher-order theoretical model. But it also provides generic heuristic resources for framing specific first-order hypotheses and predictions about specific ontogenetic or phylogenetic stages. Its five layers, detailed below, can be re-ordered in formulating specific first-order hypotheses.

SCM’s central hypothesis is that associations underwriting predictive simulation of effects of an agent’s own movement, for instrumental control functions, can also yield mirroring and “reverse” simulation of similar perceived movements by others. Mirroring allows ends/means associations with instrumental control functions to be accessed for simulative functions bilaterally, so causes of observed movements can be simulated, as well as effects of intended acts. Such bilaterally accessible simulations of instrumental structure can provide enabling information for deliberation, imitation, and understanding the instrumental acts of others. Shared dynamics for action and perception can provide the foundations of shared dynamics for self and other, and of the self/other and actual/possible distinctions characteristic of human cognition.

Simulation has a generic sense throughout, including, but broader than, that in simulationist theories of

mindreading (Gallese 2003; see Goldman 1989; 1992 on *process-driven* simulation). Simulation *uses* certain processes to generate related information, rather than *theorizing about* them in separate meta-processes. Effects or causes can be simulated, online or offline. Simulation can be subpersonal or personal; in SCM it is subpersonal. Subpersonal processes that predict results of movement online can also generate information about results of possible movements offline. Subpersonal mirroring that enables copying can also generate information covertly about observed movements or their goals, without overt copying (cf. Barkley [2001] on executive functions as covert behavior).

2. Part 1. Review

I begin in Part 1 by reviewing recent work, strands of which are knitted together by SCM in Part 2.

2.1. Behavior

Imitative learning is a sophisticated form of social cognition. It requires *copying* in a generic sense: Perception of behavior causes similar behavior by an observer, and the similarity plays a role – not necessarily consciously – in generating the observer’s behavior. *True imitation*, restrictively understood, requires novel action learned by observing another do it, plus instrumental or means/ends structure: the other’s means of achieving her goal is copied, not just her goal or just her movements. The concept of true imitation is contested, given the different aims and methodologies of imitation researchers (Byrne 2005; Heyes 1996; 2001; Rizzolatti 2005).

Other forms of social learning can seem similar to imitation, but should be distinguished. In *stimulus enhancement*, another’s action draws your attention to a stimulus, which triggers an innate or previously learned response; but a novel action isn’t learned directly from observation. Bird A’s pecking may draw bird B’s attention to a food, which evokes pecking in bird B. In *goal emulation*,³ you observe another achieving a goal by certain means, find that goal attractive, and try to achieve it yourself. Monkey A may use a tool in a certain way to obtain an attractive object, leading monkey B to acquire the goal of obtaining a similar object. Through his own trials and errors, monkey B may arrive at the same type of tool use to obtain the object. Emulation is found in macaques, who have not shown imitative learning. In *movement priming*, bodily movements are copied, but not as learned means to a goal. Primed movements can be innate, as in contagious yawning.

Goal emulation and movement priming provide the ends-and-means components of full-fledged imitation. Ends and means can be relatively distal or proximal; the distinction is relative, not absolute. Misunderstandings can result concerning whether ends, means, or both are copied and hence whether imitation or emulation is present (Voelkl & Huber 2000, pp. 196, 201). A movement can be the proximal means to a bodily posture, which is regarded as the proximal end of the movement (Graziano et al. 2002, pp. 354–55), but posture can also be a means to more distal ends – effects on objects or others in social groups. Complex imitation can involve structured

sequences or hierarchies of ends–means relations: acquiring a goal, learning to achieve it via subgoals, and so on.

How are these forms of copying distributed across animals, children, and adults? Stimulus enhancement, goal emulation, and movement priming are certainly found in nonhuman animals. But careful experiments are needed to distinguish these from imitation proper and obtain evidence of the latter. The *two-action* paradigm has been the tool of choice. Suppose two demonstrators obtain an attractive result by two different means: One group of animals observes one demonstrator, and the other observes the other demonstrator. Will the observer animals tend differentially to copy the specific method they have seen demonstrated? If not – if the animals' choices of method do not reflect the specific method they have observed, say, because both groups converge on one method – they may be displaying stimulus enhancement or goal emulation plus trial-and-error learning, but not imitative learning. Even if they do differentially display the behavior demonstrated, this may be merely movement priming if the behavior is already in their repertoire. But if the behavior is differentially used in a new way to achieve a result, it expresses imitative learning (Call & Tomasello 1994; Nagell et al. 1993; Voelkl & Huber 2000).

The difference between copying ends and copying means is important for theorizing the phylogeny of imitation and action understanding. (*Action understanding* is short for understanding observed behavior as goal-directed action.) Is action understanding phylogenetically prior to imitation? This view seems to face an objection: Some animals copy movements (schooling fish), though we don't think they understand the others' actions. A response to this objection distinguishes movement copying from mirroring of goals (Rizzolatti 2005) and both from imitation. Movement copying may precede action understanding, whereas action understanding may require goal mirroring, but precede imitation. True imitation involves something phylogenetically rare: the flexible interplay of copying ends and copying means; a given movement can be used for different ends and a given end pursued by various means (Barkley 2001, p. 8; Tomasello 1999). This is something humans are distinctively good at.

It is difficult to find evidence of true imitation in nonhuman animals (Byrne 1995; Galef 1988; 1998; 2005; Heyes & Galef 1996; Tomasello 1996; Tomasello & Call 1997; Voelkl & Huber 2000; Zentall 2001). Early work with chimps seems to reveal imitation, but critics have challenged this interpretation effectively; the results of subsequent experiments were negative for chimp imitation. Skeptics about nonhuman imitation long had the upper hand; for example, Tomasello et al. (1993) found no convincing evidence of nonhuman imitative learning. They proposed that understanding behavior as intentional distinguishes human from nonhuman social learning. On this view, humans can imitate observed means or choose other means to emulate observed goals. Other animals don't distinguish means and goals this way; rather, they copy movements without understanding their relevance to goals, or learn about the affordances of objects by observing action on them. In neither case, it was claimed, do other animals learn about the intentional, means/end structure of observed action.

Many skeptics have now been won over by work on imitation in great apes and monkeys (Voelkl & Huber 2000; Whiten et al. 2005b), dolphins (Herman 2002), and birds (Akins & Zentall 1996; 1998; Akins et al. 2002; Hunt & Gray 2003; Pepperberg 1999; 2002; 2005; Weir et al. 2002). Continuities are described along a spectrum from the capacities of other social animals to human socio-cognitive capacities (Arbib 2005; Tomasello 1999). For example, innovative experiments extend the two-action method by using "artificial fruits" that can be opened in different ways to obtain a treat: Chimps tend to imitate for one aspect of a demonstrated task and emulate for another aspect, whereas children tend to imitate both aspects, even when the method imitated is inefficient. These and other experiments suggest that chimps imitate more selectively than children (Whiten 2002; Whiten et al. 1996; 2005b; see also Call & Tomasello 1994; Galef 2005; Harris & Want 2005; Heyes 1998; Nagell et al. 1993; Tomasello & Carpenter 2005).

Children have been called "imitation machines" (Tomasello 1999, p. 159). They don't always imitate unselectively and sometimes emulate goals instead (Gergely et al. 2002), but they have a greater tendency than chimps to imitate rather than emulate when the method demonstrated is transparently inefficient or futile (Tomasello 1999, pp. 29–30). After seeing a demonstrator use a rake inefficiently, prongs down, to pull in a treat, 2-year-old children do the same; they almost never turn the rake over to use more efficiently, edge down. By contrast, chimps given a parallel demonstration, tend to try turning the rake over (Nagell et al. 1993).⁴ The differential tendency of children and chimps to imitate suggests an interplay of biological and cultural influences, with a role for innate endowment enabling human imitation (perhaps a matter of articulated relations among multiple mirror subsystems, enabling recombinant structure in social learning, rather than the presence versus absence of a mirror system at all; see discussion of mechanisms in sect. 2.2).

Imitative and related behaviors appear throughout human development (Meltzoff 1988a; 1990; 1995; 1996; 2002a; 2002b; 2005; Meltzoff & Moore 1977; 1983a; 1983b; 1989; 1997; 1999; 2000). Infants younger than 1 month of age appear to copy facial gestures. By 14 months, infants imitate a novel act a week later: They turn on a light by touching a touch-sensitive panel with their forehead instead of their hand, differentially copying the novel means demonstrated, as well as the result (Meltzoff 1988a; 2005; cf. Gergely et al. 2002). They don't turn the light on in this odd way unless they have seen it demonstrated. By 15 to 18 months, infants recognize the underlying goal of an unsuccessful act they observe, and produce it: After seeing an adult try but fail to pull a dumbbell apart in her hands, they succeed in pulling it apart using knees and hands. But they don't pick up goals from failed "attempts" involving similar movements by inanimate devices, thus apparently discriminating agents from non-agents (Meltzoff 1988a; 1995; 1996; 2005; Meltzoff & Moore 1977; 1999; Tomasello & Carpenter 2005). Children's perception of behavior tends to be enacted automatically in similar behavior, unless actively inhibited; but frontal inhibitory functions are not well developed in young children (Barkley 2001, pp. 5, 22; Kinsbourne 2005; Preston & de Waal 2002, p. 5).

Adult “imitation syndrome” patients with frontal brain lesions also imitate uninhibitedly (Barkley 2001, p. 15; Frith 1992, pp. 85–86; Lhermitte 1983; 1986; Lhermitte et al. 1986, p. 330). They persistently copy the experimenter’s gestures, though not asked to, even when these are socially unacceptable or odd, such as putting on eyeglasses when already wearing glasses.

But the human copying tendency isn’t confined to the young or brain damaged! Normal adults can usually inhibit overt imitation selectively, which is evidently adaptive, but their underlying tendency to copy is readily revealed or released. Overt imitation is the disinhibited tip of the iceberg of continual covert imitation (Barkley 2001; Dijksterhuis 2005). Experiments show how action is modulated or induced by perception of similar action (Brass et al. 2001; Prinz 2002). Imitative tasks have shorter reaction times than nonimitative tasks; gestures are faster when participants are primed by perceiving similar gestures or their results or goals – even when primes are logically irrelevant to the task (W. Prinz 2005). Similarity between stimulus and response also affects which response is made. Normal adults, instructed to point to their nose when they hear “Nose!” and to a lamp when they hear “Lamp,” performed perfectly while watching the experimenter demonstrate the required performance, but made mistakes when watching the experimenter doing something else: they tended to copy what they saw done rather than to follow instructions (Eidelberg 1929; Prinz 1990). Movements can be induced by actions you actually perceive or by actions you would like to perceive – as when moviegoers or sports fans in their seats make movements they would like to see (W. Prinz 2005). Visually or verbally represented, as well as observed, actions can induce similar actions.

It is helpful to distinguish copying of specific behaviors from *chameleon effects*, where complex patterns of behavior are induced – a relevant kind of copying, if not strict imitation. In an experiment involving *specific behaviors*, when normal adults interact in an unrelated task with someone rubbing her foot, they rub their own feet significantly more. Transferred to another partner who touches his face, they touch their own faces instead. Demonstrations of chameleon effects show that exposure to traits and stereotypes automatically elicits general patterns of behavior and attitude and influences how behavior is performed (Bargh 1999; 2005; Bargh & Chartrand 1999; Bargh et al. 1996; 2001; Chartrand & Bargh 1999; Chartrand et al. 2005; Dijksterhuis & Bargh 2001). Normal adults primed with stimuli associated with traits (e.g., hostility, rudeness, politeness) or stereotypes (e.g., elderly persons, college professors, soccer hooligans) tend to behave in accordance with the primed traits or stereotypes. For example, hostility-primed participants deliver more intense “shocks” than control participants in subsequent, ostensibly unrelated experiments based on Milgram’s (1963) classic shock experiments. Priming can also affect intellectual performance: College professor-primed participants perform better and soccer hooligan-primed participants perform worse than controls on a subsequent, ostensibly unrelated general knowledge test (Dijksterhuis 2005; Dijksterhuis & van Knippenberg 1998).

Such priming results are robust across a wide range of verbal and visual primes and induced behavior, using

dozens of stereotypes and general traits, and various priming methods, including primes perceived consciously and subliminally. Whether subjects perceive primes consciously or not, they are unaware of any influence or correlation between primes and their behavior. These influences are rapid, automatic, and unconscious, apply both to goals and means, and don’t depend on subjects’ volition or having independent goals that would rationalize their primed behavior.

Copying, at various levels of generality, is thus a default social behavior for normal human adults; it requires specific overriding or inhibition (Barkley 2001, p. 22; Dijksterhuis 2005; Preston & de Waal 2002). Just thinking about or perceiving action automatically increases, in ways participants are unaware of, the likelihood that they will perform similar actions themselves. Nevertheless, these influences are often inhibited, as when goals make conflicting demands: elderly-primed participants tend to walk more slowly, but not if they have independent reasons to hurry.

2.2. Mechanisms

Copying perceived behavior seems to pose a *correspondence problem* (Nehaniv & Dautenhahn 2002): How is another’s observed action translated into the observer’s similar performance? When I copy your hand movements, I can see my own hands, though my visual perspectives on the two movements are different. But when I copy your facial expressions, I cannot see my own face. What information and mechanisms are needed to map perception to similar behavior?

Evidence that newborns copy facial gestures, though they cannot see their own faces, suggests innate supramodal correspondences between action and perception of similar action (Meltzoff 1988a; 1990; 1995; 1996; 2002a; 2002b; 2005; Meltzoff & Moore 1977; 1983a; 1983b; 1989; 1997; 1999; 2000). Although further correspondences could be acquired as imitative abilities develop, skeptics about newborn copying can also be skeptical about the need to postulate any innate correspondences (Anisfeld 1979; 1984; 1991; 1996; 2005; Anisfeld et al. 2001; Heyes 2005).

Heyes, who is one such skeptic, argues that sensorimotor associations subserving copying can be acquired through general-purpose associative learning mechanisms whereby neurons that fire together wire together. Direct sensorimotor associations between motor output and sensory feedback could result from watching one’s own hand gestures. An indirect route is needed when the agent cannot perceive her own actions, as in facial expressions: The sensorimotor association could be mediated by environmental items such as mirrors, action words, or stimuli that evoke similar behavior in the actor and in other agents the actor observes. Moreover, adults commonly copy infants, performing the associative function of a mirror. When baby smiles and father smiles back, baby’s motor output is associated with sensory input from father’s smile (Heyes 2005, p. 161; Preston & de Waal 2002, p. 8). Imitation can thus develop from interactions between organisms with associative learning mechanisms and certain cultural environments (see Heyes 2001; 2005; see also SCM, Layer 3 in sect. 3.3 here).

Common coding for perception and action has been postulated to explain human copying tendencies. On this view, perception and action share subpersonal processes carrying information about (“coding for”) what is perceived or intended, in which perception and action are not distinguished. The differentiation between perception and action is overlaid on those shared resources, so that they are informationally interdependent at a basic level. If capacities X and Y share an information space, their commonality is informationally prior to their differentiation.

Meltzoff and Moore (1997) postulate common coding of perception and action in explaining infant imitation: proprioceptive feedback is compared to an observed target act, where these are coded in common, supramodal terms. Innate common coding could initially be for relations among, say, lips and tongue; more dynamic, complex, and abstract common coding could develop with experience of body babbling. Common coding might also be acquired as in Heyes’s (2005) model.

Wolfgang Prinz (1984; 2005; cf. Bekkering & Wohlschläger (2002); Preston & de Waal 2002, pp. 4ff., 9–10) appeals to common coding of perception and action to explain the normal adult tendency to imitate and the reaction-time advantage of imitative tasks. Common coding facilitates imitation, avoiding the correspondence problem and any need for translation between unrelated input and output codes to solve it. Prinz associates common coding with what William James called *ideomotor theory*, on which every representation of movement awakes in some degree the movement it represents (Brass 1999; Prinz 1987). Perceiving another’s observed movement tends inherently to produce similar movement by the observer, and primes similar movement even when it doesn’t break through overtly. Regular concurrence of action with perceived effects allows prediction of an action’s effects and selection of action, given an intention to produce certain effects (Greenwald 1970; 1972). Thus, representation of an action’s regular result, whether proximal or distal, can evoke similar action, in the absence of inhibition.

Other sources also support the view that perception and action share processing resources. Observing an action primes the very muscles needed to perform the same action (Craighero et al. 2002; Fadiga et al. 1995; 2002). Watching an action sequence speeds the observer’s performance of that sequence; merely imagining a skilled performance, in sport or music, improves performance – is a way of practicing – as many athletes and musicians know (Jeannerod 1997, pp. 117, 119–22; Pascual-Leone 2001). Similar points concern perception and experience of emotion: Gordon argues that a special containing mechanism, which isn’t fail-safe, is needed to keep emotion recognition from producing emotional contagion. On his simulationist theory, only a thin line separates one’s own mental life from one’s representation of another’s; offline representations of others tend inherently to go online (Gordon 1995b; cf. Adolphs 2002; Preston & de Waal 2002).

Common coding theories characterize subpersonal architectures for copying functionally. What neural processes might implement such functional architectures?

Certain neurons directly link perception and action: their firing correlates with specific perceptions and specific actions. *Canonical neurons* (Gallese 2005; Rizzolatti 2005) reflect affordances (Iacoboni 2005; Miall

2003): They fire when an animal perceives an object that affords a certain type of action and when the animal performs the afforded action. *Mirror neurons* fire when an animal perceives another agent performing a type of action, and also when the animal performs that type of action itself; they don’t distinguish own action from others’ similar actions (see SCM, Layer 3 in sect. 3.3). Some fire, for example, when a monkey sees the experimenter bring food to her own mouth with her hand or when the monkey brings food to his own mouth with his hand (even in the dark, so the monkey cannot see his hand). Specificity of tuning varies.

How mirror neurons relate to imitation is of much current interest (e.g., see Frith & Wolpert 2004; Rizzolatti et al. 2002; Williams et al. 2001). It may be tempting to think they avoid correspondence problems, thus facilitating imitation: If the same neurons code for perceived action and similar performance, no translation is needed. But things are not so simple. Rizzolatti, one of the discoverers of mirror neurons, holds that imitation requires both the ability to understand another’s action and the ability to replicate it. On his view, recall, action understanding precedes imitation phylogenetically; action understanding is subserved by mirror systems, which might be necessary, but are not sufficient, for imitation. Rizzolatti (2005) suggests that the motor resonance set up by mirror neurons makes action observation meaningful by linking it to the observer’s own potential actions.

Mirror neurons were discovered by single-cell recording in macaques (Di Pellegrino et al. 1992; Rizzolatti et al. 1988; 1995), which can emulate but have not been shown to imitate in a strict sense (cf. Voelkl & Huber [2000] on marmoset imitation). Evidence for human mirror systems (Craighero et al. 2002; Decety & Chaminade 2005; Decety et al. 1997; Fadiga et al. 2002; Hari et al. 1998; Iacoboni et al. 2005; Rizzolatti et al. 1996; Ruby & Decety 2001) includes brain imaging and demonstrations that observing another person move primes the muscles needed to move in a similar manner (whether or not movements are goal directed; Fadiga et al. 1995).

Rizzolatti (2005) describes mirror neurons in monkey frontal brain area F5 as part of a circuit including parietal area PF and visual area STS (superior temporal sulcus). He regards a similar human brain circuit as a control system: Sensory results associated with certain movements are compared in PF to observed target movements, enabling imitative learning (cf. Iacoboni 2005, with regard to locating the comparator in STS). Differently structured mirror systems may explain different copying capacities across species. In monkeys, mirror neurons appear to code for the goals or results of performed or observed actions.⁵ By contrast, human mirror systems include specific movements that can be means to achieving goals (Fadiga et al. 1995). Recall how the difference between mirroring ends versus means of action matters for the view that action understanding precedes imitation phylogenetically. If seeing someone reach for an apple produces motor activation associated with the same goal in the observer (though not necessarily with the same movements in the observer), that could provide information about the observed action’s goal directness. But it wouldn’t provide information about how to achieve the goal by means of the observed movements, as in imitation.

Brain imaging suggests a division of labor in the human mirror system: Its frontal regions tend to code for goals of action, whereas its parietal regions tend to code for means (i.e., movements; Iacoboni 2005). (Monkey parietal mirror neurons seem to be goal-related; Fogassi et al. 2005; Nakahara & Miyashita 2005.) One theory of how this division of labor enables imitation relates signals generated by these brain regions to comparator circuits for instrumental motor control combining inverse and forward models. *Inverse models* estimate the motor plan needed to achieve a goal in a given context. They can be adjusted by comparison with real motor feedback, but this is slow. It is often more efficient to use real feedback to train *forward models*, which anticipate the sensory effects of motor plans, associating action with its perceived results (as do mirror neurons). Forward models combine with inverse models to control goal-directed behavior more efficiently. Forward models can predict the results of imitative motor plans for comparison to observed action, and motor plans can be adjusted until a match obtains (Flanagan et al. 2003; Iacoboni 2005; Miall 2003; Wolpert et al. 2003).

Thus, mirror neurons are arguably part of the neural basis for true imitation, though not sufficient for it. Monkey mirror neurons code for ends rather than means. Human mirror systems, by contrast, have articulated structure: some regions code for goals, whereas others code for specific movements that are means to goals. It has been suggested that human mirror systems enable imitation (not just emulation) because they code for means as well as ends (unlike the macaque's system), and that mirror neurons contribute predictive forward models to subpersonal comparator control circuits.

2.3. Functions

Human brains differ most from chimp brains in expanded areas around the Sylvian fissure which subserve imitation, language, and action understanding – where many mirror neurons are found (Iacoboni 2005). Can mirror systems illuminate the functions of imitation in relation to distinctively human capacities – for language, or for identifying with others and understanding the mental states motivating their actions? The relationships among capacities for imitation, language, and mindreading are important for understanding phylogeny and human development. Does development of either language or mindreading depend on imitation? If so, at what levels of description and in what senses of “depend”? Or does dependence run the other way? Or both ways, dynamically? Answers may differ for language and for mindreading. Issues about relations between imitation and mindreading entwine with issues about whether mindreading is best understood as theorizing about other minds or as simulating them.

I shall survey some hypothesized functions of imitation in language, cultural evolution, cooperation, and mindreading. The first three topics, discussed briefly, provide context for SCM and illustrate its broader relevance to understanding what is distinctive about human minds. Mindreading is directly related to SCM and so is discussed more fully.

2.3.1. Language. It has been suggested that “mirror neurons could . . . be an important neural stepping stone . . .

to spoken language” (Miall 2003, p. 1). Mirror systems for action goals include Broca's area,⁶ a main language area of human brains, which is active during imitative tasks. Moreover, transient virtual “lesions” to Broca's created by transcranial magnetic stimulation interfere with imitative tasks (Heiser et al. 2003; Iacoboni 2005). Nativism about language might view Broca's as the best candidate for an innate language module (M. Iacoboni, in discussion). But discovery that Broca's subserves mirror systems and has some role in enabling imitation has generated new arguments about how language acquisition could build on capacities for action understanding and imitation, in either evolutionary or developmental time frames, exploiting imitative learning rather than, or in addition to, innate linguistic knowledge (Arbib 2005; Arbib & Rizzolatti 1997; Iacoboni 2005; Rizzolatti & Arbib 1998; 1999; Stamenov & Gallese 2002). (On language and social learning, see Baldwin 1995; Barkley 2001; Christiansen 1994; 2005; Christiansen & Kirby 2003; Deacon 1997; on establishing shared reference to objects through joint attention, via gaze following and role-reversal imitation, see Tomasello 1999.)

What features of imitation and human mirror systems might language build on? First, I suggest, flexible articulated relations between means and ends in imitative learning could be an evolutionary precursor of arbitrary relations between symbols and referents. Decoupling a particular bodily movement from a given result and treating it as a potential means to various possible results in varying circumstances (see SCM, Layers 2 plus 4, in sects. 3.2 and 3.4) may be a step toward treating it as lacking an intrinsic function and so available for an arbitrarily or conventionally assigned communicative function.

Second, mirror systems provide a common code for actions of self and other, and thus for language production and perception; by enabling intersubjective action understanding, mirror systems may be the basis for the intersubjective *parity*, or sharing of meaning, essential to language (Arbib 2005; Iacoboni 2005).

Third, the flexible recombinant structure of ends and means in imitation may be a precursor of recombinant grammatical structure in language (Arbib 2005). The latter may result when creatures with recombinant imitative skills learn to pursue their goals by recombinant manipulation of external symbols.

Fourth, finding recombinant units of action in streams of bodily movement has parallels with finding linguistic units (e.g., words) in continuous acoustic streams of speech (Byrne 2005). The modular structure of skilled action facilitates flexible recombination. Patterns of action organization could be learned in program-level imitation,⁷ despite variation in implementational details, by using mirror mechanisms plus mechanisms for parsing behavior modules. Behavior parsing and the recombinant structure of program-level imitation may be precursors of human capacities to perceive underlying structures of intentions or causes in the surface flux of experience – and perhaps of syntactic parsing and the recombinant structure of language.

2.3.2. Cultural evolution. A more fundamental question is: Why might evolution favor neural structures that enable various forms of copying to begin with? Suppose individuals vary in behavioral traits that are not genetically

heritable, so some reproduce more successfully than others. Their offspring may benefit by acquiring behaviors from their successful parents by copying, as well as genetically. By copying reproductively successful parents, offspring can acquire nonheritable behaviors associated with appropriate environmental conditions. If individual learning is costly, copying may contribute more to genetic fitness.

If true imitation requires mirror circuits for means and ends to be linked in ways that give social learning recombinant flexibility, it should be harder to evolve than movement priming or emulation. And, indeed, it is found in fewer species. But wouldn't this rare development from other forms of copying to imitation be maladaptive? Recall the short-term disadvantage of children compared to chimps in two-action paradigms: Children have a greater tendency to imitate even inefficient models, whereas chimps have a greater tendency to emulate and find more efficient means to attractive goals (Nagell et al. 2003; Whiten et al. 2005b). Despite this, could the stronger imitative tendency be adaptive long-term?

Yes: via the *ratchet effect* (Tomasello 1999). Gifted or lucky individuals may discover efficient new means to goals – means that are not readily rediscoverable by independent trial-and-error learning. These would be lost without recombinant imitative learning, which preserves and disseminates valuable instrumental innovations, providing a platform for further innovation. Once imitation evolves genetically, it provides a mechanism of cultural and technological transmission, accumulation, and evolution. The effects of imitative copying and selection intertwine with those of genetic copying and selection; culture and life coevolve (see and cf. Baldwin 1896; Barkley 2001, p. 21; Blackmore 1999; 2000; 2001; Boyd & Richerson 1982; 1985; Dawkins 1976/1989; Deacon 1997; Dennett 1995; Gil-White 2005; Henrich & Boyd 1998; Henrich & Gil-White 2001; Hurley & Chater 2005a, part 4).

The capacity for *selective* imitation may have an important role in underwriting the ratchet effect (Harris & Want 2005). Imitation with selective inhibition has the advantages of theft over honest toil: Instead of letting *hypotheses* die in his stead, a selective imitator lets *others* die in his stead, reaping the benefits of success without unusual native wit while avoiding the costs of trial and error. Imitative social environments may in turn generate pressure to prevent successful techniques being appropriated cost free by competitors, resulting in capacities for covert or simulated action, shielded from potential imitative theft (Barkley 2001, pp. 9, 18–21).

2.3.3. Cooperation. As well as being subject to automatic copying influences, humans often deliberately select a behavior pattern to imitate because it is associated with certain traits or stereotypes, even if they themselves don't exemplify these traits or stereotypes. This can be benign and contribute to moral development (J. Prinz 2005); perhaps I can become virtuous, as Aristotle suggested, by behaving like a virtuous person.

But, like automatic copying, deliberate selective imitation does not always operate benignly. Selective imitation can provide “Machiavellian” social advantages (Byrne & Whiten 1988; Whiten & Byrne 1997). It can steal not only instrumental successes but also cooperative benefits from competitors. Suppose information about the

mental states of others is not transparently available. A cooperative group can share certain behaviors by which members identify one another, obtain cooperative benefits, and exclude free-riding noncooperators. Cooperators may copy such identifying behaviors from other cooperators.

Noncooperators could invade such a cooperative group by selectively copying its identifying behaviors. They could thus induce cooperation from group members while failing to cooperate in return, deceptively obtaining cooperative benefits without paying the costs. Free riding via deceptive copying partially appropriates cooperative benefits based on in-group behavioral copying. While *greenbeard genes* could produce genetically determined analogues of such free riding (Dawkins 1982, p. 149), selective copying provides the evolutionary advantages of *flexible* free riding, which is not dependent on genes for specific behaviors.

How can cooperative benefits be defended against free riding through deceptive copying? An arms race between behavioral signaling and deceptive copying in cooperative games arguably produces pressure for imitative and mindreading abilities. As a result, certain solutions to cooperative games, which require mindreading rather than mere behavior prediction, may become available. Mindreading can be based on behavioral evidence yet still have functional advantages over behavior prediction (Hurley 2005a).

To elaborate: To counter invasion by increasingly sophisticated deceptive mimics, mutual recognition processes among cooperators would move progressively further from copying and detecting superficial behaviors and toward more subtle and covert imitation and detection of underlying mental causes of behavior. Mere behavior reading would move toward ever-smarter reading of behavioral evidence for intentions. Mere copying would in turn become more creative and flexible, with means/ends structure: imitation. This arms race could produce capacities for mindreading and intersubjective identification via covert mirroring, albeit based on subtle behavioral perceptions (cf. Krebs & Dawkins 1984).

The advance from cooperation plus deceptive copying via imitation to mindreading is significant for enabling cooperation and obtaining its benefits. Certain solutions to collective action problems effectively require recognizing and identify with others' mental states. A simple self-referential *mirror heuristic*⁸ for non-iterated Prisoners' Dilemmas (PDs) says: cooperate only with any others you meet who act on this same rule (Howard's *mirror strategy* [Howard 1988]; Danielson's *self-same cooperation* [Danielson 1991; 1992⁹]). When another player doesn't share your mirror heuristic, you don't cooperate with him. Famously, Tit-for-Tat can outperform Defection in iterated PDs, where given players meet repeatedly; but mirror heuristics outperform Defection even in non-iterated PDs, where given players *don't* meet again.¹⁰

Mirror heuristics effectively require mindreading: discovering another player's intention, not simply predicting his behavior (Danielson 1992, pp. 75–82; Schmitt & Grammar 1997).¹¹ They are *conditional metaheuristics*: they explicitly condition cooperation on the other's operative heuristic itself, not on his predicted behavior. (Tit-for-Tat requires not mindreading but memory of a given

player's behavior in past games.) Employing a mirror heuristic requires discerning, more or less reliably, whether others are operating on a mirror heuristic – a general intention or rule of choice. Which choices mirrorers should make are not determined by predicting what others will do; mirrorers need to know whether others have the intentions of a mirrorer before they can determine what to do. Participants in mirror-based cooperation must not only *be* mindreaders, but also be able to *identify*, more or less reliably, other mindreaders.

In non-iterated games, the mindreader has not previously played with the player she is mindreading, so cannot refer to memories of their past play. In informationally clouded social environments, mindreading is based on evidence from observing behavior, which may be subtle and/or deceptive. But mindreading need not be foolproof to provide mirror-based cooperative benefits to individual mirrorers and groups of them; the benefits would vary with the accuracy of mindreading (cf. Danielson 1992, pp. 157ff.).

What is the difference between genuine mindreading and mere smart behavior reading? Many social problems that animals face can be solved via behavior-circumstance correlations and behavioral predictions, without postulating mediating mental states. What problem-solving pressures are addressed by additionally attributing mental states to explain observed behavior? (Cf. Call & Tomasello 1999; Hare et al. 2000; 2001; Heyes 1998; Heyes & Dickinson 1993; Hurley 2006b; Povinelli 1996; Povinelli & Vonk 2006; Sterelny 2003, pp. 67ff.; Tomasello & Call 2006; Whiten 1996; 1997.)

Mental state attributions may support more flexible behavior prediction in novel conditions. But mirror meta-heuristics show that mindreading's function in enabling cooperation goes beyond providing better predictions of behavior. As explained, mirror metaheuristics do not require predicting other players' behavior per se, but rather, ascertaining the heuristic they use. Such mindreading may be done by *observing* others' behavior, but that does not mean that its function is only behavior *prediction*, or that it has the same functions as behavior prediction. Mindreading can *function* to enable cooperation in a way that merely predicting behavior cannot (Danielson 1992, p. 82), even if mindreading is based on behavioral *evidence*. This is why the emergence of mindreading, via imitation, from an arms race between cooperative and deceptive copying is significant for enabling cooperation.

2.3.4. Mindreading. What more can be said about the possible functions of imitation in relation to mindreading?

Human mirror systems may be part of the mechanisms for understanding observed actions and intersubjective empathy. Observing another act primes your motor system to copy, even if overt copying is inhibited. Covert copying is a kind of process-driven simulation, which uses offline the processes that would be used actually to copy the observed action, but it inhibits motor output. This direct resonance with another's action provides a fundamental similarity between yourself and other agents that enables the understanding of another's actions as instrumentally structured. Mirror systems also provide a plausible neural basis for emotional empathy and understanding (see Adolphs 2002; Decety & Chaminade 2003; 2005; Gallese 2001; 2005; Gallese & Goldman

1998; Goldman 2005; Gordon 1995a; 1995b; Iacoboni 2005; Iacoboni et al. 2005; Meltzoff 2005; Preston & de Waal 2002; Rizzolatti 2005; Williams et al. 2001).

Within this broad perspective, I shall compare the views of Gallese, Meltzoff, Gordon, and Tomasello on simulation theory versus theory theory and on relations between imitation and mindreading. And I will preview how SCM reconciles opposed views on both topics. An outline follows.

2.3.4.1. Simulation theory (ST) versus theory theory (TT).

Gallese views mirror systems as enabling broad interpersonal empathy by implementing primitive intersubjective information, prior to differentiation of self from other.

Meltzoff views early imitation as foundational for the ability to understand other agents: In imitation my acts are directly, noninferentially identified with others' acts; I then associate my acts with my mental states and infer a similar association in others.

TT accounts of mindreading invoke laws and inferences about mental states and behavior, whereas on ST accounts mindreaders use their own psychological processes offline to attribute similar mental states or actions to others.

Undersimilarity between one's own and others' acts is shared ground between Meltzoff's TT account of mindreading based on early imitation and ST accounts based on inhibited copying.

Gordon criticizes first-person- to third-person-inference accounts of early imitation's role in mindreading, for taking the self/other distinction for granted.

In Gordon's ST view of how mindreading involves offline imitation, "constitutive mirroring" "multiplies the first person" by reference to a shared scheme of reasons.

First reconciliation of ST and TT: Foundations of intersubjectivity and the self/other distinction can be provided by simulative mirroring (SCM's layers 3 and 4), although richer self/other and other/other distinctions depend on interpretation, theorizing, and inference (layer 5 and beyond).

2.3.4.2. Relations between imitation and action understanding.

Tomasello and Carpenter's view that imitation depends on action understanding contrasts with views of action understanding as depending on imitation.

Second reconciliation, concerning relations between imitation and action understanding: Simple mirroring, of goals or movements, can express a fundamental intersubjectivity, enabling simple forms of action understanding and providing elements of more complex imitative mirroring with flexible instrumental structure, which in turn contributes to more articulated, instrumentally structured understanding of other agents and their minds (layers 3 and 4).

2.3.5. Simulation theory (ST) versus theory theory (TT).

On Gallese's (2001; 2005) *shared manifold hypothesis*, mirror systems enable various aspects of interpersonal understanding and empathy. Mirror systems develop from the way biological control systems model interactions between organisms and their environments. They provide

the neural basis of a primitive intersubjective information space or “shared manifold” that is prior to a self/other distinction both phylogenetically and ontogenetically, yet preserved in human adults (SCM’s layer 3 incorporates this feature). The shared manifold underwrites automatic intersubjective identifications across different perceptual modalities and action, but also for sensations and emotions: There is evidence of mirror mechanisms for pain and disgust, and hearing anger expressed increases the activation of muscles used to express anger. Empathy involves a common scheme of reasons under which persons, self and others, are intelligible, rather than recognition that others’ bodies also have minds.

Meltzoff (2005) argues that early imitation and its enabling mechanisms begot understanding of other agents, not vice versa. In his view, ability to understand other minds has innate foundations, but develops in stages, in which imitation plays a critical role. Infants have a primitive ability to recognize being imitated and to imitate, and hence to distinguish people from other things and recognize equivalences between acts of self and other. This initial bridge between self and other provides privileged access to people that we don’t have in relation to other things; it develops in three stages. First, own acts are linked to others’ similar acts supramodally, as evidenced by newborns’ imitation of others’ facial gestures. Second, own acts of certain kinds are linked bidirectionally to own mental states of certain kinds, through learning. Third, others’ similar acts are linked to others’ similar mental states. This early process is conceived not as formal reasoning, but as processing the other as “like me.” It gets mindreading started on understanding agency and closely associated mental states: for example, intentions, emotions, and desires. Meltzoff emphasizes that mindreading isn’t all or nothing. (Tomasello [1999] makes similar claims for nonhuman animals.) Understanding mental states further from action, like false beliefs, comes in later development.

Meltzoff’s view of mindreading is usually put in the *theory theory* (TT) rather than *simulation theory* (ST) category, but it has ST as well as TT aspects. TT regards commonsense psychology as a proto-scientific theory. It represents knowledge as laws about mental states and behavior that can be known innately or discovered by testing hypotheses against evidence. Specific mental states and behaviors are inferred from other mental states and behaviors by means of such laws; the process does not depend on copying. On ST, mindreading starts with taking another’s perspective and generating “pretend” mental states or behavior that match the other’s. These offline states are not objects of theoretical inference. Rather, they are entered into the simulator’s own psychological and decision-making processes, which are held offline to produce further simulated mental states and behavior that are then assigned to the other. Further behavior by the other can be predicted, or mental states attributed that explain his observed behavior. Such simulation is an extension of practical abilities rather than a theoretical exercise: it copies the other’s states and uses the copies in the simulator’s decision-making equipment, instead of using laws to infer the other’s states (Davies & Stone 1995a; 1995b).

Meltzoff’s three-stage process can be restated in explicitly TT terms. First, innate equivalence between my

own and others’ acts (exploited by early imitation and recognition of being imitated) provides a fundamental, underived similarity between some acts (by myself) and other acts (by another). Second, first-person experience provides laws linking one’s own acts and own mental states. Third, it is inferred that another’s acts and mental states are lawfully linked in the same ways as my similar acts and mental states are linked. Proceeding through stages 2 and 3, we find inferences from first-person mind-behavior links to similar third-person links as in traditional arguments from analogy.¹² “The crux of the ‘like-me hypothesis’ is that infants may use their own intentional actions as a framework for interpreting the intentional actions of others” (Meltzoff 2005, p. 75). For example, 12-month-old infants follow a model’s “gaze” significantly less when the model’s eyes are closed rather than open, but only similarly refrain from following the “gaze” of a blindfolded model after they are given first-person experience with blindfolds.

However, as Meltzoff points out (personal communication), there is no first-person to third-person inference at stage 1. The initial bidirectional self-other linkage, expressed in early imitation and recognition of being imitated, is via a supramodal common code for observed and observer’s acts that’s direct and noninferential (Meltzoff & Moore 1997). Stage 1 of Meltzoff’s view thus has important common ground with ST: In covert offline copying, direct noninferential resonance with another’s action with inhibited motor output enables understanding of the other’s action. But such direct noninferential resonance can also occur in overt copying, as Meltzoff postulates; copying can provide information for understanding another’s actions, even when not inhibited and serving other functions. Mindreading’s foundation at Meltzoff’s first stage is noninferentially direct, not theoretically derived. His view shares this nontheoretical basis, at its first, online copying stage, with ST views of mindreading as based on offline copying, though they diverge on how mindreading develops further. If mindreading develops in stages, theoretical inference can enter later, increasing with development.

While Meltzoff’s theory theory involves first-person to third-person inference, Gordon’s “radical” simulation theory (see Gordon 1986; 1995; 1995a; 1996; 2002; 2005) explicitly rejects it, and provides a different view of the relations between imitation and mindreading. In *constitutive mirroring*, a copied motor pattern is part of the perception of another’s action, though overt movement may be inhibited. Gordon finds constitutive mirroring in Gallese’s primitive intersubjective “we”-space, the basis of empathy that implicitly expresses *similarity* of self and other rather than their *distinctness*. When constitutive mirroring imposes first-person phenomena, a process of analysis by synthesis occurs whereby another’s observed behavior and the self’s matching response – part of the very perception of the other’s behavior – become intelligible together, in the same process. When I see you reach to pick up the ringing phone, your act and my matching response are made sense of together within a scheme of reasons that is fundamentally common to persons. As Gordon (2005) puts it, I don’t infer from the first to the third person, but rather multiply the first person. To understand what I or another believes, perceives, or intends, I look out at the world and the

reasons it provides, though in the case of others I imaginatively recenter to the other's perspective (Gordon 1995a).

Gordon criticizes first-person to third-person inference in Meltzoff's account not because it attributes *similarity* to one's own and others' acts or experiences, but because it requires that they be *identified* and *distinguished*. In Meltzoff's stage 1, there is innate equivalence between acts of self and other; this stage may involve constitutive mirroring, as in Gallese's primitive shared manifold. But later stages of Meltzoff's account, where analogical inference occurs, require that self and other also be distinguished: If this kind of act *by me* is linked to *my* mental states of a certain kind, then a similar (as per stage 1) kind of act *by another* is also linked to *her* mental states of a similar kind. Gordon explains that I cannot infer analogically from *a* to *b* unless I can distinguish *a* from *b*. He is skeptical that infants have this capacity, though mature imitative mirroring may involve such inference (Gordon 2005).

Pure ST views of mindreading are standardly criticized for lacking resources to explain how mature mindreaders distinguish and identify people and track which actions and mental states are whose. Gordon suggests that multiple first persons are distinguished and tracked in the personal-level process of making others intelligible, avoiding incoherence under the common scheme of reasons (see also Hurley 1998, part 1; 1989). Mental states that don't make sense together are assigned to different persons. Can this be done in pure simulation mode, without theorizing? Simulation supposedly *uses* practical abilities rather than *theorizing about* actions. How does interpreting an action to make sense of it differ from theorizing about it? When I use practical reason offline in interpretative mindreading, I don't formulate normative laws from which I make inferences; rather, I activate my normative and deliberative dispositions. As Millikan might say (Millikan 2005), my thought about another's action isn't wholly separate from my entertaining that action.

SCM will suggest a reconciliation between ST and TT. The fundamental *similarity between self and other* is understood in terms not of theorizing but mirroring (as in Gallese's shared manifold, Gordon's constitutive mirroring, Meltzoff's innate self-other equivalence, and SCM's layer 3). Such primitive intersubjectivity persists into adulthood, providing a basis for mature empathy and mindreading, as Gallese holds. The informational origin of the *self/other distinction* is understood in terms of monitoring whether mirroring is inhibited (layer 4). As mindreading develops, it also employs a richer self/other distinction, as when children come to distinguish imitating from being imitated (see Decety & Chaminade 2005), or to attribute beliefs different from their own to others. Mature personal-level mindreading requires abilities to distinguish, identify, and track multiple other persons, to assign acts and mental states to them in an interpretative process, and to entertain multiple possible acts by multiple other persons (layer 5). If decentering from me-here-now creates a trail to others and other possible actions, mature mindreading creates multiple branching and interacting trails. Negotiating these, by using the full range of distinctions and identifications required by mature mindreading, probably demands theoretical resources, even though the subpersonal enabling foundations of intersubjectivity are found in mirroring, and

of the self/other distinction in monitored simulation. SCM explains how mirroring and simulation can provide foundations for mindreading on which theorizing builds.

2.3.6. Relations between imitation and action understanding. How are imitation and action understanding related to each other? On *imitation-first* views, imitation underwrites early mindreading abilities. Gallese, Meltzoff, Gordon, and Goldman stress the contribution of imitation to understanding other agents. By contrast, *understanding-first* views emphasize the way imitative learning depends on action understanding and intention reading (Carpenter et al. 1998; Rizzolatti 2005; Tomasello & Carpenter 2005). Recent paradigms with children where the demonstrated action is unsuccessful or accidental (Meltzoff 1995) distinguish imitation from other forms of social learning more clearly than the two-action method does. If the observer copies what was intended even though it wasn't achieved, as opposed to copying only the observed movements or the observed though unintended result, that suggests the observer understands the intentional structure of the observed action. Tomasello and Carpenter argue that intention reading is needed to explain what is copied by imitators when the modeled behavior is the same across conditions while the modeled intention varies. In their view, results from various paradigms are most parsimoniously explained by holding that children use their understanding of intentions to imitate.

Imitation-first and understanding-first views are not necessarily opposed; each may tell only part of the story. SCM provides a framework for their reconciliation, accommodating both views at different points in its layered architecture. Different types of copying, and covert forms of each that enable corresponding types of understanding, can dovetail over evolution and development, building on one another reciprocally, with increasing instrumental structure in both action and understanding over successive stages. A simpler form of copying can precede a simpler form of understanding, which precedes a more complex form of copying, which precedes a more complex form of understanding (the ordering can be interpreted phylogenetically or ontogenetically). Start, say, with goal mirroring and emulation. Covert goal mirroring can then enable understanding the goals of observed action. Such goal understanding, along with mirroring of movements, may be needed for instrumentally articulated imitation (understanding first). But richer instrumental understanding, of how observed means contribute to observed ends, may involve covert imitation (imitation first). In SCM, self-other similarities expressed by mirroring, whether more or less structured, are informationally prior to the self/other distinction required for understanding action as another's.

3. Part 2. The shared circuits model

The shared circuits model (SCM) shows how subpersonal resources for control, mirroring, and simulation can enable the distinctively human sociocognitive skills of imitation, deliberation, and mindreading. The model has intertwined empirical and philosophical aims. One aim is to provide a unified framework for the various strands of

empirical evidence and theorizing surveyed thus far. Another is to illustrate the philosophical view that embodied cognition can emerge from active perception, avoiding the “classical sandwich” architecture, which insulates central cognition from the world between twin buffers of perceptual input and behavioral output (Hurley 1998; 2001). It does this, recall, by addressing a *higher-order* theoretical question, about how it is *possible* for subpersonal processes to enable certain personal-level abilities: in particular, how it is possible to build subpersonal resources for sociocognitive skills on those for active perception.

SCM thus provides a generic heuristic framework for specific first-order hypotheses, about how particular socio-cognitive capacities map onto specific layers of the model or develop in phylogenetic or ontogenetic time. SCM itself does not articulate specific first-order hypotheses, but does make some general predictions: for example, of neural systems for mirroring based on those for instrumental prediction, and of the priority of online over offline mirroring. Nor is SCM exclusive; important work in enabling persons’ cognitive capacities is done by other processes, including linguistic processes. The point is to illustrate how it is *possible* for important cognitive resources to emerge from active perception. Although SCM is therefore somewhat abstract, in accord with its higher-order theoretical aim, I suggest in this target article how it lends itself to more specific empirical predictions, and hope the commentaries will offer suggestions as well. Details follow, layer by layer.

3.1. Layer 1: Basic adaptive feedback control

SCM begins with specific comparator feedback control systems. A comparator system generates outputs that are means to a target, by establishing an instrumental association between outputs and their results. For example, a thermostat compares a target signal with an input signal. If they don’t match, system output is adjusted and the resulting change in input signal or feedback is tracked. Input continues to be re-compared with target and output readjusted, to minimize mismatch to target. The elements of such control are (Fig. 1):

1. A target or reference signal (e.g., target room temperature for a thermostat).
2. An input signal (e.g., actual room temperature), the joint result of elements 3 and 5.
3. Exogenous environmental events (e.g., nightfall).
4. A comparator, which determines whether target and input signals match and the direction and degree of any mismatch or error (e.g., the room is still five degrees below target temperature).
5. The output of the control system (e.g., the level of heat output), regulated by comparison between target and input signals (e.g., heat output is increased if measured room temperature is below target).
6. A feedback loop by which output has effects on succeeding input signals (e.g., measured room temperature rises when heat output increases).

Feedback control is adaptive; output is adjusted to compensate for changing exogenous influences, keeping sensed input close to target. Under different exogenous influences, feedback calls for differing outputs to achieve the target; when the weather changes, a thermostat

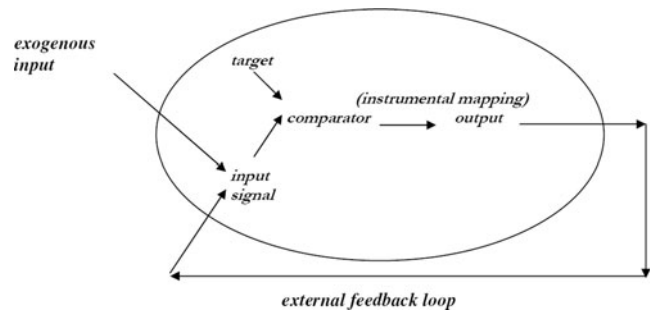


Figure 1. Layer 1: Basic adaptive feedback control.

adjusts heat output to maintain the target temperature. Feedback at layer 1 operates in real space and time, and therefore can be slow (e.g., a room takes time to warm up after the heat is turned up). A control system implements a mapping from the target, in the context of actual input, to output, thus specifying the means for approaching the target in given circumstances. *Inverse model* is engineering terminology for this instrumental mapping.

Net sensed input results from the system’s output plus independent environmental influences. In organisms, *reafferent* feedback carries input resulting from the organism’s own activity, whereas *exafferent* input results from exogenous events. Reafference includes visual and proprioceptive inputs resulting from movements of one’s hands, movement through space, manipulation of objects, and so on. Exafference includes visual inputs resulting from environmental events, such as movements by others in a social group. However, at layer 1, information distinguishing reafference from exafference is not available.

Feedback control is a cyclical and dynamic process, with no nonarbitrary start, finish, or discrete steps; input is as much an effect as a cause of output (Marken 2002; Powers 1973). Control depends on dynamic relations among inputs and outputs. Information about inputs is not segregated from information about outputs; this blending of information is preserved and extended in the informational dynamics of further layers. Perception and action arise from and share this fundamental informational dynamics (Hurley 1998; 2001).

Specific means/ends associations or instrumental mappings can be chained (output A is the means to controlled result B, while B in turn is the means to controlled result C, and so on) or organized into hierarchies. There are independently determined evolutionary, developmental, and individual differences in the grain and complexity of the possible control sequences and hierarchies of different creatures.

3.2. Layer 2: Simulative prediction of effects for improved control

Real-time feedback can be slow and produce overshooting. Control functions can be speeded and smoothed by adding simulative predictions to a comparator system (Grush 2004; Miall 2003): Instrumental output-result associations can then be activated predictively, simulating the effects of specific outputs for informational purposes.

Over time, associations are established between output and subsequent input in certain contexts, so that copies of motor-output signals can evoke associated input signals. An inner loop maps copies of output signals directly onto “expected” input signals – means to results (Fig. 2; new aspects italicized). *Forward model* is engineering terminology for this mapping; copies of output signals in organisms are called *effeference copy*. This subpersonal process simulates feedback – predicts the results of output on input. Prediction can occur during actual action, to smooth a behavioral trajectory by anticipating feedback, or prior to action, to provide information about alternative possible actions (layer 4).

A general improvement in instrumental control results. A control system with predictive simulation no longer need await actual feedback. A thermostat can turn the heat down before the target is reached, avoiding overshooting; hand movement can be initiated in accord with predictions of retinal signals based on eye movements. When real and simulated results don't match, a local switch can default back to actual feedback control while predictive simulations are fine-tuned to improve subsequent predictions.¹³ For purposes of online instrumental control, the system need not monitor continuously or access globally whether it is using actual or simulated feedback.

Comparisons can now be made not just between targets and actual results of action, but between the latter and anticipated results. This permits refference to be distinguished from exafference: Information about the organism's goal-directed behavior can be distinguished from information about environmental events. Consider the familiar ambiguity: When my train pulls out of the station, I register movement relative to the train on the next platform, but this does not itself provide information about whether my train or the neighboring train is moving. Comparison of predicted feedback from action (effeference copy) with actual feedback provides resources to resolve an analogous subpersonal ambiguity (between refference and exafference), and hence to distinguish the self's activity from environmental events.¹⁴ This subpersonal information can contribute to enabling the personal-level distinction between action of the self and perception of the world. The perception/action distinction emerges from subpersonal informational dynamics between world-to-animal inputs and animal-to-world outputs. However, perception and action don't map, respectively, onto input and output (Hurley 1998). Rather, layer 2

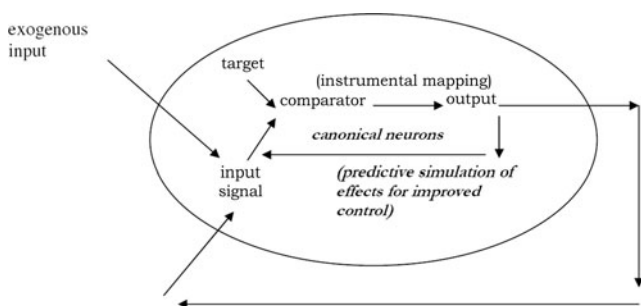


Figure 2. Layer 2: Simulative prediction of effects for improved control.

inherits unsegregated information about inputs and outputs from layer 1, and uses this blended information in enabling the perception/action distinction. Perception and action share these basic informational dynamics and processing resources. SCM thus provides a dynamic process version of a common coding view of perception and action (sect. 2.2).

However, the system does not yet provide information about similarities between the agent's actions and actions by others, nor information distinguishing the agent's actions from similar actions by other agents (as opposed to distinguishing the agent's actions from environmental events in general). This suggests that a basic distinction between action by self and perception of world, associated with instrumental control functions, can be available to creatures still lacking intersubjective information, a self/other distinction, or mindreading abilities. There are more and less fundamental layers of information about self (*self* in SCM is neutral between persons and other animals).

Some general predictions derive from layer 2. First, neural mechanisms that implement sensorimotor affordance associations (such as canonical neurons [sect. 2.2]) are predicted. Suppose an animal typically acts in ways afforded by certain kinds of object: for example, eating a particular food in a specific way. Copies of motor signals for eating movements will be associated with a multimodal class of exafferent and refferent inputs deriving from such objects and the agent's eating of them. Cells mediating this sensorimotor association could thus have both sensory and motor fields and carry information about objects' affordances. Second, deficits in predictive simulation functions should be associated with deficits in distinguishing self from world and action from perception.¹⁵ A specific first-order prediction relating to layer 2 might be that capacities requiring information that distinguishes self from world are phylogenetically prior to capacities for social learning and action understanding.

3.3. Layer 3: Mirroring for priming, emulation, and imitation

SCM next postulates that instrumental output-result associations can be activated bilaterally, from effect to cause, as well as from cause to effect. Not only do copies of motor signals predictively simulate input signals (layer 2), but input signals can evoke corresponding motor signals.¹⁶ To put it more technically, not only does effeference copy produce simulated refferent input in forward models, but input signals can evoke mirroring effeference or motor output. Mirroring in effect runs the predictive simulations of forward models in reverse (Fig. 3).

Observed actions are thus mirrored in the observer; if mirroring is sufficiently strong and not inhibited, overt copying results. (*Mirroring* is here a functional subpersonal rather than neural description of behavioral priming produced by observing action, and may be implemented in neural mirror systems.) Mirroring within specific control structures of differing grain and structure would enable different copying capacities, observed across various social species: mirroring of basic movements in priming (Rizzolatti's “low-level resonance”), mirroring of goal-directed action or emulation (Rizzolatti's “high-level resonance”), and even full-fledged imitation (if mirrored

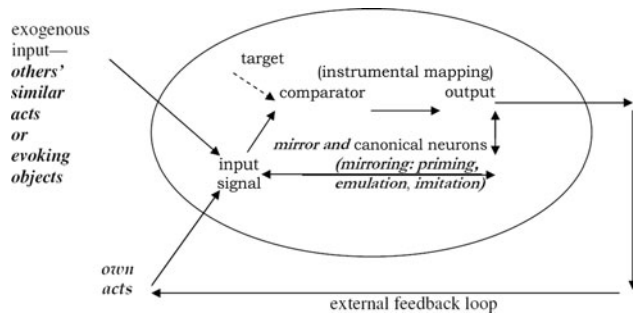


Figure 3. Layer 3: Mirroring for priming, emulation, and imitation.

elements of control structures are sufficiently articulated and flexibly linked to provide the information needed for social learning of novel means to a goal [see further on in this section]).

Note the intimate relationship between the sharing of circuits for self and other and for action and perception: Layer 3's shared informational dynamics for intersubjectivity presupposes layer 2's shared informational dynamics for perception and action, which builds on layer 1's generic informational dynamics for sensorimotor control. SCM explicitly builds shared resources for self and other on those for action and perception. It thus integrates W. Prinz's shared information for perception and action (though in terms of functional dynamics rather than coding) with Gallese's primitive intersubjective "we"-centric information. By bringing together information about motor causes of one's own and others' similar observed actions, mirroring enables simulation of means/ends associations from either direction: Observed action retrodicts motor activation in the observer via mirroring of causes, which are associated with further results via simulative prediction of effects. But although mirroring makes information about action's instrumental structure accessible bilaterally (from acting and from observing others act), mirroring does not yet distinguish own action from observed action. At layer 3, self and other share informational resources: intersubjective information is subpersonally prior to the self/other distinction (as Gallese holds [sect. 2.3]).

Instrumental mapping and mirroring both map input to output; SCM distinguishes them functionally (neural implementations may overlap). Instrumental mappings have control functions: given certain inputs, they select motor outputs that in turn produce inputs matching a target. Although mirroring exploits instrumental control structures and also produces motor outputs, given certain inputs, it does not itself select outputs as means to inputs that match observed action – or any other target (cf. Peterson & Trapold 1982). Rather, SCM postulates, mirroring is a by-product (via reversal) of predictive simulations, which do have instrumental control functions. However, the resulting automatic copying tendency has evolutionary functions, and copying can be exapted for cognitive functions associated with imitation, action understanding, or signaling; these in turn can enable advanced social ("Machiavellian") forms of instrumental control (see sects. 2.1 and 2.3). Whether specific mirroring capacities are adaptive depends on their potential

functions for different social species under different evolutionary pressures.

How could mirroring arise? Consider first movements that produce visual reafference. When creature A sees her own hand movements, associations form between copies of motor signals for these movements and visual reafference from these movements. Cells mediating this association can acquire congruent sensory and motor fields. These cells would also fire if creature A receives similar visual inputs from creature B's similar hand movements; the cells wouldn't distinguish observer's action from observed action producing similar inputs. So, like mirror neurons, they would fire both when A acts and when she observes such similar action by B. They mediate associations between copies of motor signals and a class of inputs including both characteristic reafference from the agent's movement and similar exafference from others' similar movements.

How could mirroring arise for movements not seen by their agents? This requires a correspondence between one's own and others' similar acts, without refferent feedback from own acts in the same modality as observations of others' acts. Facial movements normally produce proprioceptive rather than visual feedback: How can one's own facial movements be associated with visual information about similar observed facial movements, to enable copying of facial expressions?

Several answers are possible (sect. 2.2), all compatible with SCM. Some supramodal correspondences may be innate (as in newborn copying; Meltzoff 2005; Meltzoff & Moore 1997). Some may be acquired through experience with mirrors, or with being imitated (Heyes 2005). Some could be established via stimulus enhancement, as follows. Suppose a creature repeatedly sees conspecifics act a certain way; their actions draw its attention to the typical objects of their actions, which evoke in the observer an innate or previously acquired response. An indirect association results between visual observations of others' actions and one's own similar action. This isn't copying initially; the object independently evokes others' and own similar acts. But the indirect, object-mediated association between own and others' similar acts may become direct with repetition. Cells mediating it could thus acquire similar sensory and motor fields, so that observing another's act primes similar action by the observer. Therefore, stimulus enhancement could develop into copying, and an indirect link via an enhanced stimulus into a direct sensorimotor mirroring link.¹⁷ SCM is neutral about whether such correspondences are innate, acquired, or both.

SCM does not describe one all-inclusive structure, but has multiple instances for specific movements and results, at various points along different means/ends chains (cf. Fogassi et al. 2005; Wolpert et al. 2003). The ends/means distinction is relative and applies along spectra of means/ends links in which basic movements are means to proximal results that are means to more distal results. Such control spectra can vary in grain. SCM could apply at successive points along a control spectrum, or between spectra (with the right neural connections); one circuit's target could be the means to next circuit's target. A network of control spectra could support hierarchical control and flexible recombination

of means and ends. With mirroring added, it could enable movement priming, emulation, or imitation.

SCM predicts that capacities for specific kinds of copying and social learning will vary across species and development with

- (A) the grain and complexity of instrumental control capacities, and
- (B) which of these have associated mirroring functions and how richly and flexibly they are linked (see discussion following; Csibra 2005).

How do these two influences operate?

(A) Different animals are equipped to different degrees with capacities for instrumental control and associated predictive simulation. This variation reflects potential means/ends chains of differing grains and lengths and with differing degrees of lateral connectivity to enable novel combination of means and ends as instrumentally appropriate. Animals suitably equipped with control functions by evolution and development could form chains of simulative predictions, resulting in information such as: this tail movement has that effect on weight over legs, facilitating this movement trajectory in relation to that gazelle, and so on – with eating the prey later in the chain. A similar chain could lead from knee movement to winning a slalom race.

Mirroring operates such instrumental associations in reverse. Mirroring the cause of another's movement, or resulting relationship to an object, could enable movement priming, goal emulation, or even full-fledged imitation (if instrumental control and associated mirroring functions are sufficiently articulated and flexible). Combined with further information distinguishing self from other (layer 4), simulative mirroring can provide information to enable understanding of others' observed movements as instrumental actions with intentional, means/end structure. Simpler control and predictive simulation capacities, with shorter, coarser, means/ends chains, limit correspondingly an animal's potential for mirroring and related functions.

Whether observed movement primes copying or is recognized as goal-directed depends on, *inter alia*, the instrumental capacity potentially available for mirroring and simulative functions. Mirroring and simulation might provide information about the goals of certain observed movements, given fine-grained, complex means/ends associations, but not given coarser control capacities. No doubt a monkey can move her hand to grasp a piece of sushi and move it to her mouth to eat it. But I can move my hand to operate chopsticks to pick up sushi to dip it in soy sauce and then move it to my mouth to eat it, in order to impress my boss; given associated simulative mirroring functions, I may start to resent you for eating the last piece of sushi as soon as you reach for your chopsticks.

(B) While mirroring of instrumental associations can provide information about the instrumental structure of observed movements, mirroring does not itself determine which instrumental associations, or predictive simulations thereof, are in place and hence potentially exploitable by mirroring. Control processes can be neurally distributed, with components of an articulated mean/ends chain processed in different brain areas – control of fine movements versus gaze versus posture versus whole body movement versus external objects. Some such neural areas may have mirroring as well as predictive simulation capacities, whereas others do not. Whether mirroring is

associated with particular control structures will vary across species and development, as a result of evolutionary and developmental processes, yielding different capacities for copying and for generating information about goals of observed actions.

We can understand the enabling of movement priming, goal emulation, or imitation in terms of mirroring that exploits different control structures. Mirroring associated with more basic means/ends links (Rizzolatti's 'low-level resonance') predicts priming of basic movements; mirroring associated with less basic means/ends links predicts priming of less basic movements or results of more basic movements. Finger movements are means to chopstick deployment; when I can control chopsticks, chopstick deployment is a means to sushi eating, which could be a means to some further social result. Suppose I eat sushi by deploying chopsticks by moving my fingers a certain way. You watch. If seeing me move my fingers generates mirroring motor activation, then you are primed to move your fingers similarly. Such movement priming predicts interference if you watch me doing X while you are doing Y.

If you can already control chopsticks, less basic mirroring and prime chopstick deployment can occur. Such priming could be goal-mediated: Your chopstick deployment could mirror mine when sushi-eating results (even if no sushi is visible), but not when the results are unrelated to sushi (cf. Umiltà et al. 2001). Goal emulation could be enabled by mirroring midway along a control spectrum of proximal results of movements (rather than basic movements themselves) that are in turn means to more distal results (cf. Rizzolatti's 'high-level resonance'). Such midway mirroring would generate motor activation associated with a corresponding midway goal for the observer (rather than with similar basic movements in the observer). As well as enabling emulation, this would contribute information for understanding observed action as goal directed (layer 4).

The phylogenetically rare capacity for *imitative learning* requires flexible means/ends associations; priming and emulation, respectively, provide its ends and means components (sect. 2.1). Articulation within and linkages between mirror circuits determine whether mirroring enables imitative learning. An animal with mirror circuits along a chain of means/ends associations may never have used a certain means to a given goal. But if an observed novel means to that goal is mirrored, and neural links permit specific mirroring activations to be flexibly combined with targets, that goal and those mirrored means may be newly associated, capturing information about novel instrumental structure in observed action and enabling imitative learning. You might learn to use chopsticks by watching me; emulation plus trial and error would be bettered by mirroring that primes your finger movements towards those you see me make that are associated with the target of chopstick deployment.

Children's greater imitative tendencies, compared with chimps, may depend not on the presence versus absence of a mirror system, but on articulated relationships among multiple mirror circuits, permitting recombinant structure in social learning (sect. 2.1; cf. Barkley [2001] on recombinant structure in executive functions and their relation to imitation). SCM predicts correctly that

imitation should be found in fewer species than is movement priming or emulation separately, because imitation additionally requires linkages supporting flexible instrumental associations between mirrored means and ends at a relatively fine grain. Flexibly linked mirror circuits could also generate behavioral building blocks combined in *program-level imitation* of sequential or hierarchical structure. And they could allow infants to form three-way associations among observed behavior by parents (who have survived to reproduce, so may have adaptive behaviors, not all of which are heritable), observed circumstances of such parental behavior, and infants' own similar behavior, enabling *contextual imitation*: act like *that*, when the environment is like *this*. A network of linked mirror circuits could also permit mirroring activation to spread and generalize automatically, as in *chameleon effects* (see sects. 2.1 and 2.3).

Some general predictions derive from layer 3. First, neural systems are predicted that implement mirroring functions and don't distinguish own and others' actions. Mirror neurons need not *individually* implement mirroring and simulation functions. Specific hypotheses might concern how mirror neurons with varying sensorimotor congruence contribute to the functions of mirror systems (cf. Csibra 2005). Second, functional associations are predicted between deficits in mirroring and in predictive simulation for control, which reflects shared circuits for these functions. Third, implementational associations are predicted between neural mirror systems and neural systems for comparator control and predictive simulation (sect. 2.2). Fourth, automatic behavioral priming and copying tendencies are predicted at varying grains, reflecting the articulation and complexity of control functions across species and development (as in movement priming, emulation, human infant copying, human perceptual induction effects, imitative interference and reaction-time effects, and chameleon effects). Overt automatic copying tendencies should be greater where inhibition is weaker, that is, in young children or imitation syndrome patients (sect. 2.1). Fifth, the phylogenetic rarity of imitation as opposed to movement priming and emulation is predicted (sect. 2.1).

Another thought: Recall that canonical neurons mediate sensorimotor affordance associations; for example, between copies of motor signals for eating and a class of inputs associated with objects that afford eating and the eating of them. As a result of such associations, observing an object may prime action it affords and produce a tendency to automatic action on affordances where inhibitory function is reduced – in young children, utilization syndrome patients, or subjects with attention deficit hyperactivity disorder (Barkley 2001, p. 15; Lhermitte 1983; 1986; Rietveld, in preparation). Whether inhibitory functions are specific to imitation or action on affordances is a further question (Brass et al. 2003).

3.4. Layer 4: Monitored output inhibition combined with simulative prediction and/or simulative mirroring

Information about the instrumental structure of observed action provided by a flexibly articulated network of mirror circuits can not only enable imitative learning, but also contribute to enabling the understanding of another's actions as instrumentally structured, including in novel

or complex ways. SCM's layer 4 introduces the capacity to inhibit actual output and monitor this inhibition while instrumental associations are activated. This capacity for monitored inhibition could combine with layer 3's mirroring to enable action understanding. Or, it could combine with layer 2's online predictive simulations to enable offline instrumental deliberation. Or both.

Take instrumental deliberation first. Layers 2 and 4 could combine functionally to distinguish actual from possible actions. At layer 2, simulative predictions improve online control of ongoing action. This online function does not require the control system to monitor whether it is currently using actual or simulated feedback, as long as it can switch between them as needed to achieve the target. However, simulative predictions of results could also function offline, with actual motor output inhibited. Multiple simulative predictions could provide information about results of alternative possible actions, rather than anticipating results for ongoing action. Simulated results of alternative possible actions could be compared for the closest match to a target prior to actual action. Layers 2 plus 4 could thus provide information for "trials and errors in the head" (Millikan 2006) prior to actual trials and possibly fatal errors, allowing simulations to die in a chooser's stead. They could thus enable counterfactual instrumental deliberation and choice among alternative possible actions.

Enabling these capacities requires more than comparing simulated results of different acts with a target. It also requires monitoring whether motor output is inhibited, to track the distinction between actual and possible actions. The predicted results of actual actions will call for a very different response to the predicted results of possible actions. A creature unable to distinguish these two types of predictions wouldn't be long for this world. (Barkley 2001). Layer 4's monitored inhibition provides a basis for this distinction: Simulated results with output inhibition provide information about possible actions, whereas simulated results without output inhibition provide information about actual actions. Multiple predictive simulations provide information about the consequences of various actions by the agent, whereas monitoring of output inhibition provides information that these are possible actions, not actual ones.

Whether such a subpersonal informational structure corresponds directly to the personal-level sense of ability to do otherwise – of having alternative possible actions open to choice – is a further question. The point here is to explain how information for an actual/possible distinction and counterfactual practical reasoning emerges in SCM. Although this information concerns agency, it may provide a basis, when combined with language, for counterfactual reasoning in theoretical contexts.

Now consider action understanding. Layers 3 and 4 could combine functionally to distinguish self from other – more precisely, to distinguish one's own action from another's. At layer 3, observing another's action primes similar action by the observer, through mirroring. At layer 4, the observer's similar action is inhibited; observed behavior isn't actually copied. Copying behavior can be beneficial (especially for young organisms), but unselective overt copying would often have disastrous results for copiers; a prey that chases predators won't survive for long. The capacity to inhibit copying is adaptive and

should be expected to evolve (Barkley 2001). Offline mirroring simulates in the observer the causes of observed action, reversing the direction of simulative prediction: Instead of simulating feedback that would result from motor activations, mirroring simulates motor activation that would produce results similar to those observed, with actual motor output inhibited. Simulative mirroring can thus provide information for understanding the instrumental structure of observed actions. In effect, offline copying enables action understanding (Fig. 4).

Mirroring of means/ends associations for observed actions isn't enough to enable understanding action as another's. This also requires monitoring as to whether motor output from mirroring is inhibited, to separate information about others' actions from information about one's own. Layer 3 doesn't do this job. But use of information about actions to understand others has different consequences, and makes different demands on subsequent behavior, than does use of information about actions actually to copy them. Therefore, it would be adaptive to track the distinction between own and others' actions. Layer 4's monitored inhibition provides an informational basis for this distinction, which overlays the shared informational dynamics for own and others' actions at layer 3: simulative mirroring with monitored inhibition provides information about another's action, not one's own. Thus, simulative mirroring can provide information about the causes and instrumental structure of observed action, while monitoring of output inhibition can provide information that such actions are another's, not one's own. This is how information for a self/other distinction emerges in SCM.

The length and grain of chains of means/ends associations and the flexibility of linkages between them should affect not just the types of copying enabled by mirroring (as explained at layer 3), but also the types of action understanding enabled by simulative mirroring. Goal-mediated simulative mirroring would provide information about goals of others' movements, enabling an early stage in understanding others as acting on intentions (hence, in mindreading). If mirror circuits are sufficiently articulated and flexibly recombinant to enable imitative learning, then monitored inhibition of imitative mirroring would capture the means/end structure of novel observed action more fully and flexibly, enabling more sophisticated mindreading. Depending on the articulation of control structures

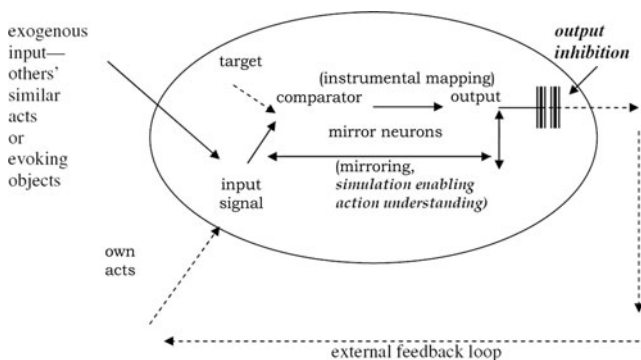


Figure 4. Layer 4: Simulative mirroring (or prediction) combined with monitored output inhibition, enabling action understanding (or instrumental deliberation).

with associated mirroring, adding layer 4 to layer 3 could enable different mindreading abilities.

Recall the discussion (sect. 2.3) of imitation-first versus understanding-first views. SCM reconciles them, by showing how different types of copying and action understanding can be built up, enabled by differently articulated mirroring structures combined with monitored inhibition. Mirroring can enable goal emulation without a capacity for imitation; with monitored inhibition, it can enable understanding the goals of another's action. Instrumentally articulated imitation may require understanding of goals plus mirroring of movements; covert imitation can then enable fuller understanding of the instrumental structure of observed actions.

Whether such informational structures correspond directly to personal-level empathic understanding of another's actions or to knowledge of other minds are further questions. Layer 3's "first person plural" is informationally prior to layer 4's self/other distinction. Subpersonally, the problem of "knowledge" of other minds is reconfigured: it is neither one of starting from information about self and constructing a bridge to information about others, nor one of starting from information about others and from these resources generating information about self. Rather, the similarity of own and others' acts comes first, with mirroring. Monitored inhibition then distinguishes subpersonally between instrumentally structured action "centered" on self versus "decentered" onto another; self-centering and other-centering of agency arrive together. SCM gives concrete subpersonal form to the interdependence and parity of information about self and other intentional agents. The subpersonal job that remains is not to bridge a gap between self and others, but to track distinctions among them, especially when multiple other agents are in play.

Is the subpersonal priority of intersubjective information reflected at the personal level, and if so, how? Is the personal-level problem of knowledge of other minds similarly reconfigured, to avoid both first-person to third-person and third-person to first-person inferences, that is, both the argument from analogy and from behaviorism? Not necessarily. Further questions arise, about phylogeny, development, the structure of mature capacities to understand others, and the epistemology of such understanding. The role of subpersonal information in the epistemology of other minds raises general philosophical issues about the roles of reliable information and justification in knowledge: for instance, can reliable subpersonal information support knowledge of other minds? SCM does not in itself answer these questions. Rather, it provides generic, adaptable tools for framing specific hypotheses. Care is needed: We shouldn't assume an isomorphic projection from the subpersonal level to the personal level. But we should allow that understanding subpersonal processes can contribute to understanding the personal level; this doesn't require interlevel isomorphism.

Gordon's simulation theory (sect. 2.3) suggests responses to such questions. It has affinities with SCM: both build mindreading on resources for perception and action. But there are also differences: In SCM, the priority of "first-person plural" information over the self/other distinction is subpersonal, whereas Gordon's account of how constitutive mirroring "multiplies the first person" links

subpersonal mirroring directly to personal-level understanding of other minds: Observed behavior and mirroring are made sense of together, under a common scheme of reasons, and incoherent mental states assigned to different persons (Gordon 1995a, pp. 56–58, 68; 2002; 2005; cf. Hurley 1989; 1998).

Actual/possible and self/other distinctions are necessary for much explicit theorizing and for aspects of the normativity and intersubjectivity of the personal level. SCM thus helps to understand how features of the personal level could be informationally enabled by subpersonal processes. It suggests that these distinctions share an informational basis in monitoring of motor inhibition: theoretical informational resources arise from practical ones. But SCM does not specify a phylogenetic or developmental priority between the subpersonal actual/possible and self/other distinctions, which is left to specific first-order hypotheses: Layers 2 and 4 could combine even if layers 3 and 4 do not or if layer 3 is missing altogether; and layers 3 and 4 could combine even if layers 2 and 4 do not. Hence, in different species or stages of development, we might find creatures with action understanding but not instrumental deliberation, or vice versa.

Some general predictions derive from layer 4. First, there could be deficits in inhibitory capacities although capacities to copy, or to act on affordances, are intact (as in Lhermitte’s imitation and utilization syndrome patients). Second, the priority of mirroring to inhibition of copying predicts that people with intact inhibitory capacities should nevertheless retain underlying default tendencies to copy (as in perceptual induction, imitation interference effects, priming, and chameleon effects; see sect. 2.1). In addition, SCM provides generic, adaptable tools for framing specific hypotheses. A specific first-order hypothesis might be that capacities for instrumental deliberation and for understanding others’ actions are enabled by shared inhibitory resources at layer 4, predicting that deficits in these capacities should be associated (Frith 1992, pp. 81–83, 93; cf. Barkley 2001). Another first-order hypothesis might predict pathological dissociations of capacities supported by layers 2 plus 4 from those supported by layers 3 plus 4. Other specific first-order predictions might concern the interleaving of various copying abilities with various types of understanding of others’ actions, in an ontogenetic or phylogenetic progression leading to imitative learning and mindreading capacities. Copying without inhibition should come earlier in such progressions; capacities to inhibit copying and to distinguish others’ intentional actions from one’s own should develop together. Yet another hypothesis might be that adding layer 4’s monitored inhibition of output to layer 3’s mirroring enables a transition from behavior copying to mindreading that enables effective use of mirror heuristics by cooperators (sect. 2.3).

3.5. Layer 5: Counterfactual input simulation

Finally, the system can be taken offline for input as well as output (Fig. 5). Counterfactual inputs can simulate different possible acts by others and their results. Monitored simulation of inputs to control circuits with simulative prediction and mirroring functions can provide information distinguishing between *others’* actual and possible acts. This social extension of counterfactual information,

combined with simulation of different possible acts by self and their results, provides information about how possible acts by self could result in possible acts by others with further results, and vice versa. These combined functions provide enabling information for strategic game-theoretic deliberation, coordination, and cooperation. To distinguish possible from actual acts by others, information is needed about whether inputs are simulated; so simulation of inputs must be monitored. Simulative informational resources for strategic in addition to instrumental deliberation enrich the practical foundations for more general capacities to manipulate counterfactual information and theorize counterfactually.

Recall the discussion of TT versus ST in section 2.3. Again, SCM shows how they can be reconciled. Differentiating and tracking interacting means/ends relations for multiple possible acts by self and by multiple other agents make acute informational demands. Despite their foundational role, simulation mechanisms may be insufficient to provide information for such multi-agent, multi-possibility tracking with ramifying paths of decentering. Meeting these demands, and further demands in differentiating the epistemic states of multiple others, probably requires SCM’s practical simulative informational basis for understanding other agents to be supplemented by language-dependent functions and theorizing. Mindreading, like social learning and instrumental control, is a graded phenomenon, not all or nothing (Tomasello 1999). Language can build on SCM’s foundational actual/possible and self/other distinctions to enable interpretative understanding of multiple others with multiple alternatives and varying beliefs. SCM hypothesizes that mindreading has practical foundations, in simulative mirroring of means/ends relations, but allows that mature mindreading with all the bells and whistles (including understanding false beliefs) requires both simulation and language-based theorizing.

Specific predictions deriving from level 5 could concern ontogeny or phylogeny. First, since understanding goals and action is foundational for mindreading, it should be prior to understanding others’ epistemic attitudes (on phylogeny, see Tomasello & Call 1997; on ontogeny, see Rakoczy et al. 2007). Second, understanding the instrumental structure of observed action can precede, and may be a phylogenetic precursor of, understanding linguistic structure. The instrumental recombinant structure of imitative learning may combine with learned manipulation

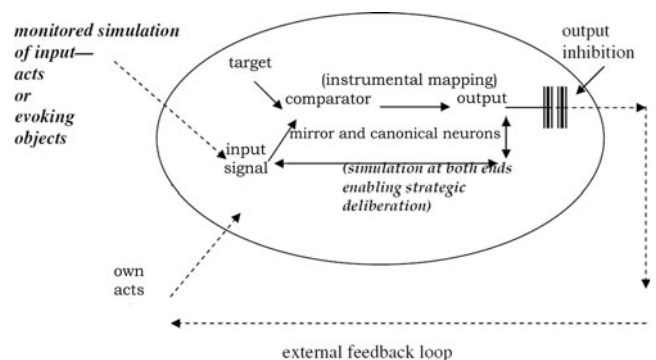


Figure 5. Fifth layer: Counterfactual input simulation enabling strategic deliberation.

of external symbols to support the richer recombinant structure of language. Third, the fundamental transition from agents “playing against nature,” instrumentally, to agents playing against one another, strategically, can be supported by simulative mechanisms. Nevertheless, simulative mechanisms without linguistic capacities may not enable advanced mindreading, such as multi-person strategic deliberation or false-belief attribution.

4. Conclusion

SCM describes a functional subpersonal architecture, at a level above that of neural implementation but below that of the conscious and/or normative contents of persons’ mental states. I now step back from the details and review the functional relations among SMC’s layers:

Layer 1: SCM’s starting point is *dynamic online motor control*, whereby an organism is closely attuned to its embedding environment through sensorimotor feedback.

Layer 2: Next, add *online predictions of sensory feedback from ongoing motor output*. Online predictive simulation improves instrumental control and provides information distinguishing action by the agent from perception of the world (Table 1).

Layers 2 + 4: Combining *predictions of feedback* with layer 4’s capacity for *monitored inhibition* of output has further benefits. Offline predictive simulation distinguishes actual from possible acts and provides information about results of alternative possible acts, for offline/counterfactual instrumental deliberation.

Layer 3: *Mirroring* reverses layer 2’s predictive associations, so that observing movements generates motor signals in the observer that tend to cause similar movements. Various mirroring structures can enable various forms of copying, with various functions. If mirroring preserves novel means/ends structure of observed actions, it can enable imitative learning. But mirroring provides intersubjective information in a subpersonal “first-person plural,” without distinction or inference between own and others’ similar acts.

Layers 3 + 4: Combining *mirroring* with *monitored inhibition* of overt copying does distinguish own from

others’ acts. When the goals or means/ends structure of observed actions are mirrored, this combination provides information for various levels of understanding of another’s action.

Note that layer 4’s monitored inhibition can combine independently with prediction of effects or mirroring of causes (or both), providing information for various capacities: see Table 1.

Layer 5: SCM’s last layer adds the function of monitored simulation of input specifying possible observed actions. This extends counterfactual information about actions socially, providing information about possible acts by others. This function combined with inhibited mirroring (layers 3 plus 4) of possible actions can generate information about possible (as opposed to actual) actions by others (as opposed to self), and possible causes and effects of such possible actions. Linguistic and theoretical resources can be added to simulative foundations, thereby enabling deft manipulation of combined actual/possible and self/other distinctions and tracking of interacting means/ends relations among multiple possible acts by self and multiple others. Strategic social intelligence is thus enabled, whereby agents can play against one another rather than merely “against nature,” a nonagent.

Specific instances of SCM layers could combine into sequences, hierarchies, or networks permitting flexible decomposition and recombination of particular links. The numbering of layers is largely heuristic and does not necessarily represent the order of evolution or development. First-order hypotheses can map the layers onto specific phylogenetic or ontogenetic progressions, and the combination of layers can vary across particular empirical applications. For example, in the absence of layer 3’s mirroring function, layers 2 and 4 could combine to provide information about results of different possible actions. Nonnegotiable features of SCM are its explanation of mirroring as an exaptive reversal of online prediction, and the way the actual/possible and self/other distinctions arise as online processes are overlain by monitored inhibition (on impulsiveness as default, cf. Barkley 2001, pp. 5, 22).

SCM shows how information for important cognitive capacities of persons can have a foundation in the dynamic co-enabling of perception and action. Its layered build-up of functions illustrates a horizontally modular architecture, in which rich cognitive resources emerge without a classical sandwich. Specifically, materials for active perception can generate cognitively significant resources: the action/perception, self/world, actual/possible, and self/other distinctions, intersubjective information enabling social learning and mindreading, and counterfactual information enabling instrumental and strategic deliberation.

4.1. Levels versus layers

Unlike some of the work surveyed earlier, SCM distinguishes neural, functional subpersonal, and personal levels of description. Each of SCM’s functionally described *layers* raises questions at the *level* of neural implementation, as well as providing information enabling personal

Table 1. *Online/offline by prediction/mirroring*

Shared circuits model’s (SCM’s) middle layers	Prediction	Mirroring
Online	Layer 2: Online instrumental control	Layer 3: Copying, including imitation
Offline (with monitored inhibition)	Layers 2 + 4: Counterfactual instrumental deliberation (actual acts vs. possible acts)	Layers 3 + 4: Action understanding (own acts vs. others’ acts)

level capacities (Table 2). Clarity and progress are served by distinguishing levels and framing issues explicitly at a given level, or as concerning interlevel relations. Sliding between levels on a priori assumptions of isomorphism is unjustified. Nevertheless, one level can shed light on another. We can look “down” a level, seeking neural implementations of aspects of SCM’s functional architecture, or “up” a level, considering what SCM’s functional architecture would enable persons to do.

4.1.1. Looking down. SCM’s implementation is distributed across neural processes and embodied activity in environments, especially social environments. It predicts neural systems mediating affordance and mirroring functions, and has a heuristic role in generating specific first-order hypotheses concerning the following:

The location of postulated comparators and simulators.
 The division of neural labor in mirroring ends and means and in inhibition.
 The role of mirror neurons in Broca’s area and their relation to linguistic capacities.
 How the compositionality of imitation relates to the compositionality of language.
 While SCM is described cybernetically, dynamic systems theory could represent interactions of its implementing neural processes and embodied activity over time as evolution of a phase space, and investigate its attractor structure.

4.1.2. Looking up. SCM explains how distinctive features of the personal level can be informationally enabled. It provides theoretical resources for addressing further questions concerning the following:

Table 2. *The shared circuits model: Layers and levels*

Interlevel relations		Layer 1	Layer 2	Layer 3	Layer 4	Layer 5
	Personal/ animal level:	Adaptive motor control, maintain target against disturbance	Instrumental action, self/world and action/perception distinctions	Social learning, behavior copying (movement priming, emulation, imitation); intersubjective empathy; automatic priming & copying & interference effects; chameleon effects	Deliberation about own possible acts understanding others’ instrumental actions, actual/possible and self/other distinctions	Deliberation about others’ possible acts, strategic social reasoning about own and others’ possible acts
	Subpersonal functional level: SHARED CIRCUITS MODEL	Comparator feedback-control system	Simulative prediction (from cause to effect), smooths & speeds instrumental control, shared information for perception and action	Mirroring, shared information for own and others’ actions	Simulative prediction (from cause to effect) and/or simulative mirroring (from effect to cause), with monitored inhibition of output	Monitored simulation of input
	Subpersonal neural level:	Various neural comparator systems, exafference vs. reafference	Efference copy, neural systems for sensorimotor affordances (canonical neurons)	Neural mirror systems	Neural inhibitory and monitoring mechanisms	*? Neural imagery mechanisms

*JK&AC: Why the question mark? We think that Susan meant to indicate that the search for the neural basis of layer 5 might begin with mechanisms that underpin imagery, but that this was just the beginning. Just how the brain might go about simulating inputs was one of the many open empirical questions that SCM raises, which Susan hoped might be taken up in the commentaries or in future work developing SCM.

- Relations among distinctively human capacities for imitation, deliberation, mindreading, and language;
- Relations between personal-level action/perception, self/other, and actual/possible distinctions, and whether they reflect subpersonal structure.
- How SCM's layers map onto evolutionary or developmental stages.
- Relations between subpersonal simulation of possible actions and the personal-level sense of being able to do otherwise.
- Whether knowledge of other minds requires first-person to third-person inference or can bottom out in reliable subpersonal information at layer 3.
- Whether the subpersonal priority of intersubjective information in SCM is reflected in personal-level epistemology of other minds.
- Whether social cognition is related to aspects of consciousness, given the roles of comparator and simulation structures in accounts of consciousness (Frith et al. 2000a; Gray 2004; Hesslow 2002; Jeannerod 1997; 2001; Milner & Goodale 1995, p. 64).
- Whether SCM can contribute to distinguishing conscious from unconscious processes.
- Whether SCM can extend from instrumental to expressive action, including facial expressions of emotion and emotional mirroring (cf. Adolphs 2002; Decety & Chaminade 2005; Gallese 2005; Iacoboni 2005; Preston & de Waal 2002¹⁸; Rizzolatti 2005).
- Whether extending SCM to expressive action can illuminate relations between social cognition and language or consciousness.

SCM has a cybernetic rather than conceptual structure, yet can provide information for cognitive skills – for example, mindreading; deliberation – with personal-level conceptual structure. It thus probes the kind of intelligibility to be found in explanations of how subpersonal resources enable personal-level capacities. By showing how subpersonal resources for cognition can build on those for active perception, SCM illustrates that informationally enabling subpersonal structure need not recapitulate personal-level conceptual structure in any explicitly isomorphic way. It is an empirical question, to be decided case by case, whether enabling subpersonal structure corresponds isomorphically to personal-level structure, conceptual or otherwise. Interlevel isomorphism should not be required or denied a priori. Personal-level content can remain distinctively conceptual and normative, while it nevertheless becomes intelligible how the minds of persons can arise from interactions of embodied brains with environments, including social environments.

In explanations of some systems' dynamical behavior, higher-level structure corresponds to lower-level structure. But the dynamical behavior of other, complex systems cannot be so explained: System behavior can result deterministically from nonlinear relationships among lower-level factors although its structure does not correspond to lower-level structure. Brain-body-environment systems are sufficiently complex and nonlinear that emergent structure without isomorphic lower-level structure should not surprise us.

I end by highlighting some noteworthy aspects of this model:

1. SCM avoids the traditional conception of cognition as sandwiched between separate perception and action

systems. Rather, it understands perception and action as enabled by shared subpersonal dynamics, and builds subpersonal resources enabling cognition on shared resources for perception and action.

2. Shared processing of actions by self and by others in social cognition is a special aspect of shared processing of action and perception in dynamic control. I perceive your action by means that engage my capacity for similar action, enabling me to copy or understand your action.

3. These shared resources are prior to self/other and actual/possible distinctions that provide information for action understanding and instrumental deliberation. The shared processing of action and perception in dynamic control is preserved when an actual/possible distinction is overlaid via inhibition of overt action. Similarly, the shared processes of action and perceiving others' action is preserved when a self/other distinction is overlaid via inhibition of overt copying.

4. The subpersonal basis of counterfactual deliberation and mindreading is simulation of instrumentally structured agency. Linguistic and theoretical resources can build on practical simulative foundations to enable more advanced counterfactual reasoning and mindreading.

ACKNOWLEDGMENTS

For helpful comments and discussions, I thank Mark Ashton Smith, John Bargh, Paul Bloom, Jeremy Butterfield, Nancy Cartwright, Nick Chater, John Cummins, Chris Frith, Vittorio Gallese, Philip Gerrans, Alvin Goldman, Robert Gordon, Jeffrey Gray, Rick Grush, Cecilia Heyes, Marco Iacoboni, Dorothee Legrand, Andrew Meltzoff, Alva Noë, Hanna Pickard, Joëlle Proust, Nicholas Rawlins, Simon Saunders, Marleen Schippers, John Schureman, Nicholas Shea, Evan Thompson, members of various audiences, and anonymous referees.

NOTES

1. Enabling subpersonal dynamics can include bodily or environmental loops, as well as neural processes (on vehicle externalism, see Hurley 1998; in press).

2. Like Dennett (1991) and Millikan (1991; 1993), I am wary of projecting properties or structure between the personal and subpersonal levels. Like McDowell (1994), I use "enable" for a making-possible relationship between subpersonal and personal levels and deny interlevel isomorphism requirements (such as a language of thought requirement). But I allow that enabling explanations can sometimes contribute to personal-level intelligibility, and find philosophical interest in subpersonal explanations per se.

3. I use "emulation" in the well-established social learning theory sense (Call & Carpenter 2002; Tomasello 1998; 1999), not in the different sense used by Grush (1995; 2004). I, like many others, use "simulation" to include "emulation" in Grush's sense.

4. Cf. Povinelli (2000, Ch. 6), where chimps are not permitted trials and errors with rake orientations.

5. Paul Harris, at the conference "Perspectives on Imitation: From Cognitive Neuroscience to Social Science" held at Royaumont Abbey, France, May 23–26, 2002, suggested an experiment to assess how far monkey mirror neurons subserve action understanding. They fire when the monkey reaches for an apple or she sees the experimenter reach for it. The same mirror neurons fire when, after a screen has occluded the apple, the monkey sees the experimenter's hand reach behind the screen to where the apple is. They don't fire when the monkey sees that there is no apple before the screen descends and then sees the experimenter's hand reach behind the screen in the same way (Umiltà et al. 2001). These neurons thus code for the action's goal or result. Harris suggests a variation that

provides a neural “false-belief” test. Suppose both monkey and experimenter see a nut, and see the screen descend to occlude it. The experimenter leaves the room. The monkey is permitted to remove the nut. The experimenter returns and the monkey sees the experimenter reach behind the screen for the nut, which the monkey knows is no longer there. Will the monkey’s mirror neuron for “reaching for the nut” fire? This might suggest the neuron codes for a goal of nut grasping, since the experimenter “doesn’t know” the nut is no longer there. Or will it not fire, because the nut isn’t actually there, so nut grasping cannot result? That is, does the mirror neuron code for the intended goal of the observed action, or merely its result? Note that even chimps fail nonverbal false-belief tests (see Call & Tomasello 1999; Call et al. 2000; and cf. Hare et al. 2000; 2001).

6. Or its monkey homologue. Mirror neurons are also found in frontal area 6 (Buccino et al. 2001) and posterior parietal cortex, area PF (Miall 2003).

7. Byrne (1998; 1999; 2002a; 2002b; Byrne & Russon 1998) argues that this is found in gorilla food processing.

8. Howard (1988) and Danielson (1992) implement mirror heuristics computationally, avoiding computational loops and regresses. Danielson’s technique matches quotations of another player’s program and one’s own; others’ programs are read offline to determine a match, but not executed (Danielson 1992, p. 82ff).

9. Danielson’s self-same cooperators cooperate just with exact syntactic copies of themselves. He also discusses more flexible and selective metaheuristics (Danielson 1992, pp. 130ff, 140).

10. Danielson (1992, p. 45ff) runs Prolog implementations against one another in tournaments where a heuristic’s stability against invading strategies does not depend on same players meeting one another repeatedly (as Tit-for-Tat’s stability does). Similar points apply to Howard’s (1988) Mirror Strategy.

11. The mindreading needed for mirror heuristics is of intentions rather than beliefs – the ontogenetically and phylogenetically earlier capacity (Rakoczy et al. 2007; Tomasello 1999; Tomasello & Call 1997; Whiten 1997, p. 167).

12. Space disallows discussion of Goldman’s (1989; 1992; 2005) important version of ST, which shares this first-person to third-person feature. Arguments from analogy have been criticized, for example, for generalizing without warrant from one case (oneself) and on grounds that self-recognition by a subject requires a contrast with other subjects.

13. Cf. Wolpert (1997), Wolpert and Kawato (1998), Haruno et al. (2001), and Wolpert et al. (2003) on selection among forward models; Grush (2004) on continuous combination of real and simulated input; and Grush (1995) linking motor control and mindreading. SCM’s layer 2 comparator is similar to that in Gray’s account of schizophrenia (Gray 1991, pp. 11ff). Layer 2’s predictive mechanisms are similar to those Grush (2004) argues can be run offline to account for motor and visual imagery; his commentators discuss possible links to mirror neurons, but he doesn’t postulate anything similar to layer 3’s mirroring. See Note 5 on terminological differences.

14. See Hurley (1998, pp. 140ff). Frith explains schizophrenic auditory hallucinations as involving defects in self-monitoring via efference copy, that is, failure to distinguish perception from action via predictive simulation: “Brain structures responsible for willed actions no longer send corollary discharges to . . . parts of the brain concerned with perception . . . In consequence self-generated changes in perception are misinterpreted as having an external cause” (Frith 1992, pp. 93, 81–83).

15. See Frith (1992, pp. 81–83, 93). Gray’s comparator model of schizophrenia emphasizes failure to integrate memories of input regularities with ongoing motor programs, predicting close association of cognitive and motor disorders (Gray 1991, pp. 1, 11, 19).

16. Gallese and Goldman (1998, p. 498) suggest something like this reversal; thanks to Vittorio Gallese for discussion here.

Blakemore and Decety (2001, p. 564) suggest a related reversal more explicitly. Whether empirical evidence for this hypothesized reversal will emerge, and what neural mechanisms may underwrite it, are open questions. Perhaps co-firing associated with the forward model strengthens and unmasks backprojection. Cf. Oztop et al. (2005), whose model involves no reversal of forward models or mirroring, but rather prediction and gradient descent.

17. Heyes (2005) suggests a mediating role for words in acquired equivalence learning, though she allows that the third term can be nonlinguistic.

18. SCM is what Preston and de Waal call a “perception-action model” (PAM). They apply PAM to empathy and emotional expression, whereas SCM details the development of mirroring and simulation from instrumental control.

Open Peer Commentary

The relationship between conscious phenomena and physical reality in behaviour control: The need for simplicity through phenomenological clarity

doi: 10.1017/S0140525X07003135

Ralf-Peter Behrendt

MRCPsych, The Retreat Hospital, York YO10 5BN, United Kingdom.

rp.behrendt@btinternet.com

Abstract: How can “predictive simulation” – as a conscious phenomenon, related to goal imagery – be interchangeable with mirroring, which is an automatic response that, from a first-person perspective, enters awareness only after the act? The correspondence between perception of another’s action and execution of one’s similar action may be an example of a general perception-motor interface that maps perception onto behaviour or disposition towards action, without the need for simulation.

We think of a certain result and our muscles produce this result, though we did not really mean to do this act ourselves. The thought arouses the movement because it has previously been linked with the movement. A thought which has previously served as the stimulus to an act will tend to have this effect again, unless inhibited by some contrary stimulus. There is no need of a definite consent to the act, provided there is nothing present to inhibit it.

— Woodworth (1926, p. 528)

Efference copies are unlikely to play a role in the awareness of one’s movement. William James (1890) recognised that consciousness of movement is an afferent, not efferent, sensation; it is a consequence, not antecedent, of the movement itself. He thought that consciousness of muscular exertion is impossible without movement being effected somewhere. Thus, our awareness of movement is secondary to the actual occurrence of movement in the physical realm. Motor acts enter awareness not at the point of execution, but only as they are *perceived*. Hurley argues that “copies of motor signals predictively simulate input signals” (sect. 3.3, para. 1). Yet, if motor output were to be unconscious, how could an efference copy of motor output simulate sensory consequences of an action or produce imagery of action effects?

Hurley thinks that through association of an action with its result, we can anticipate the sensory effects of a motor plan. Certainly, what precedes goal-directed action is an “anticipatory image of the movement’s sensible effects” (James 1890). Woodworth (1926), too, recognised that voluntary action is preceded by imagination of some change to be effected, but he also saw

that such imagery of action effects could not amount to the *prediction* of action effects. The role of what Hurley calls “predictive simulation” may *not* be to provide feedback onto “input signals” but – by acting as a stimulus – to determine instrumental action in itself. Already James (1890) considered that thoughts about actions automatically predispose to these actions; indeed, that representations of movement induce the movements they represent. Elsner and Hommel (2001) suggested that perception of an event that resembles a known action effect can automatically activate the corresponding action. Furthermore, imagery of the “intended and expected action effects,” representing the anticipated goal of a voluntary action, can – by acting as a stimulus – elicit the response that produces the “to-be-expected effect” (Elsner & Hummel 2001). Thus, anticipating the consequences of an action may in itself serve as precipitant for action. Imagery of a goal can cause goal-directed action, inasmuch as perception of a goal can have this effect.

There may be no need for a comparator mechanism: Perception and imagery are constructed by attentional mechanisms already in a way that makes them sufficient determinants of motor behaviour in accordance with the situation. In order to determine a motor response, goal perception or imagery is contrasted with awareness of body movement and position (as constrained by reafferent proprioceptive and visual sensory input). Movement plans (*intentions*), which specify a target and the movement required to achieve it, are formed automatically – upon perception of salient events – in the posterior parietal cortex, following which they activate connected premotor areas, thus predisposing to a motor response (Colby & Goldberg 1999). The representation of a stimulus in the posterior parietal cortex allows premotor areas to determine the coordinates of an action in response to the stimulus, although this does not imply that an action will be produced (Colby & Goldberg 1999).

The significance of premotor cortex activation during observation of another’s purposeful action (Grezes & Decety 2001), which occurs even when the observed interaction is partly hidden from vision and can only be inferred, may be that it produces *response facilitation* (Rizzolatti et al. 2001), which explains our tendency to imitate an observed movement. During observation of goal-directed action, the observer may shift attention to the other’s explicit or implicit goal, whereupon the goal is perceived or imagined; which, in turn, primes motor activity instrumental for obtaining the goal. Action understanding was shown to involve activation in premotor areas – more so than the mere imitation of others’ actions (Rizzolatti et al. 2001). Therefore, understanding of observed action, presumably including understanding of others’ speech, involves dispositions to own action.

Many social behaviours are automatically induced by the perception of others’ actions (Ferguson & Bargh 2004). People tend to mimic gestures or adopt the accent of a conversation partner without being aware of this. Behaviours in a conversation partner that connote high or low status automatically induce people to adopt opposite postures that connote submissiveness or dominance, respectively (Ferguson & Bargh 2004). Conversational partners use automatic *interactive alignment* of their response dispositions on multiple linguistic levels – a process that is related to imitation. A speaker’s words, sounds, grammatical forms, and meanings activate matching linguistic representations in the partner and thus automatically influence the latter’s linguistic productions. Speakers automatically re-use linguistic structures that they have just perceived as listeners (Garrod & Pickering 2004).

Hurley suggests that mirroring reverses predictive simulation (Layer 3). Assuming that we simulate “effects of intended acts” (in the sense of anticipatory imagery of a goal for action), can we conclude that “causes of observed movements can be simulated” (target article, sect. 1, para. 9), too? Does perception of another’s movement produce simulation in the sense of conscious imagery of a goal or cause of another’s action, or does it in itself lead to an automatic disposition to act? Hurley considers

that, when observing another’s behaviour, the self’s matching response is part of the very perception of the other’s behaviour. In other words, “a copied motor pattern is part of the perception of another’s action, though overt movement may be inhibited” (sect. 2.3.5, para. 6). However, perception of others’ behaviour may not give us conscious access to others’ goals or motivations, but only indirectly through our automatically elicited behaviour, including verbal expression. Mirrored behaviour may impress after its execution as being in tune with others’ attitudes and pursuits, inasmuch as it will also appear to be conforming to the social situation and context in general.

Exposure to others’ traits and stereotypes *automatically* elicits patterns of behaviour and attitude in accordance with the primed traits or stereotypes, whereby subjects are “unaware of any influence or correlation between primes and their behavior” (sect. 2.1, para. 13). Indeed, “thinking about or perceiving action automatically increases, in ways participants are unaware of, the likelihood that they will perform similar actions themselves,” yet why should this process involve copying “at various levels of generality,” particularly the *copying of goals* (sect. 2.1, para. 14)? Social situations determine our behaviour rapidly and automatically. When we imitate behaviour patterns associated with others’ traits and stereotypes, we may do so as a way of seeking social approval, or to display social submission or dominance, although these are by no means explicit goals. Motivation of social behaviour is unconscious and, insofar as one can speak of goals, these are unconscious, too. The absence of explicit goals in patterns of social behaviour raises the possibility that perceived/imagined events or objects generally translate into behavioural dispositions or overt behaviour without preceding reference to goals, meaning that even instrumental behaviour is proximally determined by *imagination of some change to be effected*, not by explicit goals.

Finally, Hurley argues that mirroring, which produces a “similarity of own and others’ acts” (sect. 3.4, para. 9), is prior to the self/other distinction (Layer 4). “Monitoring of output inhibition” is required in combination with simulative mirroring in order to “separate information about others’ actions from information about one’s own” (sect. 3.4, para. 6). The assumption that “the origin of the *self/other distinction*” lies in “monitoring whether mirroring is inhibited” (sect. 2.3.5, para 9, emphasis in original) implies that self experience hinges on others’ presence and is a social phenomenon even in its most fundamental dimension. It also demands, rather awkwardly, that there should be a fundamental difference between agent/world and self/other distinctions. More parsimoniously, agent/world and self/other distinctions can be conceptualised on a more basic level of sensorimotor control, independently from mirroring (Behrendt 2005).

Mirroring cannot account for understanding action

doi: 10.1017/S0140525X07003147

Jeremy I. M. Carpendale^a and Charlie Lewis^b

^aDepartment of Psychology, Simon Fraser University, Burnaby, BC, V5A 1S6, Canada; ^bDepartment of Psychology, Fylde College, Lancaster University, Bailrigg, Lancaster LA1 4YF, United Kingdom.

jcarpend@sfu.ca c.lewis@lancaster.ac.uk

<http://www.psyc.sfu.ca/people/faculty.php?topic=finf&id=67>

<http://www.psych.lancs.ac.uk/people/CharlieLewis.html>

Abstract: Susan Hurley’s shared circuits model (SCM) rightly begins in action and progresses through a series of layers; but it fails to reach action understanding because it relies on mirroring as a driving force, draws on heavily criticized theories, and neglects the need for shared experience in our grasp of social understanding.

Hurley has addressed the challenging problem of getting from the subpersonal level to the personal – from casual processes to knowledge. We agree with her starting point being in action and her use of layers, but we see two types of problem with her approach.

One is the error of omission, in that Hurley endorses positions that have been extensively critiqued. She buys into a reconciliation of theory theory (TT) and simulation theory (ST), and thereby inherits the problems inherent in both. We, and others, have identified fundamental flaws in TT and ST (e.g., Carpendale & Lewis 2004; 2006), but they are still taken for granted and seem to be immune to criticism.

Second, in attempting to forge links between layers, Hurley makes a series of errors of commission. The layers are bridged by jumps that fudge the distinction between information and knowledge. At the level of causal processes we can talk about information in one sense of the word, such as a digital camera recording a scene. The camera contains information, but this is very different from the knowledge that a person might acquire through observing the same scene. It is knowledge that Hurley's model has yet to account for. And this is the start of her problems.

There are consequences of both types of error. Central to the transition between layers, for Hurley, is involved the miraculous shift from copying or “mirroring” to “action understanding.” The oft-repeated mantra is: “I perceive your action by means that engage my capacity for similar action, enabling me to copy or understand your action” (sect. 4.1.2, para. 4). That is, I see an action, I have a tendency to mirror it, and through this process I understand the action. We will set aside the problems of how mirror systems develop and whether they should be thought of as a source or an outcome of development, and instead point out that in most circumstances this would not work. For example, I see pointing, and I have a tendency to imitate it; that is, a tendency for my arm and index finger to extend. But this would not help in understanding the action.

Consider the following examples. At a picnic on the beach, my (Carpendale's) wife looked at me and pointed to a friend's bag. It was clear that she wanted to direct my attention, but why did she want me to look at the bag? Was there something special about it? Did she want me to do something with it, get something out of it? Give it to her? How was I to understand this gesture? If I copied, or mirrored, this arm movement, it would not help in understanding her action. A little more information about the situation might help. We were picking things up, preparing to walk further down the beach. Yet still I failed to understand her, although by now readers may be thinking of possible meanings. In fact, I had to ask her what she meant. It turned out that she knew that the bag was heavy and she wanted me to carry it for our friend. Given shared experience in helping others, it is, in fact, reasonable to expect that this gesture might be understood. This need for experience in shared routines in order to understand gestures is clarified by considering research with chimpanzees.

Although chimpanzees can follow an experimenter's gaze direction and pointing gestures and end up looking at a bucket where food is located, they do not *understand* the gesture (Hare & Tomasello 2004). Chimpanzees do not understand that the experimenter is trying to tell them where the food is located because they have never experienced such a form of cooperative activity with others; they have never had a conspecific indicate food for them. However, they do have plenty of experience with competition, and if the event is transformed into a competitive encounter – if they know they are competing for food with the experimenter – then if they see the experimenter reaching for a bucket they immediately know where the food is. So what is needed in addition to the action in order to understand that action? The animal needs some experience with that form of interaction. No amount of mirroring or copying, even if it was “offline copying,” would help the chimpanzee.

For Hurley, action understanding is a foundation for later “mindreading,” which she distinguishes from “mere behavior prediction” (sect. 2.3.3, para. 4); that is, “discovering another player's intention, not simply predicting his behavior” (sect. 2.3.3, para. 7). But what else is there to predict other than people's action, broadly conceived to include verbal action? The very term “mindreading” conflicts with her foundation in action, and to do this Hurley buys into the mysticism of contemporary theory and its dualist assumption that new substance is added in addition to action. She later acknowledges that mindreading has to be based on observation of behavior, but, if so, what is the means by which the transition to it occurs? We need an account of social understanding that does not rely upon the crumbling foundations of behaviorism or dualism, both of which split intentional activity into two parts.

Our point here is that nothing about the physical movement, whether it is mirrored or not, would help in understanding the action of pointing in various situations. Instead, what is needed is experience in similar sorts of routines. This is missing from Hurley's model. She does write about various forms of “action understanding,” so perhaps she has a loophole, and these might be graded in terms of their complexity. Her intuition about understanding others' actions might work in simpler situations such as seeing another reach to grasp something. But it might work in such cases only because of the shared experience in this form of action. This is overlooked in her model and then she encounters problems with the slightly more complex cases that we discuss. What are the simpler forms of action understanding and how can the transition from simpler to more complex forms be accounted for in Hurley's model? These are crucial questions. Although layering is a useful metaphor to help us envisage these questions, Hurley's model does not help bridge these layers.

Can the shared circuits model (SCM) explain joint attention or perception of discrete emotions?

doi: 10.1017/S0140525X07003159

Bhismadev Chakrabarti and Simon Baron-Cohen

Autism Research Centre, Cambridge CB2 8AH, United Kingdom.

bhisma@cantab.net <http://people.pwf.cam.ac.uk/bc249>

sb205@cam.ac.uk <http://www.autismresearchcentre.com>

Abstract: The shared circuits model (SCM) is a bold attempt to explain how humans make sense of action, at different levels. In this commentary we single out five concerns: (1) the lack of a developmental account, (2) the absence of double-dissociation evidence, (3) the neglect of joint attention and joint action, (4) the inability to explain discrete emotion perception, and (5) the lack of predictive power or testability of the model. We conclude that Hurley's model requires further work before it could be seen as an improvement over earlier models.

In the shared circuits model (SCM), Susan Hurley provides an impressive, overarching multi-level, multi-layered heuristic model to explain how we perceive intentional action, and to explain imitation and mindreading. Wide-ranging in its coverage of philosophy of mind, comparative and developmental psychology, and neuroscience, the SCM warrants a critical evaluation from within each of these domains.

The SCM proposes five layers of “functional subpersonal” description. The author explicitly resists stating any phylogenetic or ontogenetic relations amongst them. Our first question is to ask what, if any, developmental progression exists between these five layers. Since overt imitation is found in human newborns (Meltzoff & Moore 1983b; 1989), does this mean for example that layer 3 exists at birth? Mindreading abilities are

typically observed by the age of 3, and appear in layer 5. Is Hurley's claim that one cannot pass from layers 3 to 5 without passing via layer 4? The developmental differences from one layer to the next are in need of considerable further specification.

Our second question is to ask whether, in humans, any double dissociations exist among these layers in certain neurodevelopmental conditions. These are not discussed in Hurley's account. For example, Autism Spectrum Conditions entail deficits in both mindreading (Baron-Cohen 1995a; Frith 2001) and affective empathy (Baron-Cohen & Wheelwright 2004). In contrast, psychopathy entails intact mindreading abilities (Richell et al. 2003) but a deficit affective empathy, since they show little psychophysiological response to signals of distress when compared with matched controls (Blair et al. 1997; Prinz 2005). If these two conditions represent a double dissociation between mindreading and affective empathy (Blair & Perschardt 2003; Chakrabarti & Baron-Cohen 2006), how does the SCM explain such fractionation? This dissociation would suggest that intact layers 3 and 4 are not essential for an intact layer 5.

Our third question is regarding where joint attention and joint action are situated in the model. These emerge by the age of 9–14 months in typically developing infants (Scaife & Bruner 1975). Humans may be unique among primates in developing joint attention. They use their index finger to point to share interest, without any special teaching, in every culture where it has been studied (Tomasello et al. 2005). Joint action is also made possible by joint attention and intention detection. Imagine the classic scenario of needing to move a heavy log through a narrow exit, and achieving this with a conspecific but without using language. Person 1 has to attract the attention of person 2, and then succeed in drawing person 2's attention to the far end of the log rather than his own end of it. This would entail glances and gestures that convey (without words) "I'll pick up my end and you pick up your end." Next, person 1 has to make clear to person 2 that he wants him to move his end of the log leftwards, while person 1 moves in the opposite direction, so as to rotate the log (*role reversal*). We guess this thought experiment involving joint action could be achieved effortlessly between two typical humans from age 18 months – and all without language – and some relevant data exist to support this prediction (Carpenter et al. 2005; Phillips et al. 1995; Ross & Lollis 1987). Other models already exist to account for such coordinated action, such as the shared attention mechanism (SAM) within the human mindreading system (Baron-Cohen 1995b). If the SCM cannot account for such a capacity, which arguably lies at the root of later social cognition, then what does the SCM add over earlier models? Notice that a mirror-neuron system might not be up to the task of explaining such joint action, since this might lead person 2 to move left when person 1 does (to mirror his or her movement), whereas person 1 may have been intending person 2 to recognise that the goal was to rotate the two ends of the log in *opposite* directions to get it out of the exit. Reading another's intention during coordinated action may involve goal representation that cannot be explained by mirroring alone.

Our fourth question is concerning how the SCM applies to the processing of facial expressions of *discrete* emotions. In a recent neuroimaging study in our lab, involving passive viewing of dynamic facial expressions of emotion, we found that trait empathy correlated maximally with different brain regions for different emotions (Chakrabarti et al. 2006). A similar result (Lee et al. 2006) showed that different brain regions were maximally active during explicit imitation of different emotion expressions. Interestingly, both of these studies found a common, emotion-independent role for the inferior frontal gyrus. This could possibly be accommodated by the SCM, with the assumption that the perception of emotion-independent facial muscular movement is mapped onto inferior frontal gyrus activity. However, the finding that both passive viewing and explicit imitation of *different* emotion expressions are

associated with activity in *different* brain regions raises a problem for the SCM framework, which makes no provision for how perception of different emotions would recruit layers 4 and 5 to different extents.

Our final concern is the broad question of what testable predictions the SCM makes. In science, models are useful if they make novel, falsifiable predictions. While we reiterate how the SCM provides an impressive review integrating large literatures relevant to studies on action perception, we conclude that it leaves questions of central importance unaddressed and it is not clear how it is an improvement over earlier models.

The neural underpinnings of self and other and layer 2 of the shared circuits model

doi: 10.1017/S0140525X07003160

Linda Furey and Julian Paul Keenan

Cognitive Neuroimaging Laboratory, Montclair State University, Upper Montclair, NJ 07043.

furey1@mail.montclair.edu Keenanj@mail.montclair.edu

http://www.cogneurolab.com

Abstract: Differentiating self from other has been investigated at the neural level, and its incorporation into the model proposed Hurley is necessary for the model to be complete. With an emphasis on the feed-forward model in layer 2, we examine the role that self and other disruptions, including auditory verbal hallucinations (AVHs), may have in expanding the model proposed by Hurley.

In Susan Hurley's shared circuits model (SCM), we see a model that is useful in its intended purposes, though somewhat incomplete both in application and in description, specifically at the neural level. Hurley predicts that disruptions within layer 4 could lead to cognitive deficits, specifically in the capacities for instrumental deliberation and for understanding others' actions. Although her model is successful in predicting some cognitive dysfunction, she does not note relevant problems that can arise when a person fails to adequately differentiate between self and other. Nor does she explore fully the useful role that SCM's layer 2, or, more specifically, the feed-forward model that underlies it, may play both in explicating the self/other delineation and in encouraging the search for the neural correlates involved in the dynamic.

For example, Feinberg and Keenan (2005) found that the parietal lobe, as well as the right frontal regions, are critical in disorders of patients suffering from a loss of self. In delusional misidentification syndrome (DMS), patients consistently and adamantly misidentify persons, places, objects, or events. In delusional reduplication syndrome (DRS), patients reduplicate or double the misidentified entity. Data from such patients revealed that the greatest number of cases is associated with right frontal and right parietal damage. Disturbances of the self/other relationship are not unique to DMS/DRS, however, but occur in a number of neurological disorders. For example, patients with delusions of control confuse self-produced and externally produced actions and sensations; such delusions of alien control are hallmark symptoms of schizophrenia. Hyperactivity of the parietal cortex and cerebellum occurs in such patients, suggesting that over-activation of a cerebellar-parietal network during self-generated actions is associated with the misattribution of those actions to an alien, external source (Blakemore 2003). Another instance of deficits in the ability to differentiate self from other that is unexamined yet germane to the target article concerns auditory verbal hallucinations (AVHs) and inner speech. Individuals who experience AVHs report hearing speech in the absence of any external stimulation; that is, they hear in their head a voice or voices other than their own.

Although AVHs are classified as a first-rank symptom of schizophrenia, they may not necessarily signify pathology and may best be understood within the wider context of the development of inner speech (Jones & Fernyhough 2007a; 2007b). In Vygotsky's theory of the social origins of higher mental processes, inner speech represents the end point of a developmental process in which external conversation gradually becomes internalized to form verbal thought (Vygotsky 1934/1987). Like its semi-covert developmental precursor, private speech, inner speech retains the dialogic nature of the external discourse from which it derives. Fernyhough's four-stage model of the development of inner speech as conceptualized by Vygotsky suggests two distinct forms of dialogic inner speech (Fernyhough 2004): *expanded inner speech*, where the give-and-take quality of external conversation permeates the verbal mentation; and a *condensed* variety of inner speech, where inner speech becomes "thinking in pure meanings" (Vygotsky 1934/1987), having lost most of the acoustics and structure of external dialogue. According to Fernyhough's (2004) theory of AVHs, which draws on Vygotskian ideas about the developmental significance of inner speech, AVHs result from the temporary re-expansion of condensed inner speech, particularly under conditions of stress and cognitive challenge. The acoustic properties of the voices in inner dialogue are thus not attenuated but are experienced fully. The question then is how it is possible that cognition (inner voice) produced by self may be experienced as produced by other.

The cognitive dysfunction that results in the failure to differentiate self from other in inner speech may be explained by a forward model similar to the one underpinning Hurley's layer 2. SCM relies on the forward model of motor control as proposed by Miall (2003) to postulate the subpersonal process that predicts the consequences of motor commands and compares them with the desired state. In her article on delusions of alien control, Blakemore (2003) uses this model to explain how an internal predictor uses information about intentions to enable the distinction between self-generated and externally generated sensory events. The forward model is dysfunctional when it cannot accurately predict the sensory consequences of a movement based on the efference copy of the motor command. This results in sensory discrepancy and a failure to cancel the reafference or actual feedback, so that the self-produced movement feels externally caused (Blakemore 2003; see also Frith et al. 2000b). Although developed to explain abnormalities involving overt actions, this forward model has recently been applied to inner speech (Jones & Fernyhough 2007b). Jones and Fernyhough's application proposes a direct causal mechanism leading from a malfunction of the predicted state to the experience of inner speech as being of alien origin. When the brain either produces a degraded predicted state or fails to produce a predicted state at all from the initial inner speech motor command, the consequence is that an emotion of self-authorship is not felt and instead the inner speech is experienced as authored by an other.

For any model of the mind or cognitive functioning to be complete, it must relate to the brain. Thus, we need to understand the neural underpinnings of the predicted-state mechanism proposed by the forward model. This may require investigating networks, such as the interactions between perceptual and motor areas (Jones & Fernyhough 2007b). For example, Leube et al. (2003) have suggested that neurological activity associated with a deficit in the efference copy mechanism may involve the cortical network that de Vignemont and Fourneret (2004) found implicated in action attribution, including the prefrontal and the parietal cortex, the supplementary motor area, and the cerebellum. In terms of AVHs, Shergill et al. (2000) examined functional magnetic resonance imaging (fMRI) scans of patients with schizophrenia made while the patients were experiencing AVHs. They noted that the pattern of activation observed during AVHs was remarkably similar to that seen when healthy volunteers engaged in auditory verbal imagery (AVI), which is produced when one imagines being spoken to by another person.

Specifically, Shergill et al. (2000) observed common activation of bilateral frontal and temporal gyri, along with right-sided precentral and inferior parietal gyri. Increased supplementary motor area activation was associated with healthy participants generating auditory verbal images; however, the supplementary motor area (SMA) was only weakly activated during AVHs. Other studies have suggested a role for the right anterior cingulate gyrus (see Jones & Fernyhough 2007a and studies cited therein). Given that the parietal and cingulate cortices subservise attention to internal and external bodily space and the attribution of significance to sensory information, they provide a plausible neural substrate for the misattribution of self-generated inner speech to other (see Spence et al. 1997).

Shared circuits in language and communication

doi: 10.1017/S0140525X07003172

Simon Garrod^a and Martin J. Pickering^b

^aDepartment of Psychology, University of Glasgow, Glasgow G12 8QB, United Kingdom; ^bDepartment of Psychology, University of Edinburgh, Edinburgh EH8 9JZ, United Kingdom.

simon@psy.gla.ac.uk <http://www.psy.gla.ac.uk/~simon/>

martin.pickering@ed.ac.uk

http://www.psy.ed.ac.uk/people/martinp/index_html

Abstract: The target article says surprisingly little about the possible role of shared circuits in language and communication. This commentary considers how they might contribute to linguistic communication, particularly during dialogue. We argue that shared circuits are used to promote alignment between linguistic representations at many levels and to support production-based emulation of linguistic input during comprehension.

Hurley's shared circuits model (SCM) provides a framework for investigating the role of emulation and imitation in social cognition. The SCM builds on two recent developments in cognitive neuroscience: Grush's (2004) notion of an emulator (originating from motor control theory) and the discovery of mirror and canonical neurons in monkeys. The target article specifically concentrates on the role of shared circuits in imitation, deliberation, and mindreading. However, it says little about their role in language and communication, which presumably underpin many of the cognitive abilities that Hurley focuses on.

Section 2.3.1 of the target article discusses various hypotheses about how imitation might support language. For example, Hurley argues that the "flexible articulated relations between means and ends in imitative learning could be an evolutionary precursor of arbitrary relations between symbols and referents" (para. 2) and that "mirror systems provide a common code for actions of self and other, and thus for language production and perception" (para. 3). Finally, she suggests that the "flexible recombinant structure of ends and means in imitation may be a precursor of recombinant grammatical structure in language" (para. 4).

However, section 3 contains surprisingly little about the relationship between the SCM (and its various layers) and language processing. In fact, it is only when discussing layer 5 (the full-blown model) that language is considered at all. This is in relation to how imitative learning together with learned manipulation of external symbols could support the rich structure of language. Hurley also speculates that language could assist layer 5 circuits in taking input off-line, thereby allowing for more advanced mindreading (e.g., in multi-person strategic deliberation).

By contrast, we suggest that lower layers of the SCM may play a crucial role in language processing, in particular during

interactive dialogue, which is the most basic setting for linguistic communication. Notice that some of the strongest evidence for perception priming action is from the language domain. For example, there are now a number of demonstrations of the priming of articulators during speech perception using transcranial magnetic stimulation (TMS) and electromyography (EMG) (Fadiga et al. 2002; Watkins et al. 2003). We have argued that, during dialogue, interlocutors tend to align their mental states at many levels, and that such alignment is largely a result of priming (Pickering & Garrod 2004). Indeed, successful communication appears to occur when interlocutors align their models of the situation under discussion. So it would be surprising if the “shared circuits” underlying imitation and mindreading did not also play an important role in this process. In fact, good evidence suggests that alignment of the situation model is supported by rapid and largely automatic alignment at many linguistic levels, such as sound (e.g., Pardo 2006), syntax (Branigan et al. 2000), and meaning of expressions (Garrod & Anderson 1987).

Such linguistic priming would arise at layer 3 of the SCM, just as it does for the chameleon effect (Chartrand & Bargh 1999). Hurley notes “the intimate relationship between the sharing of circuits for self and other and for action and perception: Layer 3’s shared informational dynamics for intersubjectivity presupposes layer 2’s shared informational dynamics for perception and action” (sect. 3.3, para. 3). We argue that just such a relationship holds between shared circuits for linguistic representations in communicators and shared informational dynamics for language production and comprehension (Garrod & Pickering 2004). In other words, covert and overt imitation (i.e., imitative production) at various linguistic levels promotes alignment or intersubjectivity between linguistic representations at those levels.

It is not only in relation to imitation that dialogue processing involves shared circuits. There is increasing evidence that language comprehension like action observation may use production-based (i.e., action based) emulation. In particular, we have argued that comprehension uses predictions based on simultaneous involvement of components of the language production system in the form of a Grush-style emulator (Pickering & Garrod 2007). Such an emulator uses the production system to make predictions (at various linguistic levels) about the input to the comprehension system and runs those predictions in real time. In this way, the system facilitates rapid interpretation and is robust in dealing with ambiguous or noisy language input. At the same time, by priming the production system, the emulator facilitates the rapid switching between comprehension and production during dialogue. Although this production-based emulator is used for comprehending speech, it is built out of exactly the same action-perception components as used in layer 3 of the SCM.

Incorporating control systems into shared circuits for social cognition is a welcome theoretical development. Here we have argued that such shared circuits can also be used to explain how interlocutors align their linguistic representations during dialogue, which ultimately supports successful communication. Indeed, communication is about sharing (with the Latin *communicare* meaning “to share” or “to make common”), so it should come as no surprise that linguistic communication depends upon shared circuits.

Does one size fit all? Hurley on shared circuits

doi: 10.1017/S0140525X07003184

Alvin I. Goldman

Department of Philosophy and Center for Cognitive Science, Rutgers University, New Brunswick, NJ 08901-2992.

goldman@philosophy.rutgers.edu

http://fas-philosophy.rutgers.edu/goldman

Abstract: Hurley’s high level of generality suggests that a control-theoretic framework underpins all of the phenomena in question, but this is problematic. In contrast to the action-perception domain, where the control-theoretic framework certainly applies, there is no evidence that this framework equally applies to feelings and emotions, such as pain, touch, and disgust, where mirroring and simulational mindreading are also found.

Hurley’s target article is pitched at a high level of generality. It speaks broadly of shared circuits, control, mirroring, simulation, mindreading, and so forth, giving the impression that its major theses apply equally across all applicable types of cognition. But there is good reason to doubt that this is accurate, and it is not entirely clear whether Hurley really intends it. Important sub-themes of the target article seem principally aimed at the relation between action and perception – for example, the falsity of the “classical sandwich architecture” (sect. 3, para. 1). Is everything she says about action, perception, and feedback supposed to apply equally to other domains in which shared circuits, mirroring, and mindreading are found? The article’s level of abstraction leaves the distinct impression that the theses advanced at the various layers of analysis cut across all the domains, but that is dubious.

My chief worry centers on the relation between shared circuits (or mirroring) and control theory. Hurley is not alone in emphasizing such a connection (Gallese 2003; Wolpert et al. 2003). However, the case for tying the control-theoretic perspective to shared circuits, mirroring, and simulation is based mainly on the action-perception domain, where there is specific physiological, theoretical, and experimental evidence for efferent copy and reafferent input. Nothing of this sort exists, however, for a number of other domains where shared circuits and simulation are found.

To be specific, mirroring phenomena exist in several areas of cognition in addition to the motoric: in sensation, including pain (Jackson et al. 2004; Singer et al. 2004) and touch (Keyers et al. 2004), and in emotion (most clearly, disgust; see Wicker et al. 2003). But in these domains, there are no established feedback or control-theoretic phenomena of comparable importance – or any sort at all. Here is a brief review of the shared circuits (or mirroring) findings across multiple domains. The shared areas or circuits for action are the premotor cortex and inferior parietal lobule interconnected with the superior temporal sulcus (STS)/middle temporal gyrus (MTG); for disgust, the insula; for fear, (possibly) the amygdala; for pain, the anterior cingulate cortex (ACC) and anterior insula; and for touch, the somatosensory cortices. In all cases, observing what other people do or feel is transformed into an inner representation of what we would do or feel in a similar, endogenously produced, situation. In many of these cases, moreover, evidence drawn from lesion studies and imaging studies indicates that mirroring produces mindreading of others’ mental states (Goldman 2006; in press; Goldman & Sripada 2005). However, only in the case of action is there clear evidence of feedback loops that fit the control-theoretic framework. So the notion that systematic relationships between shared circuits, simulation, and mindreading crucially depend on control-theoretic mechanisms is unsupported. Yet that is what Hurley suggests, since her architecture of social cognition is erected on a control-theoretic foundation.

Hurley writes that “the shared circuits model (SCM) shows how suppersonal resources for control, mirroring, and simulation can enable the distinctively human sociocognitive skills of imitation, deliberation, and mindreading” (sect. 3, para. 1). Her two bottom layers of analysis highlight adaptive feedback control and prediction of effects for improved control, and the three higher layers are explained in terms of these lower-level mechanisms. She makes no attempt, however, to explain how feedback and control account for simulational, empathic, or mindreading properties related to sensation and emotion. Indeed, the latter are barely mentioned. The explananda listed at her top level, the personal-animal level, all involve action and behavior; yet the

social-cognitive phenomena featuring shared circuits include feelings and emotions rather than just action. If a single unifying framework underpinning all shared circuits phenomena is feasible, a different framework seems to be called for.

An alternative approach to many of the same phenomena is the Hebbian learning approach developed by Keyser and collaborators (Keyser & Gazzola 2006; Keyser & Perrett 2004), closely aligned with Heyes's (2005) associative learning approach. When they are young, monkeys and humans spend a lot of time watching themselves. Neurons in the premotor cortex responsible for the execution of a hand-grasping movement will be active at the same time as the visual neurons in the STS respond to the sight of grasping. Given that the STS and area F5 are connected through area PF, ideal Hebbian learning conditions are met: what fires together wires together. Hence, the synapses going from STS grasping neurons to PF and then F5 will be strengthened as the grasping neurons at all three levels are repeatedly coactive. Given that many neurons in the STS show viewpoint-invariant responses, the sight of someone else grasping in similar ways suffices to activate F5 mirror neurons. The same Hebbian argument can be applied to sensations and emotions. While seeing oneself being touched, somatosensory activations overlap in time with visual descriptions of an object moving towards and touching our body. The Hebbian learning approach also has the virtue of not assuming that a particular modality is crucial to shared circuits. Damasio (2003) emphasizes the importance of somatosensory representations, but somatosensory representations do not seem to be important for the representation of action or emotion. Analogously, Hurley does not make a strong case for a single control-theoretic explanation of all types of shared circuitry or the phenomena arising from them (e.g., empathy, mindreading). Indeed, she hardly mentions the existence of emotion and feeling cases. On the surface, these are just as arresting a set of cognitive phenomena as the motor-theoretic ones, and equally in need of explanation.

Can it be argued that the Hebbian associationist perspective is just another version of the control-theoretic one? After all, control theory also postulates association-based learning. True, but Hebbian learning does not posit comparable use of neural comparator systems, forward and inverse models, or other apparatus characteristic of control theory. Also, newly discovered properties of mirroring, such as modulation of mirroring by social relations between individuals (Singer et al. 2006), aren't obviously explainable in control-theoretic terms. But this goes beyond the present purview.

Imitation as a conjunction

doi: 10.1017/S0140525X07003196

Cecilia Heyes

Department of Psychology, University College London, London WC1 6BT, United Kingdom.

c.heyes@ucl.ac.uk

<http://www.psychol.ucl.ac.uk/celia.heyes/netintro.htm>

Abstract: The conjunctive conception takes imitation to be a combination of observational learning and copying. In the target article, and elsewhere, this conception generates problems in (1) explaining the copying of intransitive actions, (2) elucidating the potential functions of imitation, and (3) recognising when the correspondence problem has been avoided rather than solved. Hurley's careful use of subpersonal and personal levels of explanation shows us how to tackle these and other questions about imitation.

Rarely does an empirically minded philosopher – or, for that matter, a scientific specialist – achieve the kind of breadth, balance, and penetration of a scientific literature evident in Hurley's review of research on imitation. In nearly all respects,

the shared circuits model's (SCM's) embeddedness in this literature is a huge asset, to the model and to the field. However, the model has adopted one feature of this literature – a complex, conjunctive conception of imitation – that is problematic.

Hurley takes imitation to be a phenomenon in which action observation causes the observer both (1) to learn an instrumental relationship between a body movement and its effect, and thereby (2) to perform the observed body movement. Here I shall call the first of these conditions *observational learning* and the second *copying*. An alternative conception of imitation, increasingly common in the literature, distinguishes more firmly between observational learning and copying, and calls the latter *imitation*.

Empirical evidence and everyday experience confirm that observational learning and copying can occur independently. As a spectator at Wimbledon, I can learn about the relationship between body movement dynamics and ball dynamics without being able to copy an ace serve. Conversely, when we are talking, I may copy the way in which you tug your earlobe or jiggle your foot, without knowing what, if any, are the effects of these actions (Chartrand & Bargh 1999).

There is no doubt that observational learning and copying can be linked in the manner assumed by the conjunctive conception of imitation; observing your action sometimes both *enables* me to copy, by showing me a new sequence of movements, and *motivates* me to realise this potential, by showing me that the sequence has a particular, desirable outcome. And the functional properties of this conjunction have long been of interest to researchers seeking the origins of distinctively human sociocognitive abilities. Hurley's decision to root the SCM in the conjunctive conception of imitation is, therefore, consistent both with the purposes of the SCM and with the approach of many scientists in the field. But it leads to some problems that might be avoided in future development of the model if observational learning were more firmly distinguished from copying.

The first problem, concerning the explanatory range of the SCM, is openly acknowledged in the target article. Because the SCM assumes that learning about means-ends relationships is an essential component of imitation, and because it treats "ends" as effects of body movement on the world (note the "external feedback loop" in layer 1), it is not clear "whether SCM can extend from instrumental to expressive action" (sect. 4.1.2, para. 1). In other words, the model does not currently account for copying of intransitive actions, such as facial and hand gestures – a type of copying that includes many paradigmatic examples of imitation in the colloquial sense, and appears to be more closely related than copying of instrumental actions to the functioning of the human mirror system (e.g., Fadiga et al. 1995).

The second problem relates to the functions of imitation. Section 2.3 of the target article begins with a finely turned pair of questions: "Does development of either language or mind-reading depend on imitation? If so, at what levels of description and in what senses of 'depend'?" (sect. 2.3, para. 1). (Many of us, tempted to make crude claims about the evolutionary-developmental consequences of imitation, would do well to use this quote as a banner screensaver.) However, having stated the question so clearly, Hurley is compelled by the conjunctive conception of imitation to be, on occasions, less lucid in suggesting and endorsing relationships between imitation and other sociocognitive functions. Take language as an example. She distinguishes four senses in which language may depend on imitation: two, relating to means and ends, that connect language with observational learning (and possibly with all goal-directed action) but not with copying; and two, relating to common coding and parsing of body movement sequences, that connect language with copying, but not with observational learning. When copying and observational learning are firmly distinguished, it is possible to see that one or the other may be involved in the evolution and/or development of language, but the conjunctive conception of imitation gives the impression that, if one is involved, then so is the other.

The third problem is that the conjunctive conception of imitation tends to make researchers lose sight of the correspondence problem – to forget that copying often requires the observer to map visual information from action observation onto matching motor output, under conditions where it is far from obvious how she could have acquired the information necessary to achieve this mapping. The classic examples are facial movements – the sensory input that I receive when I see you, for example, curling your lip, is in a different modality and coordinate frame from the sensory feedback I receive when I produce the same gesture (Brass & Heyes 2005).

The SCM does *not* neglect the correspondence problem. Indeed, it is the principal function of layer 3 to solve this problem, and Hurley and I agree that the solution comes from associative learning (e.g., Catmur et al. 2007; Heyes 2001). However, under the influence of the conjunctive view of imitation, which focuses attention on the impetus or motivation for copying, rather than on its epistemic base, Hurley lets other models off the hook a bit too easily. Explaining imitation with reference to mirror neurons, or to the common codes postulated by ideomotor theory, does indeed “avoid” the correspondence problem, and that is entirely reasonable given that neither mirror neuron research nor ideomotor theory is intended primarily to explain copying/imitation. However, it is important to recognise that avoiding is very different from solving. Mirror neurons and common codes merely move the correspondence problem from the personal to the subpersonal level. Instead of asking how the *observer* knows that lip-curling is the same when observed and executed, we have to ask how mirror neurons or common codes get this information.

One of the greatest strengths of Hurley’s article is the way that it gently, but firmly, encourages those of us involved in research on imitation to clean up our act with respect to levels of explanation. The SCM is a model, not just of imitation-related functions, but of how to respect the three boundaries between subpersonal (neurological and functional) and personal levels of explanation. By teasing them apart, Hurley reveals the complex and sometimes tenuous nature of many pre-existing hypotheses, illuminates questions for further empirical research, and leaves us no excuse for muddled thinking about the mechanisms and functions of imitation.

Shared circuits, shared time, and interpersonal synchrony

doi: 10.1017/S0140525X07003202

Michael J. Hove

Department of Psychology, Cornell University, Ithaca, NY 14853.

mjh88@cornell.edu

Abstract: The shared circuits model (SCM) is a useful explanatory framework that can be applied to interpersonal synchrony by incorporating temporal dynamics. Temporally precise predictive simulations and mirroring enable interpersonal synchrony. When partners’ movements are highly synchronous, the self/other distinction can be blurred.

The shared circuits model (SCM) presented in Susan Hurley’s target article provides a useful framework for understanding imitation, deliberation, and mindreading. The shared informational dynamics of perception/action and self/other not only enable copying and understanding others’ actions, but also enable interpersonal synchrony. The temporal dynamics of perception and action are essential in interpersonal synchrony, and incorporating these aspects could further elucidate two key facets of the SCM: prediction and the self/other distinction. The temporal precision of predictive simulations, integration, and mirroring enables

interpersonal synchrony. Tightly coupled interpersonal synchrony can blur the self/other distinction and potentially increase interpersonal empathy.

Interpersonal synchrony and imitation are examples of social coordination, but differ in temporal aspects. Synchrony, by definition, occurs in shared time, whereas imitation occurs after some delay. The time course in imitation has been investigated. Meltzoff (1988b), for example, employed *deferred imitation* as a measure of memory in infants. In order to investigate perception/action links more directly in imitation, Iacoboni et al. (1999), for example, utilized immediate observation-execution, and thereby mitigated intermediary processes such as memory and interpretation. The delay can and must be excised completely in joint action and synchrony tasks that rely on prediction.

In order to synchronize actions with another person, one cannot simply react to their partner’s actions; instead, one must predict what the other will do and then plan and execute accordingly (Sebanz et al. 2006). In a joint task requiring precise temporal coordination, Knoblich and Jordan (2003) had partners track a circle on the screen with each partner controlling one tracking direction. Successful tracking required anticipatory coordination. The results showed that anticipatory coordination can become as effective with timing feedback from a partner’s action, as when performing the task alone. Action planning was based on the prediction of the joint effects of self and other. This supports the notion of shared representations for self and other and that predictive simulations of others’ actions are integrated with temporal precision.

Similar predictive mechanisms enable interpersonal synchrony in ensemble music performance. Musical synchrony is extremely precise despite the common expressive timing deviations from isochrony (e.g., Rasch 1988). Keller et al. (2007) suggest that such precision is possible by predictively simulating the actions of others. In their study, pianists synchronized more precisely with recordings of themselves after a delay of several months than with recordings of others. Presumably, the predictive simulations were more accurate for self-generated performance because they were carried out on the same perception/action system (with all its idiosyncratic constraints). As Hurley writes, “I perceive your action by means that engage my capacity for similar action” (sect. 4.1.2, para. 4, point 2); and when I perceive my own action, this resonance would be strongest.

The tendency to synchronize with others is well established (e.g., Schmidt et al. 1990). A recent electroencephalographic (EEG) experiment provided evidence for mirror system involvement in interpersonal synchrony (Tognoli et al. 2007). In this interpersonal finger-tapping study, two EEG oscillatory components, whose topographies were consistent with the mirror system, distinguished coordinated from uncoordinated tapping. The authors suggest that the EEG component during phase-locked coordination could be associated with mirror system *enhancement*, whereas the component during uncoordinated tapping could be associated with mirror system *inhibition*. Although the movements and visual input were the same in both cases, the purported mirror system rhythm emerged only when self/other activity was coupled in time.

With shared circuits for perception and action and for self and other, we must somehow distinguish self-produced from other-produced action. However, in interpersonal synchrony, this distinction can become difficult. One way to distinguish self from other is based on the predicted efference copy (SCM’s layer 2): If predicted and actual sensory consequences of an action closely correspond, then the action can be attributed to the self. Temporal correspondence is a key factor in attributing actions to oneself (Sato & Yasuda 2005). But when another’s movement is similar to one’s own in both form and timing, sensory consequences from the other’s movement overlap with one’s own movement prediction and therefore can render self/other attributions ineffective. Another basis for distinguishing

self-produced from other-produced actions is based on monitored output inhibition (SCM's layer 4). Observing another's action maps onto one's own action system and primes similar action, but motor output is inhibited, so the action is not overtly copied. This monitored inhibition during mirroring infers that the observed movement is externally generated, whereas lack of motor output inhibition infers that the movement is self-generated. However, during interpersonal synchrony, mirroring others' actions is not associated with motor output inhibition; hence, one may attribute others' actions to oneself. In interpersonal synchrony, these mechanisms for distinguishing self-generated from other-generated actions are less effective, which in essence blurs the self/other distinction.

Mirror systems and shared intersubjective information prior to distinguishing self and other provide a plausible neural basis for interpersonal empathy (Gallese 2001). Mimicry can lead to affiliation between people because of shared self/other codes as suggested by Chartrand and Bargh (1999). By extension, interpersonal synchrony could produce even stronger affiliation effects because the representational overlap additionally incorporates temporal alignment. Recent data from an interpersonal finger-tapping study support this notion (Hove & Risen, submitted). The degree of temporal synchrony between co-actors predicted subsequent affiliation ratings. Similarly, ensemble musicians and coupled dancers often report affiliation and empathy with partners. Indeed, Walter Freeman (2000) proposed that music and dance evolved as a technology of social bonding. Shared representations, accurate predictions, and temporal alignment can lead to interpersonal empathy and understanding.

In summary, the temporal dynamics during interpersonal synchrony offer an avenue to elucidate the SCM's key aspects of predictive simulation and the self/other distinction. The precise time-course of predicting and integrating other's actions via mirror systems enables interpersonal synchrony. This synchrony can render self/other distinctions ineffective and thereby potentially increase interpersonal empathy. The inclusion of temporal aspects could make the SCM an even more inclusive explanatory framework.

Mesial frontal cortex and super mirror neurons

doi: 10.1017/S0140525X07003214

Marco Iacoboni

Department of Psychiatry and Biobehavioral Science, Semel Institute for Neuroscience and Human Behavior, Brain Research Institute, David Geffen School of Medicine at UCLA; and Ahmanson-Lovelace Brain Mapping Center, Los Angeles, CA 90095.

iacoboni@loni.ucla.edu <http://iacoboni.bmap.ucla.edu>

Abstract: Depth electrode recordings in the human mesial frontal cortex have revealed individual neurons with mirror properties. A third of these cells have excitatory properties during action execution and inhibitory properties during action observation. These cells – which we call super mirror neurons – provide the neural mechanism that implements the functions of layers 3 + 4 of the shared circuits model (SCM).

The process of monitored output inhibition at layer 4 of the shared circuits model (SCM) predicts inhibitory and monitoring mechanisms at the neural level. The question that SCM leaves unresolved is whether, among these inhibitory neural mechanisms, there may be some neurons that are specialized in the inhibition and monitoring of mirroring cells. Clearly, layers 2 + 4 cannot be exclusively implemented by specialized inhibitory mirroring neurons. Indeed, the function of offline predictive simulation distinguishing actual from possible acts can be applied to all sorts of potential actions, including those directed at inanimate objects (say, a mug), which we plan for ourselves when

we are alone (say, in our office). The neural mechanisms implementing layers 3 + 4, however, could be either general-purpose inhibitory mechanisms that may also be applied to the inhibition of overt copying or specialized inhibitory mechanisms for mirroring.

We have performed depth electrode recordings in the mesial wall of the human frontal cortex in patients with epilepsy undergoing pre-surgical evaluation of the foci of epilepsy (Mukamel et al. 2007). From a total of 14 patients, we have recorded the activity of approximately 500 neurons located in three sectors of the mesial frontal cortex: the ventral and dorsal sectors of the anterior cingulate cortex and the pre-supplementary motor cortex (SMA)/SMA proper complex. Activity from individual human neurons was recorded while subjects were performing and observing hand-grasping actions, performing and observing facial emotional expressions, and during control conditions. Mirror neurons were defined as follows: Reliable firing-rate changes were measured during execution and during observation of hand-grasping actions or of facial emotional expressions, but not during the control conditions. We found that approximately 12% of all recorded mesial frontal neurons had mirror properties. Individual neurons with mirror properties were observed in all recording sites in the mesial frontal cortex. This suggests that mirror neurons are widespread in the human frontal lobe. Among these cells, approximately 50% were mirror neurons for hand-grasping actions, whereas the other 50% were mirror neurons for facial emotional expressions. One-third of mirror neurons had excitatory responses during both action execution and action observation. This is the most typical pattern of firing-rate changes observed in monkeys. One-third of mirror neurons, however, had inhibitory responses during both action execution and action observation. This pattern has also been occasionally observed in monkeys, but much less frequently. The remaining third of mirror neurons in the human frontal cortex had a pattern of firing-rate changes that has never been observed in monkeys, at least not so far. The large majority of mirror neurons (more than 80%) have excitatory responses during action execution and inhibitory responses during action observation. Few of these neurons have the opposite pattern, with decreased firing rate during execution and increased firing rate during observation. We call these cells *super mirror neurons* (Iacoboni & Dapretto 2006), not because they have super powers, but because they seem to have a modulatory role over activity in more “classical” frontal mirror neurons, that is, those mirroring cells located in the lateral inferior frontal cortex (Rizzolatti & Craighero 2004).

The mesial frontal areas we recorded from are anatomically connected with the lateral inferior frontal areas containing “classical” mirror neurons (Rizzolatti & Luppino 2001). The physiological properties of the mirror neurons in mesial frontal cortex and the anatomical connectivity between these areas and the lateral inferior frontal cortex containing classical mirror neurons suggest that this mesial frontal mirror neuron system has prevalently inhibitory functions, such that overt copying is inhibited. In line with other models (Iacoboni 2008), SCM suggests that this inhibition of overt copying allows the distinction between the actions of self and other. Several imaging studies that investigated the neural basis of complex self-related concepts have suggested a critical role of mesial frontal areas in implementing such concepts (e.g., Uddin et al. 2007; Vogeley & Fink 2003). Taken together, these theoretical considerations and empirical data support the view that the practical simulative foundations of the SCM's subpersonal functional level may be used to build explicit reasoning and theoretical deliberation.

The physiological properties of the human mesial frontal mirror neurons and their widespread anatomical location support the concept of pervasive mirroring. A fundamental way of connecting with others and even defining the self is by means of mirroring people (Iacoboni 2008).

Flexibility and development of mirroring mechanisms

doi: 10.1017/S0140525X07003226

Matthew R. Longo^a and Bennett I. Bertenthal^b

^a*Institute of Cognitive Neuroscience and Department of Psychology, University College London, London WC1N 3AR, United Kingdom;* ^b*Department of Psychological and Brain Sciences, College of Arts and Sciences, Indiana University, Bloomington, IN 47405.*

m.longo@ucl.ac.uk bbertent@indiana.edu

http://www.homepages.ucl.ac.uk/~ucjtml0/

Abstract: The empirical support for the shared circuits model (SCM) is mixed. We review recent results from our own lab and others supporting a central claim of SCM that mirroring occurs at multiple levels of representation. By contrast, the model is silent as to why human infants are capable of showing imitative behaviours mediated by a mirror system. This limitation is a problem with formal models that address neither the neural correlates nor the behavioural evidence directly.

Hurley's shared circuits model (SCM) is an ambitious attempt to systematize a large body of recent research. One key prediction is that mirroring should occur at multiple grains, or levels of representation in the motor hierarchy. Recent results from our own lab, as well as others, confirm this prediction. Several studies have shown that mirroring is dependent on the presence of the observed action in one's own motor repertoire (e.g., Calvo-Merino et al. 2005). We recently used this finding to examine the level of abstraction at which mirroring occurs, and whether this can be manipulated by instructions (Longo et al., in press). We used a paradigm developed previously (Bertenthal et al. 2006) in which participants observe a video image of a hand at rest with fingers spread apart. The hand is shown from the perspective of someone else facing the participant; the participant responds by pressing a button with the right index finger if the stimulus finger appearing farther to the left moves, and with the right middle finger if the finger farther to the right moves. With a left hand as the image, the stimulus and response fingers match anatomically (e.g., index finger response to an index finger movement); with a right hand, the stimulus and response fingers differ anatomically (e.g., index finger response to a middle finger movement). Responses are faster when there is an anatomical match between the stimulus and response fingers than when there is not, reflecting mirroring, or automatic imitation, of the perceived finger movements.

We used this paradigm to investigate the representational level of abstraction at which mirroring occurs by presenting images of a computer-generated model of a hand, the joints of which could be configured flexibly, enabling us to present finger actions which were either biomechanically possible or impossible. Importantly, the impossible actions were impossible only in terms of the *manner* in which they were performed (i.e., the joints bent in impossible ways), but were perfectly possible in terms of *what* was performed (i.e., tapping a surface). Thus, these actions are impossible at one level of the motor hierarchy (i.e., movements), but possible at a higher level (i.e., goals).

In a first experiment, in which no mention was made of different types of movements, comparable automatic imitation of possible and impossible actions was observed, though participants generally were aware of the difference between the stimuli. This suggests that mirroring involves a common representation at the level of goals. In a second experiment, in contrast, in which attention was explicitly drawn to the manner in which the actions were performed by mentioning the two types of movements during instructions, automatic imitation was completely eliminated for the impossible, but not possible, movements. This latter result suggests that actions were being coded at the level of movements. Together, these results demonstrate that mirroring can occur at more than one level of the motor hierarchy, either in terms of goals or in terms of movements – what

Rizzolatti et al. (2002) referred to as high-level resonance and low-level resonance, respectively.

Similar relations between mirroring and motor ability as those just described have been observed in young infants (e.g., Longo & Bertenthal 2006; Sommerville et al. 2005). These developmental findings are also relevant to our evaluation of the SCM because Hurley acknowledges evidence of mirroring by human infants, but her model remains agnostic as to its origins and prerequisites. By contrast, we contend that the evidence reveals that mirroring or imitation is present from birth, but limited to actions already available to infants.

We (Longo & Bertenthal 2006), for example, used the Piagetian A-not-B error to examine mirroring in 9-month-old infants. This error reflects the tendency of infants at this age to perseverate in searching to a location where they have previously found a hidden object (A), even after having seen it hidden at a new location (B). We found that infants "perseverated" in reaching to the A location, even when they had merely observed an experimenter retrieve the object there, but had not reached themselves. Furthermore, infants were significantly more likely to perseverate when the experimenter had reached ipsilaterally (without crossing the body midline), than when they had reached contralaterally (across the midline). This pattern reflects the difficulty infants of this age show in performing contralateral reaches – what Bruner (1969) referred to as the "mysterious midline barrier" – and demonstrates that mirroring in infants – as in adults – is systematically related to motor skill level. Whereas our results show an effect of action observation on motor performance, the flip side of mirroring is reported by Sommerville et al. (2005), who show that manipulating infants' ability to perform actions alters their perception of those actions when performed by another agent.

Although such results show that mirroring mechanisms are operative quite early in human ontogeny, strong inferences regarding the origins of such abilities must come from studies of younger infants still. In this light, the numerous experiments demonstrating imitation of facial and manual gestures by human neonates are key, suggesting that the neural circuits necessary for mirroring are present at birth. Indeed, given the reported lack of imitation in adult chimpanzees and monkeys, the finding of neonatal imitation in neonates of both species (e.g., Myowa-Yamakoshi et al. [2004] and Ferrari et al. [2006], respectively) is especially striking. Such neonatal imitation disappears over the first few months of life in humans as well as primates, suggesting that rather than reflecting a precocial social-communicative ability, overt mimicry represents an inability to inhibit automatic priming of motor representations. This pattern highlights the fact that at least some forms of imitation are not abilities reflecting long-term learning over time, but are rather automatic tendencies which must be inhibited in order to interact effectively with the environment.

Thus, there is a clear developmental progression of inhibitory control over mirroring responses. Whereas neonates show overt automatic imitation, reflecting very weak inhibitory control, older infants do not compulsively imitate, but are biased in their overt search behaviour by previously observed action. Mirroring in adults is more implicit still, generally manifesting itself in priming of motor responses, rather than their overt imitation (though overt imitation has been reported when attention is diverted [e.g., Chartrand & Bargh 1999; Stengel 1947]). This pattern suggests that much of the development of mirroring *responses* reflects changes in inhibitory control, rather than changes in mirroring *representations*, per se.

In conclusion, the model proposed by Hurley is in the tradition of competence versus performance models. The difficulty with such a model is that it provides a mere skeletal structure that has to be fleshed out in greater detail. Until some critical mass of details has been provided, the validity and usefulness of this model will remain an issue.

Failure, instead of inhibition, should be monitored for the distinction of self/other and actual/possible actions

doi: 10.1017/S0140525X07003238

Takaki Makino

Division of Project Coordination, Tokyo University, Chiba 277-8568, Japan.

mak@scint.dpc.u-tokyo.ac.jp

http://www.scint.jp/~mak/

Abstract: I suggest that layer 4 of the shared circuits model (SCM) should monitor the failure of performing an action, instead of output inhibition, to obtain actual/possible and self/other distinctions. The target article's assumption of selective inhibition leaves some questions unanswered, such as the criteria for the selection. Monitoring failure can answer these questions because failure does not require selection. It also provides a basis for more likely explanation for the phylogenetic and ontogenetic origin of both monitoring and output inhibition.

Susan Hurley has assembled an impressive work describing the model of social cognition. In particular, I was surprised and gratified that the theory we have proposed can be fit into the shared circuits model (SCM) hypothesis without inconsistency. We have been studying mathematical models for mutual and recursive mindreading relations (Makino & Aihara 2003; 2006; Makino et al. 2005), and we have proposed the *self-observation principle* (SOP), stating that, to achieve mindreading, one needs to develop a prediction model of the observation of movement of oneself, which can be driven only from the observation of oneself without an efferent copy of motor commands. The SOP can be regarded as a natural extension of the predictive simulation circuit in the SCM's layer 2 and can provide the basis of the "first-person plural" mirroring system in layer 3. I am happy that our work can be placed within the common framework provided by the SCM.

However, from the aspect of SOP, I found one possible point of improvement for the SCM, namely, monitoring output inhibition. Although I agree that human adults have selective inhibition of imitation, as demonstrated in Lhermitte's imitation syndrome (Lhermitte 1986), the same may not be true for a phylogenetic (and possibly ontogenetic) explanation. In the following, I discuss why it may be better to avoid monitoring selective inhibition, and I propose an alternative: that is, monitoring the failure to perform actions.

It is clear that the output inhibition in the SCM needs to be selective. The target article assumes that the mirror/canonical neurons in layer 3 hold the representation of actions, without information on whether it is the action of oneself or the observed action of another. Hence, the output inhibition, which operates somewhere in the route from the representation of actions to their motor outputs, needs to be selective; otherwise, no action could be performed at all.

However, the target article does not discuss what causes this selectivity in inhibition. Two questions remain unanswered: (1) Upon what criteria is the inhibition selected? (2) Why is the monitoring on the output inhibition, rather than on the selection of inhibition, when the latter would be an easier alternative?

Regarding question (1), the criteria cannot depend on the self/other distinction because it is introduced by monitoring the inhibition. One possible criterion might be the estimation of benefit; that is, an action is inhibited if it is estimated to be non-beneficial or hazardous. However, I am skeptical that a creature that cannot distinguish its own action from that of another would be able to distinguish its own benefit from that of others.

I argue that both the questions are answered if one assumes that *failure* of performing actions, instead of output inhibition, is monitored. It is not difficult to imagine a failure to perform an action. Consider a creature with a primitive, immature mirror system, in a phylogenetically transitional phase from

layer 2 to layer 3. The primitive mirror system would be activated when observing another's action, and produce some weak and partial representation for the action within the shared circuit. In cases where the representation is close enough to the full representation for the observer's equivalent action and the contextual input matches with the action (including posture and other environmental conditions), as well, the shared circuit would cause the represented action to be performed, resulting in priming or imitation. These cases would be rare, however, because the mirror system is still primitive. In most cases, a partially represented action or mismatched contextual input would cause the incomplete performance of action, resulting in failure. If the representation is weaker, or if the context is too distant (that is, if the mismatch is too big), then the shared circuit would totally fail to trigger the action, and, as a result, no imitation would occur at all.

As an answer to question 1, failure gives a good criterion for inhibition. If a creature has only a partial representation or mismatched context for an observed action, it would be more likely to bring undesirable results for itself; so the creature would be better off inhibiting the imitation of the action. Failure in triggering the action implements this in a simple way.

Question 2 is also answered because the failure is not controlled or selected. One can know the failure of action not by monitoring the control signal, but by monitoring the result of the action, including motor output and its reafferent feedback input. Note that this requires a change in the SCM, which originally monitors only the motor output, but I believe that this change is consistent with the design of the SCM.

Moreover, the phylogenetic origin of both the monitoring and the output inhibition can be explained better if one assumes that the action failure is monitored. I suggest that the failure is monitored as a result of exaptation from the detection of the prediction error. This may be more likely because error detection within the learning of simulative prediction in layer 2 is essentially the same information process as failure monitoring. The failure monitoring can also explain output inhibition in humans. Since it is better to inhibit actions that are about to fail, it is reasonable to assume phylogenetic development of the output inhibition function by using monitored inhibition. Such an inhibition would naturally be extended to be more selective, possibly by using self/other distinction.

The discussion so far, about the inhibition of mirrored action in layer 3, can be applied to the inhibition of simulated action in layer 2. The same questions, about the criteria for selecting inhibition and the reason not for monitoring selection, will be answered by assuming failure monitoring. Failure would occur in a primitive version of instrumental deliberation in layers 2 + 4, which might sometimes succeed to take the simulated action in advance, but would fail in most cases because of partial representation of the action or contextual mismatch. In such cases, the action should fail to avoid undesirable results, but the failure would be monitored. Later in phylogenetic time, the monitored failure is used to distinguish actual from possible actions, as well as to selectively inhibit simulated actions to be actually performed.

My point is that layer 4 should not depend on inhibition. Rather, failure monitoring can be used for both the actual/possible and self/other distinction. This provides a more concrete basis for the SCM than does the original formulation, which uses monitored inhibition for these distinctions.

Some predictions, different from those in section 3.4 of the target article, derive from failure monitoring. First, if some species have copying without inhibition, they rarely show such copying, unlike patients with imitation syndrome or echopraxia. Second, there may be creatures with capacities to inhibit copying, but not with self/other distinction. Another hypothesis might include that the imitation probability depends on contextual difference.

ACKNOWLEDGMENTS

I express my deep gratitude to Prof. Kazuyuki Aihara for his valuable advice and for his patience. I am indebted to Prof. Toshihisa Takagi, Prof. Steven Kraines, and Mr. Yohei Akada for their kind support.

The social motivation for social learning

doi: 10.1017/S0140525X0700324X

Mark Nielsen

School of Psychology, University of Queensland, Brisbane QLD 4072, Australia.

nielsen@psy.uq.edu.au

<http://www.psy.uq.edu.au/people/personal.html?id=636>

Abstract: Through the second year, children's copying behaviour shifts from a focus on emulating to a focus on imitating. This shift can be explained by a change in focus from copying others to satisfy cognitive motivations to copying in order to satisfy social motivations. As elegant and detailed as the shared circuits model (SCM) is, it misses this crucial, motivation-based feature of imitation.

There have been considerable advances in the comparative study of primate social learning during the last two decades. As Hurley outlines, one of the key findings to emerge in this time has been the clear and consistent documentation of differences in the ways human children and chimpanzees respond to another's modelled actions. Young children tend to fixate on copying the specific behavioural means used by the demonstrator (i.e., the person modelling the target actions), even if a simpler method is available. By contrast, chimpanzees tend to focus on outcomes, preferring to discover their own means of bringing about the demonstrated end result. This difference is representative of the oft-debated distinction between imitation and emulation. Hurley presents a detailed, thought-provoking model that is aimed, in part, at identifying the possible neural underpinnings of these alternative approaches to social learning. But as elegant as her model is, it misses a crucial component of human imitation: motivation.

A critical means of distinguishing imitation from emulation is to identify whether an observer preferentially aims to reproduce the specific behavioural means a demonstrator used to bring about an outcome or whether she chooses to use her own means. Does the observer focus on copying actions or outcomes? This notion of separating actions from outcomes when evaluating copying behaviour was presaged by Uzgiris (1981), who drew attention to what she saw as two core functions of copying behaviour: a cognitive function that promotes learning about events in the world and an interpersonal function that promotes children's sharing of experience with others. According to Uzgiris, young infants are primarily driven by a need to acquire new skills and behaviours and, as such, when they are shown how to do something, they focus on what was done (i.e., the outcome). However, as they move into their second year, infants become increasingly motivated to engage in social interaction and hence, as a means of realizing the congruence that exists between themselves and others, they begin to focus on the way something was done (i.e., the means used). To put it another way, young infants *emulate* out of a motivation to learn about the world, whereas toddlers show an increasing proclivity for *imitation* based on a desire to interact with, and to be like, others (for a similar view on why adults imitate, see Dijksterhuis 2005).

Recent studies have provided evidence for this proposal of an age-related shift when copying from a focus on outcomes to a focus on actions (Nielsen 2006; Tennie et al. 2006). In a cross-sectional study, Nielsen (2006; Experiment 1) tested 12-, 18-, and 24-month-olds. An adult demonstrated how to open a series of novel boxes (which contained a desirable toy) by using a miscellaneous object to activate a switch located on the front

of each box. The 24-month-olds *imitated* in attempting to open the boxes by using the object, as was shown to them. In contrast, the 12-month-olds *emulated* the demonstrator's actions and only attempted to open the boxes with their hands (18-month-olds showed reactions that were intermediate between the older and younger age groups). In a follow-up experiment (Nielsen 2006; Experiment 2), 12-month-olds did imitate the adult's object use, but only after she had "attempted but failed" to activate the switches by hand. Thus, it appears that 12-month-olds did not fail to imitate because they could not use the object, but rather because they did not interpret this action to be the most efficient alternative available (see also Gergely 2003; Gergely et al. 2002).

Following Uzgiris (1981), I reasoned that the 24-month-olds might persist in imitating a demonstrator's inefficient object use in order to satisfy social motivations. Testing this interpretation, Nielsen et al. (in press) compared the responses of 24-month-olds to live and videotaped demonstrators on the boxes task used in the Nielsen (2006) study. The rationale for using videotaped demonstrators was that they can act in a social and engaging manner but, by virtue of the medium, do not afford opportunity for spontaneous, contingent interaction. If the social motivation hypothesis is valid, children should be less inclined to imitate when the opportunity for social interaction is reduced. They should be less inclined to imitate a videotaped adult than one who is available for interaction. This is exactly what happened. The children imitated the adult's object use significantly less when she appeared on video compared to when she was "live" (Experiment 1). Critically, in a second experiment, when given the opportunity to interact with the adult on a TV monitor via a closed-circuit system (i.e., where socially contingent interaction could take place), the amount of imitation children exhibited returned to "live" levels, indicating that it was the nature of their interaction with the demonstrator that affected the children's copying behaviour, not the medium.

Hurley makes a laudable effort at trying to account for human imitative behaviour by integrating complex motor ability and a capacity for goal reading into the shared circuits model (SCM). Both elements are certainly crucial in determining how we copy others. Nevertheless, as attested to by the previously discussed studies, human social learning can be strongly impacted by interpersonal motivations. These motivations are all too frequently neglected in discussions of, and attempts at explaining, imitation. Unfortunately, the SCM is no exception. There is focus on the way in which a capacity for imitation may get developed and, in layer 5, on how this could then lead to a faculty for understanding other minds. But this is not the same as acknowledging the strong interpersonal motivations that can drive imitative behaviour in the first place.

A growing number of experiments have provided remarkable insights into the neural substrates of imitation. The SCM offers a means of unifying much of this literature and promises to make a major contribution to the field. Nevertheless, one must not lose sight of the fact that human copying behaviour is extremely complex. Its expression is affected by multiple factors and here I have tried to draw attention to interpersonal ones. If we continue to ignore these factors, our understanding of the mechanisms that lie at the heart of human imitation is destined to remain incomplete.

What kind of neural coding and self does Hurley's shared circuit model presuppose?

doi: 10.1017/S0140525X07003251

Georg Northoff

Laboratory of Neuroimaging and Neurophilosophy, Department of Psychiatry, Otto-von-Guericke University of Magdeburg, 39120 Magdeburg, Germany.

georg.northoff@medizin.uni-magdeburg.de

<http://www.med.uni-magdeburg.de/fme/znh/kpsy/northoff/>

Abstract: Susan Hurley's impressive article about the shared circuit model (SCM) raises two important issues. First, I suggest that the SCM presupposes relational coding rather than translational coding as neural code. Second, the SCM being the basis for self implies that the self may be characterized as format, relational, and embodied and embedded, rather than by specific and isolated higher-order cognitive contents.

In her impressive article, Susan Hurley offers the shared circuit model (SCM) as the common structure underlying perception and action, which, as such, can provide the foundation for the overlapping and shared dynamics between self and other. Without going into further details here, I comment on two important conceptual questions raised in Hurley's remarkable account. First, the SCM raises the question of the kind of neural coding that must be presupposed in order to make the SCM and its shared dynamic between perception and action possible. Second, the SCM raises the question of the characterization of the self that is supposed to be based upon the SCM.

The term *code* describes a mean or measurement that captures and reflects teleologically meaningful activity in a system; this mean or measurement is implemented in certain rules and mechanisms that guide and format the system's processing of various contents (see deCharms & Zador 2000; Friston 1997). For instance, these rules and mechanisms may format and guide the neural processing of perceptual contents and action contents. Hurley's SCM, which assumes a shared dynamic and structure between action and perception, implies a common code for perception and action. Referring to the theory of event coding (TEC) by Hommel et al. (2001), Hurley mentions that there might be common coding between action and perception; but she does not elaborate on it in further detail. The TEC (Hommel et al. 2001) claims that perceived events (perception) and to-be-produced events (action) are equally represented by integrated networks and so-called event files (Hommel 2004; see also Noë 2004). What remains unclear, however, is the exact format (e.g., the formal structure) according to which these event files are coded.

Since these "event files" are supposed to be common to both action and perception, there can no longer be translation between the two for a couple of reasons. First, translation presupposes different formats (i.e., formal structures) between action and perception – or else, translation would not be needed. Second, a need for translation would imply that event files are *not* shared between perception and action. Accordingly, there must be a different kind of coding than what I call translational coding, in order to account for Hurley's SCM. How must incoming or outgoing stimuli be coded in order to allow for the SCM and the assumed common structure of perception and action? I suggest that rather than the stimuli themselves being coded, be they either perceptual or action related, it is the relation between different stimuli that is coded. That is, it is not the incoming stimulus of some perceived event that is coded in isolation but rather its relationship to actually generated motor stimuli and vice versa. Such a relationship can be coded only if translational coding is replaced by what I call relational coding (Northoff 2004). Relational coding assumes that the stimuli are formatted according to their relationship to other stimuli as, for instance, incoming sensory stimuli are set and coded in relation to outgoing motor stimuli, and vice versa.

Hurley suggests that the SCM provides the basis for constituting and distinguishing self and other. One would consequently assume that relational coding might also provide the format according to which self and other are coded. This implies not only that self and other are based upon the relation between perception and action but that our self is essentially a rather basic and relational function that is always already set in relation to others and the environment. Rather than attributing some special contents like higher-order cognitive contents to the self, this implies that the self may be considered some kind of specific format that allows for stimuli to be set in relation to each other,

which in turn implicates relation of the stimuli to the respective organism and ultimately to the environment. Instead of considering the self as a special encapsulated entity or function, our self may then be essentially relational so that one may speak of a relational self. This would be well compatible with recent suggestions of self-related processing, which implicates a subcortical-cortical midline network (Northoff et al. 2006; in press). Self-related processing concerns stimuli that are experienced as strongly related to one's own person.

Without going deeply into abstract philosophical considerations, I would like to give a brief theoretical description of what I mean by the terms *experience* and *strongly related*, while to *one's person* is meant very simply as an organism. *Experience* refers to phenomenal experience such as, for example, the feeling of love or the smell of a rose. The term *strongly related* points out the process of associating and linking interoceptive and exteroceptive stimuli with a particular person. The main feature here is not the distinction between diverse sensory modalities, but rather, the linkage of the different stimuli to the individual person, that is, to his or her self. What unifies and categorizes stimuli in this regard is no longer their sensory origin but the strength of their relation to the self. The self-stimulus relation results in what has been called *mineness*; Lambie and Marcel (2002) speak of an "addition of the 'for me'" by means of which that particular stimulus becomes "mine," resulting in "mineness."

In sum, I suggest that Hurley's assumption that the SCM provides the foundation for self and other presupposes (1) a different concept of the self that characterizes the self as format, relational, embodied, and embedded (see also Clark 1997; 1999); and (2) self-related processing rather than by specific contents, a special isolated entity or function, and higher-order cognitive processing. In other terms, Hurley's SCM provides a highly fruitful starting point for reconceptualizing our notion of self and to abandon philosophical and psychological substance-, entity-, or cognitive-based models of self – and, at the same time, for gaining some insight into the hitherto unknown mechanisms of neural coding.

How do shared circuits develop?

doi: 10.1017/S0140525X07003263

Lindsay M. Oberman^a and Vilayanur S. Ramachandran^b

^aDepartment of Neurology, Beth Israel Deaconess Medical Center, Boston, MA 02215; ^bDepartment of Psychology, University of California, San Diego, La Jolla, CA 92093-0109.

loberman@bidmc.harvard.edu vramacha@ucsd.edu

Abstract: The target article discusses a model of how brain circuits mediate social behaviors such as imitation and mindreading. Hurley suggests potential mechanisms for development of shared circuits. We propose that empirical studies can be designed to differentiate the influence of genetic and learning-based factors on the development of shared circuits. We use the mirror neuron system as a model system.

The target article describes several possible scenarios for the development of "shared circuits." For example, this mechanism could be "hardwired" by genes or acquired through learning, or a combination of both. We discuss the evidence for each claim and then suggest experiments that may disentangle the factors contributing to the development of shared circuits by using the mirror neuron system to illustrate our strategy.

As discussed in the target article, studies designed by Meltzoff and Moore (1977) provided evidence for neonatal imitation in infants as young as a few hours of age. Specifically, these infants imitated mouth opening, tongue protrusion, and hand opening. The researchers suggest that the pattern of imitation is not likely the result of conditioning or innate releasing

mechanisms. They suggest that this early imitation implies that human neonates have an innate ability to equate their own unseen behaviors with gestures they see others perform. However, it is possible that the actions investigated by Meltzoff and Moore (1977) were not, as suggested, based on an innate shared circuit, but rather could have been a reflex in response to a smile – like a sneeze in response to pepper. One way to find out would be to test whether infants can mimic an asymmetrical smile or another uncommon action. This would eliminate the “reflex” explanation and implicate a more sophisticated hardwired mechanism based on preexisting rules of translating visual appearance of the body into motor output, leading to accurate imitation.

This type of shared circuit may also be based on a form of associative learning. For example, every time the monkey’s motor command neuron fired to reach for a peanut, the visual appearance of the monkey’s hand reaching activated visual neurons, such that the firing of the two neurons (motor and visual) become linked through Hebbian association. The net result is that the motor neuron itself is activated by the visual image of peanut grabbing, even if the visual image is of another monkey’s hand.

This Hebbian association hypothesis has been suggested by those who argue against the claim that mirror neurons are performing a complex remapping of other’s representations onto one’s own motor system. There are reasons for rejecting this argument (Ramachandran & Oberman 2007). First, given that only a portion of F5 neurons have “mirror” properties, why do these neurons “learn” while others do not? If mirror neurons were set up purely through Hebbian associations, one would predict that all neurons in that region would have mirror properties, but this is not the case. This shows that there are specialized mechanisms and hardwired constraints that characterize the subset of neurons we refer to as mirror neurons. Additionally, the Hebbian hypothesis cannot account for the facial mimicry literature, because, when an infant smiles, the brain receives no visual feedback on which to build an association. It is still possible that this behavior is reflexive and the mirror neuron system may not be mediating it; however, the Hebbian hypothesis is no better at explaining this behavior.

It is also possible that these shared circuits take time to develop. They may require pre-existing hardwired scaffolding that is then “educated” by learning before being fully functional. This does not speak to whether the development results in the motor neuron being converted into a “mirror neuron,” capable of doing a complex self-other algorithm, or whether a motor neuron is simply responding to the visual stimulus as a result of Hebbian associative processes. Thus, the question of innateness/learning and the question of the nature of the computation that is being performed are logically separable. Furthermore, the necessary empirical studies to answer these questions have yet to be conducted.

To answer the question of innateness, one could record from area F5 (a region already known to contain mirror neurons) in a newborn macaque and expose the monkey to several actions, including actions that he will likely be exposed to early in life (e.g., peanut breaking, grasping, etc.), as well as novel actions that are unlikely to be based on pre-existing hardwired mechanisms. If neurons in F5 respond to both the familiar and novel actions the first time they are presented, that would argue for an innate system that does not depend on Hebbian association mechanisms. If F5 neurons respond only to the familiar actions, then the same argument could be made for these findings as was made for the findings by Meltzoff and Moore, that the brain is hardwired to respond to certain evolutionarily relevant actions. Finally, if no F5 neurons respond to the observation of any actions in newborn monkeys, this would argue against mirror neurons being innate.

To test whether mirror neurons are capable of creating a self-other metarepresentation or simply a motor neuron that has

made an associative link to a visual representation, a different type of study would need to be conducted. One possible study would be to record from an F5 mirror neuron in an adult macaque while he watches either another monkey grasp a peanut or himself reaching for a peanut. In the “self” condition, it would be important that the monkey not actually reach for the peanut (enlisting other motor and sensory systems), but instead that the monkey be presented with an optically reversed image so that the inactive monkey has the visual perception of its “own” hand moving. If it is true that these neurons are set up through Hebbian associative processes, the “self” condition should elicit a greater response than the other condition, as this “egocentric view” is what the association was built on. If the metarepresentation hypothesis is correct, however, the “other” condition should elicit a greater response.

There are currently several possible mechanisms for the development of a mirror neuron system. It is our prediction that, like other systems in the brain, these types of “shared circuits” are neither purely learned nor purely innate, but a result of both hardwired and learned processes. Indeed, the circuitry underlying mirror neurons may provide an ideal model system for exploring how nature and nurture interact to create the human body and mind.

More than control freaks: Evaluative and motivational functions of goals

doi: 10.1017/S0140525X07003275

Fabio Paglieri and Cristiano Castelfranchi

Istituto di Scienze e Tecnologie della Cognizione, Consiglio Nazionale delle Ricerche, 00185 Rome, Italy.

fabio.paglieri@istc.cnr.it

<http://www.media.unisi.it/cirg/fp/paglieri.html>

cristiano.castelfranchi@istc.cnr.it

<http://www.istc.cnr.it/createhtml.php?nbr=62>

Abstract: True to its sensorimotor inspiration, Hurley’s shared circuits model (SCM) describes goal-states only within a homeostatic mechanism for action control, neglecting to consider other functions of goals – namely, evaluation and motivation. This restriction thwarts Hurley’s project of identifying the information resources enabling social cognition. In order to master intentional behavior, deliberation, and action understanding, we need to be more than just “control freaks.”

The notion of *goal* conflates three different functional roles: *control*, *evaluation*, and *motivation* (Miller et al. 1960; Rosenblueth et al. 1968). Goals describe target-states that determine appropriate adjustments in the agent’s behavior, for example, via the comparator mechanism described at layer 1 of Hurley’s shared circuits model (SCM) (control function). Goals also indicate valuable states that, even when not pursued, measure how well the current situation agrees with the agent’s interests, that is, how “good” or “bad” the world is (evaluation function). Finally, goals constitute anticipatory drives for the agent’s conduct, providing teleological reasons to perform or refrain from performing certain actions (motivation function).

Possibly because of Hurley’s enactive/embodied inspiration, the focus in her article is solely on the control function of goals. The SCM describes how increasingly sophisticated control mechanisms enable richer forms of social cognition, when paired with mirroring to provide information on the actions of others. More specifically, the sensorimotor control loop is suggested as the basis for understanding the intentional structure of actions. Although this project is valuable in its own right, we doubt it can ever succeed without first incorporating the motivational and evaluative dimension of goal-states (Castelfranchi 1998; Csibra & Gergely 2007).

Indeed, all functions of goal-states are necessary elements of the processes described by the SCM, although they are not yet properly analyzed. When current input is confronted with a target-state, the system is assumed to know what features of the environment are adequate as targets. But how does it know? In mechanical artifacts, for example, the boiler-thermostat system, the target is set by another agent. In adaptive biological systems, some basic target-states are “set” by evolution, but most of the goals relevant for the system are actually self-determined. Hence, some questions arise: How does the agent select a state of affairs as an appropriate target? Why do we use certain results as frame of reference for action control, rather than others?

The answers require invoking the evaluative function of goals: Starting from simple biologically determined goals (e.g., nutrition, proximity to conspecifics), the agent starts appraising positively certain environmental configurations (e.g., food on the table), while perceiving others to be negative (e.g., being isolated). The former are selected as potential target-states, whereas the latter are labeled as things to be avoided, possibly generating other target-states to ensure that they are indeed avoided. These more specific target-states may then acquire evaluative autonomy, if they prove successful over time in satisfying the agent’s basic needs.

Here it is worth emphasizing that evaluation and motivation are *not* exclusive features of personal-level goal-states (desires and intentions). On the contrary, these functions apply also to basic, subpersonal needs, as discussed. Hence, appeals to the functional level of explanation favored by Hurley cannot get the SCM off the hook of this criticism. Evaluation and motivation characterize both personal and subpersonal goal-states, and yet they go unnoticed in the SCM.

Significantly, the target-state is the only component that never changes its dynamics through the five layers of the SCM. This is tantamount to saying that the model does not care for goal generation and goal revision. It being so, can the SCM really claim to enable deliberation and goal-based action understanding? Can we be satisfied with the characterization of deliberation as comparison of alternative predictions – Millikan’s “trials and errors in the head” (Millikan 2006) – while nothing is said about how we compare different goals to choose our future conduct? The SCM reduces deliberation to deciding *how* to achieve a given end, whereas it remains silent on deciding *what* to achieve. This is a badly maimed picture of human deliberation: whenever decisions are made, the *what* is at least as important as the *how*.

As for understanding the intentional structure of action, this implies appreciating that intentions justify teleologically the observed behavior (i.e., motivating and controlling it; Dennett 1987), and that success or failure in achieving goals matters to the agent, that is, the world will be affectively appraised against them (Frijda 1986; Ortony et al. 1988). Moreover, mindreading should provide information on the goal of the observed action, that is, the target-state in the SCM (Gallese & Goldman 1998; Gallese et al. 1996). But there is no evidence that the model is capable of explaining anything of the sort: Mirroring, paired with action inhibition at layer 4, associates input with covert motor activation, which in turn associates with simulative prediction of the next input. Nothing of this, however, associates with the target-state, that is, the proper goal the system should recognize in the action of another agent. So it is doubtful that the SCM, in its present form, describes the informational resources enabling goal-based action understanding.

How could the SCM be amended to rectify these shortcomings? First, it should accommodate the possibility that *multiple targets* are active, so that facing the same input might yield divergent pressures on the agent’s conduct. Even simple organisms confront similar dilemmas, when a certain action (e.g., eating in the open) can have conflicting results (e.g., foraging vs. exposure to predators). Without allowing for multiple targets, the SCM could never claim to enable proper deliberation, that

is, choosing the ends as well as the means. Second, the *target-state must be incorporated in the dynamic loop of action and perception* to provide a grasp on goal dynamics (Castelfranchi & Paglieri 2007). Targets are abandoned, either in favor of better options or because they are satisfied (cyclically or permanently), while other targets emerge, either for instrumental reasons or because something unexpected and rewarding interests the agent. This requires supplementing the comparator mechanism with new functions, including the possibility of registering a mismatch between actual input and target *without* taking action – because evaluation of the world need not always trigger an attempt to change it. Third, the SCM must modify its interpretation of mirroring to account for goal understanding: To this purpose, *target-states must be adequately included in the mirror circuit*.

Ironically, the moral seems to be that the embodied/enactive perspective on social cognition should take the problem of “mind detachment” more seriously. If the aim is to prove that higher cognitive processes are grounded in bodily actions, better arguments are needed to show how mental states, goals among them, become increasingly *detached from actions*, in phylogenesis as much as in ontogenesis (Pezzulo & Castelfranchi 2007; Tomasello 1999). This is indeed a noble quest – one that the SCM has successfully pioneered, but still falls short of having completed.

ACKNOWLEDGMENTS

We are grateful to Maria Miceli, Giovanni Pezzulo, and Luca Tummolini for insightful discussion.

Putting the subjective back into intersubjective: The importance of person-specific, distributed, neural representations in perception-action mechanisms

doi: 10.1017/S0140525X07003287

Stephanie D. Preston

Department of Psychology, University of Michigan, Ann Arbor, MI 48109.
prestos@umich.edu <http://www.umich.edu/~prestos>

Abstract: The shared circuits model (SCM) relies on well-regarded theories of perception-action, mirror neurons, and forward models, but the functional/informational level of the model limits its ability to explain complex behavior such as true imitation. Data from our lab and others confirm the more general details of the model, accepted by most, but specify the neural mechanisms involved in perception-action processes.

The shared circuits model (SCM) has much in common with existing models of complex behavior and relies on some known properties of the nervous system. For example, most researchers no longer hold a pure version of the “sandwich model” and assume that perception and action overlap at the level of representation. Similarly, most agree that imitation exists on a continuum, with complex forms of true imitation relying on and evolving from more simple forms of reflexive imitation. In addition, there is general agreement that mirror neurons and forward models are relevant to questions about how we bridge the intersubjective divide and model others as we do ourselves. Therefore, most of Hurley’s theoretical review is consistent with existing theory and data on the mechanisms of complex interpersonal phenomena, making it unlikely that anyone will take issue with the basic premises of the model; on the flip side, this also means that the model is limited in its ability to stimulate new directions for the field.

The five-layer model is the unique contribution of the model. However, I think that this part of the model suffers from being

pitched at the functional/informational level. This is unfortunate and unnecessary given that much of the theory relies on very specific mechanisms for motor control and perspective taking that are precisely defined and empirically supported. Hence, the model could have been aimed at a functional neuroanatomical level, which would have made it more specific and more accurate.

The model also seems both underspecified and overspecified. In places where the literature is most agnostic on how certain processes work, Hurley is also agnostic. For example, the model intermixes concrete and abstract concepts (such as “targets,” which can be either motor goals or life plans) without specifying whether we use the same neural processes for both, or just use analogous processes when planning to reach for a cup or to overthrow the government. In contrast, the formulation of the model into five discrete layers seems ill-fated, limiting the ability of the model to accord with the structure and functions of the nervous system. For example, layers 1 and 2 likely overlap a great deal in the brain because both require the cerebellum and act in concert to control action (cf. Wolpert et al. 1998). Conversely, there is no reason from phylogeny or ontogeny to assume that these two layers of control are primary to or evolved before the mirroring mechanisms of layer 3. Layer 2 focuses on visual and tactile feedback from the periphery, which are actually slow forms of feedback that forward models were designed to surpass. Layer 4 focuses almost entirely on “monitored inhibition” to segregate activation related to self and other, but it is unclear which type of inhibition is inculcated here (spinal, brain stem, frontal?), and there are many other ways in which self and other activation can be differentiated. Thus, it seems that there are ambiguities and inconsistencies in the model that could have been rectified by making more specific reference to the existing data on how the brain processes information.

Our lab seeks to understand the ways in which people process and understand the emotions of others. Like Hurley, we believe that basic emotion processing and related intersubjective phenomena, such as empathy, rely on an evolutionary conserved and basic perception-action mechanism (PAM) whereby perception of the emotional state of another automatically activates one’s own representations for the state and situation (Preston & de Waal 2002).

Supporting Hurley’s general rejection of the sandwich model, functional imaging work on empathy has found overlap between self and other processes in regions associated with subjective feeling states (Jackson et al. 2006; Lamm et al. 2007; Preston et al. 2007; in preparation b; Singer et al. 2004; 2006). Further supporting Hurley’s application of perception-action processes to simulation, we have also found almost complete overlap in the neural substrates associated with imagining a personal past emotional experience and “trying on” the experience of another subject; however, we also found differences in self and other processes, which would not be predicted by the SCM, but are explicit in the PAM (Preston et al. 2007). In this study, the overall level of brain activation and autonomic arousal were much higher in the self-condition than the other-condition, and subjects recruited additional regions of visual association cortex when imagining another’s scenario. These data suggest that online simulation of actual, personal events can differ in both quality and quantity from that of hypothetical events. Importantly, however, we found these differences between self and other only when subjects *could not* relate well to the situation of the other; there were no differences in neural patterns or autonomic arousal when subjects selected scenarios to which they could relate strongly (Preston et al. 2007).

This latter interaction reflects an important and overlooked point about the processing of other’s actions and states: Perception-action mechanisms *require* that the subject have an existing representation for the action or state of the other. Thus, monkeys do not have mirror neurons for hand manipulations

they do not understand, babies do not imitate gestures that they cannot make, and people cannot resonate with an unfamiliar emotional state and cannot predict your response to a truly novel situation.

We have striking pilot data to support this emphasis on personal representations, as individuals with depression perceive and respond to the distress of others differently than their non-depressed counterparts – they are less personally distressed by the sadness and hopelessness of hospital patients, and they are more likely to feel empathy and offer help to patients with particularly high need (Preston et al., in preparation a).

In another behavioral study, we have found that the mere perception of an emotional facial expression not only activates mirroring in a subject’s facial muscles (cf. Dimberg & Oehman 1996), and primes the same valence in the subject (cf. Murphy & Zajonc 1993), but also rapidly activates the *semantic-level* representation for the *specific* emotion (e.g., “fear”) (Preston & Stansfield, in press) – this finding is not predicted by models that exclusively rely on motor-based, facial feedback, or mirroring of emotion processing, but it is obvious from basic facts about how information is processed from perception to concept retrieval.

It is exciting and promising to have many researchers agreeing on some basic tenants about how behavior is instantiated – however, as with all complex problems, the devil is in the details. In order to make additional headway from here on out, we must look to the data.

ACKNOWLEDGMENTS

I am indebted to Frans de Waal for encouraging this piece and to R. Brent Stansfield for his feedback and support while writing it.

In search of a conceptual location to share cognition

doi: 10.1017/S0140525X07003299

Gün R. Semin^a and John T. Cacioppo^b

^aFaculty of Social Sciences, Utrecht University, 3584 CS Utrecht, The Netherlands; ^bDepartment of Psychology, The University of Chicago, Chicago, IL 60637.

GR.Semin@psy.vu.nl <http://www.cratylus.org>

Cacioppo@uchicago.edu

<http://psychology.uchicago.edu/people/faculty/cacioppo/>

Abstract: It is argued that the multilayered model offered by the shared circuits model (SCM) falls short of capturing an essential aspect of social cognition, namely, its distributed nature. The SCM therefore falls short of modeling emergent social cognition and behavior.

Disciplinary perspectives cut the same realities in different ways, and so it is with philosophy, psychology, neuroscience, inter alia. At times, these disciplined languages cross their linguistic barriers and reach out to systematize knowledge at a level that supersedes the specific limitations imposed by their indigenous language and competence. And, even then, the particular slice of reality that we focus upon is conditioned by presuppositions about the nature of the beast we are examining.

The shared circuits model (SCM) provides a tour de force of portraying a multilayered account of *social cognition*, which is somewhat specifically grounded on imitation, whereby imitative learning is seen as a sophisticated form of social cognition. While *social* cognition appears to be a central construct for the SCM, the entire model is focused on an analysis at the *individual* level. It is this aspect of the SCM that we intend to complement in our commentary by drawing attention to the importance of distributed processes taking place between two or more individuals and the emergent quality of *social* cognition. Indeed, much of

what is presented with the SCM has convergences with a social cognition model we advanced (Semin & Cacioppo, in press a; in press b), although the social cognition model we have advanced is cast in a different mold, in particular with respect to the distributed processes taking place between two or more individuals.

However, the SCM relies on *reception*, namely, the construction of inner neural representations based on observed behavior. Second, social cognition is restricted to a *reproduction metaphor* (e.g., empathy, resonance, imitation, shared representations, or mindreading). Finally, the model attempts to provide an answer as to how intersubjectivity is achieved. However, the model remains at a purely representational level, neglecting the reciprocal nature and co-regulation of social behavior. The three R's (reception, reproduction, and representation) are conceptual consequences of relying on an individual-centered paradigm. There is a problem when social cognition, especially social cognition that emphasizes imitation, is centered on the individual, however. Cognition evolved for the control of adaptive *action*, and social cognition evolved for the control of adaptive interaction in response to evolutionary demands for the organism's survival and reproduction, which for humans always takes place in a social context (Caporael 1997; Fiske 1992), and involves the *co-regulation of action*. Imitation of a parent by an infant is not a solitary event in the service of social cognition. Instead, the infant's imitative behavior elicits an imitative or nurturant response by the parent, which not only reinforces the infant's imitative response but also establishes a connection and constitutes a co-regulation of action by the parent and infant. Any depiction of the social cognition of imitation that ignores the interaction and emergent information between individuals is incomplete.

Thus, social cognition is not driven entirely by inner processes and representations as the SCM suggests, but relies on resources that are distributed across neural, bodily, and environmental features (e.g., Agre 1997; Brooks 1999; Hutchins 1995; Kirsch 1995) with the social and physical environment supporting social action and interaction (Smith & Semin 2004). As this example illustrates, two or more individuals are capable of (a) joint work to perform a feat that supersedes their individual capabilities, and (b) co-cognition and co-regulation to achieve this joint feat. Co-regulation encompasses qualitatively different forms of co-action. The first is *entrainment* and is exemplified by periodic co-action and occurs in cycles. This can be illustrated with the example of rhythmic clapping (e.g., Neda et al. 2000). The second form is non-periodic co-action illustrated by *mimicry* or imitation (e.g., Chartrand & Bargh 1999). The third case is exemplified when people have to perform a complex task requiring interfacing each other's actions (as in open-heart surgery or playing tennis). The third case entails the execution of complementary actions, namely *coordination*, in the pursuit of accomplishing a task (e.g., successful surgery, winning in tennis).

Entrainment, mimicry, and coordination can obviously all occur simultaneously and to different degrees during social interaction. Take, for instance, a dialogue. Any dialogue features a variety of instances of multimodal coordination, entrainment, and mimicry. A dialogue can simultaneously manifest *coordination* as in the case of turn taking in a conversation (e.g., Sachs et al. 1974), or introducing a new topic, at a syntactic level (e.g., syntactic priming; Bock 1986; 1989; Bock & Loebell 1990) or at an affective level (e.g., mood contagion; Neumann & Strack 2000). Simultaneously, it is possible to see cyclically occurring instances of affective facial expressions (e.g., Dimberg et al. 2000) and breathing movements (e.g., Furuyama et al. 2005). Coordination and entrainment can converge when joint behavior is goal driven (e.g., playing tennis versus choral singing): It can be consciously accessible or escape conscious access (two people moving a heavy object versus emotional contagion), or a combination of both.

If the aim of the SCM is to fully understand the bases of emergent social cognition and behavior, then it has to incorporate a level of analysis of interacting dyads and beyond.

Goals are not implied by actions, but inferred from actions and contexts

doi: 10.1017/S0140525X07003305

Iris van Rooij, Willem Haselager, and Harold Bekkering

Nijmegen Institute for Cognition and Information, Radboud University Nijmegen, 6500 HE Nijmegen, The Netherlands.

i.vanrooij@nici.ru.nl w.haselager@nici.ru.nl

h.bekkering@nici.ru.nl

Abstract: People cannot understand intentions behind observed actions by direct simulation, because goal inference is highly context dependent. Context dependency is a major source of computational intractability in traditional information-processing models. An embodied embedded view of cognition may be able to overcome this problem, but then the problem needs recognition and explication within the context of the new, layered cognitive architecture.

Susan Hurley proposes a layered cognitive architecture to model, among other things, the human capacity for understanding people's actions. We applaud the effort because we believe cognitive science can benefit from pursuing alternatives to the traditional cognition-sandwich account, especially when it comes to higher cognition (Haselager et al. 2003; van Rooij et al. 2002). We do see one potential problem with Hurley's conception of how layers 3 and 4 of the shared circuits model (SCM) implement our ability to understand the goals that drive people's actions.

According to the SCM, people understand why people act by "mirroring" the "means/ends structure of observed actions" (sect. 4, para. 5 [layer 3]). From reading the target article, it is less than clear what mechanism underlies the activity of mirroring, but Hurley seems to have in mind a non-inferential mechanism in which goals and actions are directly coupled. According to Hurley, this is made possible by the fact that humans can reverse the direction of the goal – action associations generated by their own goal-directed actions. As a result, Hurley argues, "observing movements generates motor signals in the observer that tend to cause similar movements" (sect. 4, para. 5 [layer 3]). When the motor outputs are inhibited to prevent overt copying, then the system is able to engage in a form of "mirroring [that] simulates in the observer the causes of observed action" (sect. 3.4, para. 5, layer 4 of the SCM).

This conception of inferred goals and their relationship to observed actions is not unproblematic. It seems implausible that a simple one-to-one association between action and goal can account for the intelligent ways in which humans infer goals from observed actions. Research shows that the goals that people infer depend in complex ways on the context in which the actions are observed. For example, the action "pushing a button with one's head" can suggest the goal "that the button be pushed" (e.g., when the person's hands are occupied holding a towel), or the goal "that the button be pushed with the head" (when the hands are free to do the pushing as well). Even infants are sensitive to such contextual factors, leading them to push the button with their hands after seeing an adult push it with her head while holding a towel in her hands, but pushing the button with their heads when the adult's hands were free during the action (Gergely et al. 2002). These observations underscore the problematic nature of Hurley's idea that "observing movements generates motor signals in the observer that tend to cause similar movements" (sect. 4, para. 5 [layer 3]). From the perspective of motor plans, after all, pushing a

button with the hand is very dissimilar from pushing it with the head, yet infants will “copy” observed actions of adults in dissimilar ways if appropriate given the context.

Two defenses of the SCM could be formulated at this point: First, one could propose that the action-goal associations in the SCM are not necessarily one to one. That is, multiple goals could become associated with one and the same action (e.g., picking up a pen could be associated with writing, pointing, giving, etc.), and multiple actions could become associated with one and the same goal (the goal to go to work can be associated with walking, biking, driving, etc.). By “mirroring” one could then retrieve multiple (hypothetical) goals for any given observed action. Although it is conceivable that our brains build complexes of action-goal associations, the question remains how it selects which of the – potentially very many – possible goals is the most plausible or likely goal in the current context. It is known that context sensitivity of such abductive inferences can lead traditional information-processing models into the problem of computational intractability, be they logicist (Bylander et al. 1991), connectionist (Thagard 2000), or Bayesian models (Cooper 1990). It remains a challenge for the SCM, or other layered architectures, to incorporate abductive inference processes that can circumvent this classical intractability problem (e.g., Cuijpers et al. 2006).

Second, one could argue that from the perspective of the observer, two actions do not constitute one and the same observed action if the context of the actions differs. The argument could go as follows: the notion of “observed action” is to be understood to include relevant parts of the context (in our foregoing example: whether the hands are occupied or not); then a unique mapping from action-context pairs to goals can possibly be achieved by a mere “mirroring.” Note, however, that such a proposal serves only to move the problem from understanding the role of context in goal inference to the problem of understanding how people decide which aspects of the context are relevant parts of the current action. This is one of the many disguises in which the infamous frame-problem shows itself (Ford & Pylyshyn 1996; Haselager 1997; Pylyshyn 1987): Figuring out the proper demarcation of what constitutes an “action” is computationally no less challenging than finding the most likely goal in a set of possible goals.

By claiming that goal understanding involves in part an inferential process, we do not mean to suggest that the process is necessarily conscious, controlled, or reasoned in any way. The mechanism can be highly automatic, unconscious, and even build on associative principles. Its implementation may involve the so-called mirror neuron system (Newman-Norlund et al. 2007), but it may also draw upon different neural systems, depending on the nature or complexity of the inferential task (e.g., de Lange et al., submitted). We see it as a challenge for future research to reconcile functional, computational, and neural explanations of goal inference in a way that explains how people can effectively and efficiently make plausible inferences about other people’s goals and intentions in contexts of real-world complexity. So far, traditional information-processing models have failed in this pursuit, due to the apparently insurmountable problem of computational intractability. This is not the place for a full sketch of our views, but we would like to suggest that an embodied embedded view of cognition may prove useful in addressing this problem. First of all, Hurley’s layered (rather than “sandwiched”) view of the cognitive architecture may invite an alternative, nontraditional conception of the inferential task posed to the brain (e.g., van Dijk et al., in press). Secondly, properties of world and body can serve as cognitive resources that may reduce the computational complexity of the inferential task (van Rooij & Wareham, in press).

In sum, Hurley’s model is to be welcomed as a nontraditional model of action understanding, but the mechanisms behind layers 3 and 4 need clarification in view of the computational problems they are supposed to be solving. Embodiment and

embeddedness may help to provide clues for such clarification, although currently this is more a way to formulate the challenge than to answer it.

Imitation, emulation, and the transmission of culture

doi: 10.1017/S0140525X07003317

Andrew Whiten

Centre for Social Learning and Cognitive Evolution, and Scottish Primate Research Group, School of Psychology, University of St Andrews, St Andrews KY16 9JP, United Kingdom.

aw2@st-andrews.ac.uk <http://psy.st-andrews.ac.uk/people/lect/aw2>

Abstract: Three related issues are addressed. First, Hurley treats emulation and imitation as a straightforward dichotomy with emulation emerging first. Recent conceptual analyses and “ghost” chimpanzee experiments challenge this. Second, other recent chimpanzee experiments reveal high-fidelity social transmission, questioning whether copying fidelity is the brake on cumulative culture. Finally, other cognitive processes such as pretence need to be integrated.

Susan Hurley’s target article is an extraordinary achievement. I cannot judge her contribution from the perspective of philosophers, but she has done those of us working on the empirical sciences side of the topics she reviews a great service by pulling together potentially mutually relevant discoveries and developing specific models of how they might form evolutionary or developmental cascades. Having said this, my three sets of comments are rather along the lines of: “here are yet more empirical findings with which any modelling needs to be consistent.”

1. Imitation and emulation. Hurley presents imitation versus emulation as a clear and straightforward dichotomy, with imitation emerging at a higher level in the shared circuits model (SCM). However, recent analyses in animal social learning studies have dissected emulation into several different categories, some of which overlap with imitation under the general heading of “copying” (Whiten 2006). For example, when a chimpanzee matches whichever of two forms of tool use they see (Whiten et al. 2005a), some level of copying is implicated, and even if the matching to the model is in how the tool moves rather than the bodily actions involved, it is not so clear why this should not be regarded as a form of imitation (copying what the model does with their tool) rather than emulation (copying only the model’s effects in the world – how the tool operates). Turning to the empirical evidence, I would argue there is far less direct evidence for emulation in nonhuman animals than Hurley implies. Rather, emulation is often ascribed to an animal as a default explanation when the animal evidences social learning that goes beyond mere stimulus enhancement, but shows little or no imitative fidelity. Positive evidence for emulation is harder to come by. If one sees emulation as learning specifically about the environmental results of actions alone, then one experimental design recently developed becomes particularly significant in this regard. This is the “ghost” experiment in which we discover what observers learn from seeing only an action’s environmental effects, which is what an emulation hypothesis suggests they are learning about. In a recent experiment of this kind we obtained the striking result that chimpanzees failed to learn the complex tool technique involved (Hopper et al. 2007), despite having demonstrated earlier that they would copy the particular one of two such techniques they see performed by a chimpanzee (Whiten et al. 2005a). In short, for learning the most complex technologies they can handle, chimpanzees seem to need to see another chimpanzee perform the act. This finding and the handful of other similar “ghost” studies completed with children and nonhuman species (for a review, see Hopper

et al. 2007) may have profound significance for the kind of model that Hurley has attempted to build, and more generally for theories of embodied cognition and shared circuits.

2. Imitation and cumulative culture. Hurley embraces the hypothesis that the evolution of imitation was the critical factor enabling the emergence of cumulative culture in our lineage. Some of our recent experimental results question this idea. These concern the first “diffusion” experiments among nonhuman primates. These experiments go beyond the usual dyadic social learning paradigm (“What does B learn from A?”) to examine transmission at the group level, which is what culture requires. It is perhaps surprising that this has not been done before, given that the interest of those conducting the numerous dyadic studies in the literature is often in the functional context of cultural transmission. We have now completed seven such diffusion experiments of various designs (including Chinese whisper “chains” A to B to C, etc., and also “open diffusion,” where a model is introduced into a whole group) with chimpanzees and capuchin monkeys, and in six of these there has been a significant differential spread of the two alternative tool use or foraging techniques we seeded in one founder individual of each group (Bonnie et al. 2007; Dindo et al. 2007; Hopper et al. 2007; Horner et al. 2006; Whiten et al. 2005a; 2007). Thus, whether we describe the process as imitation or emulation, there is sufficient copying fidelity to sustain different traditions. It is therefore not so clear that the brake on cumulative culture in nonhuman primates is lack of copying fidelity. From the completely different perspective of archaeology, Mithen (1999) makes a similar point: Acheulian bifaces were complex enough stone artefacts that there can be little doubt that imitation would have been important in learning the knapping techniques required to make them, yet the technology changed little for a million years. Ergo, imitation was not the magic ingredient needed for cumulative culture.

3. A bigger picture than the SCM. Hurley draws several different cognitive systems into the SCM, such as imitation and theory of mind. In the spirit of seeing what may be a bigger cognitive picture in which these are embedded, I suggest that the relevance of the developmental theories of Leslie and Perner also be considered. Leslie (1987) first suggested a linkage between the capacity for pretence (a form of simulation) which typically emerges in 18-month-old children and the origins of theory of mind; Perner (1991) in turn went on to develop a more elaborate theory explaining this linkage and a suite of other cognitive capacities including mirror self-recognition, means-ends reasoning, and understanding invisible displacements. His theory explained these in terms of an emerging capacity for “secondary representation,” in which the child becomes able to simultaneously maintain a grasp on primary representations of reality (e.g., “This is a banana; I can see the marble”) and also entertain secondary representations involving different perspectives on this reality (e.g., “I am acting as if the banana is a telephone; you cannot see the marble that I can see”). This is particularly relevant to the attempt to identify developmental and evolutionary cognitive networks and cascades, for Suddendorf and Whiten (2001) reviewed extensive evidence that the great apes show a cluster of cognitive attributes with similarities to those analysed by Perner, including the recognition that one is being imitated and the early stages of theory of mind.

Imitation and the effort of learning

doi: 10.1017/S0140525X07003329

Justin H. G. Williams

Department of Child Health, University of Aberdeen School of Medicine, Royal Aberdeen Children's Hospital, Aberdeen AB25 2ZD, United Kingdom.

Justin.williams@abdn.ac.uk

http://www.abdn.ac.uk/child_health/williams.shtml

Abstract: Central to Hurley's argument is the position that imitation is “automatic” and requires inhibition. The evidence for this is poor. Imitation is intentional, involves active comparison between self and other, and involves new learning to improve self-other likeness. Abnormal imitation behaviour may result from impaired learning rather than disinhibition. Mentalizing may be similarly effortful and dependent upon learning about others.

At the core of Susan Hurley's thesis lies a claim that imitation is an “automatic” process. This idea was famously espoused by William James, who stated that “every representation of a movement awakens in some degree the actual movement which is its object; and awakens it to a maximum degree whenever it is not kept from doing so by an antagonistic representation present simultaneously in the mind” (James 1890, p. 1134). This is referred to as *ideomotor theory* (Brass & Heyes 2005) and has been a commonly held position in imitation and mirror neuron research in the last 10 years. Hurley takes a further step to suggest that the evolution of inhibitory processes prevents observed movements from being automatically imitated, and this underlies the capacity for imitation to serve a simulation “theory of mind.”

The notion of automatic imitation seems reasonable at a common-sense level. We often copy others without thinking, perhaps by adopting their posture or facial expressions as we chat with them. But, do we really have an urge to imitate all the actions we observe? And what about those actions that are novel? Is learning by imitation simply a matter of observing action and having these observations “awaken” previously latent action plans already encoded in the observer's brain?

An alternative view is that imitation is an active and intentional process. This was the view taken by the Liepmann (Goldenberg 2003), who conducted the first neurological studies of imitation at the beginning of the twentieth century. He saw imitation to be a form of ideomotor praxis, necessitating the implementation of an ideation formed through action observation into an intentional motor act. Studies of action imitation among infants and nonhuman primates would also suggest that imitation is an effortful and selective process (e.g., Gergely et al. 2002; Horner & Whiten 2005), partly motivated by the promise of reward.

Laboratory evidence for automatic imitation comes from several experiments (Brass et al. 2000; Kilner et al. 2003; Press et al. 2005) that show an imitative action response to a cue to be quicker or smoother than when the same response is non-imitative. For example, the cue may consist of an action executed by a mechanical object such as a robot hand. However cue salience may present an intractable problem with these studies (Aicken et al. 2007; Jansson et al. 2007). Attempts may have been made to control for differences in stimulus salience by matching the visual luminance and size of stimuli, but it remains that a human hand is probably more visually salient than a mechanical hand. Moreover, Gazzola et al. (2006) found that viewing a robotic hand activates the mirror neuron system just as well as a mechanical hand does. This suggests that differences in action responses to robots and humans cannot be explained to be the result of differences in ideomotor compatibility mediated by the mirror neuron system. Brass et al. (2005) looked at whether inhibition occurs during action observation using functional magnetic resonance imaging (fMRI). Inhibition would predict activity in the rostral anterior cingulate cortex, which monitors conflict and is active during such tasks as the Stroop test (Amodio & Frith 2006). Brass et al. found activity only in the inferior parietal cortex, reflecting greater differentiation between self and other.

Williams et al. (2007) used fMRI to compare brain activity between imitation and a condition where participants were asked to enact a learnt alternative action to the one observed.

Even at a liberal threshold, there was no greater activity for the alternative action compared to the imitation condition. This demonstrated no evidence of inhibition. In contrast, there was much greater brain activity during imitation. Clusters of activity were identified in rostral anterior cingulate cortex and lateral orbitofrontal cortex. These findings can be understood in terms of current models of motor control (e.g., Wolpert et al. 2001). Imitation is another motor skill which relies on actively developing motor control in an incremental fashion. The effects of any action plan being executed are experienced as consequences for sensory feedback, which serve to modify motor planning functions (cf. the target article). The role of mirror neurons within our model (Williams et al. 2007) is to respond to the fidelity of the *enacted* action compared to the *perceived* action. Error detection results in conflict-related activity in rostral anterior cingulate and a drive to alter behaviour emanating from lateral orbitofrontal cortex. Imitation is therefore an active process of comparison between self and other, requiring continuous modification of action planning with an aim of achieving greater fidelity between self and others. Imitation is intentional, operates within a social context, draws on a capacity for social judgment, and involves new learning.

One form of pathological “automatic” imitation occurs as echopraxia, described in people with autism or frontal lobe lesions (Lhermitte et al. 1986) or those institutionalized with schizophrenia. Such individuals may have increased suggestibility, and impaired capacity for social judgment and flexible rule learning. Williams et al. (2004) suggested that echopraxia in children may reflect *delayed* rather than deviant learning of imitation skills, because echopraxia places lower demands on new learning.

Finally, to return to the central theme of Hurley’s article, what are the implications of this for the relationship between imitation and mentalizing? The demands placed by self-other matching on new learning may also have a bearing on whether mentalizing processes occur automatically or effortfully. Both imitation and simulation “theory of mind” depend on comparing perceptions of another individual’s experience with one’s own. Understanding another individual’s thoughts or actions often requires new learning and perhaps modifying one’s own point of view. Like many other skills that are practiced daily, imitation and mind-reading may appear effortless. The evidence and common experience suggests that it is quite often otherwise.

ACKNOWLEDGMENT

I am grateful to Nina Williams for comments upon an earlier draft.

Response

Bootstrapping the mind

doi: 10.1017/S0140525X07003330

Julian Kiverstein and Andy Clark

School of Philosophy, Psychology and Language Sciences, University of Edinburgh, Edinburgh, EH8 9JX, Scotland, United Kingdom.

j.kiverstein@ed.ac.uk andy.clark@ed.ac.uk

Abstract: After offering a brief account of how we understand the shared circuits model (SCM), we divide our response into four sections. First, in section R1, we assess to what extent SCM is committed to an account of the ontogeny and phylogeny of shared circuits. In section R2, we examine doubts raised by several commentators as to whether SCM might be expanded so as to accommodate the mirroring of emotions,

and intransitive actions more generally. Section R3 responds to various criticisms that relate to the account of social-learning Hurley proposes in the target article. We conclude in section R4 by responding to a number of commentators who argued for the limitation of control theory as a framework for studying social cognition.

Introduction: Stirring the science

“Shared circuits” was Susan Hurley’s last grand project. It set the agenda that might, in another possible world, have allowed Susan and her commentators to begin to converge on a unified and integrated understanding of something quite fundamental to human thought and reason: our ability to know the minds of others, and of ourselves. Susan’s goal was to show one possible way in which a shared basic information space for perception and action might be bootstrapped, function by function, into a grip on self and other, thus providing the essential entrance ticket to the rich realm of social cognition. This project would close the circle, showing that the issues concerning embodiment and dynamics (Hurley 1998) were never that far away from those concerning social cognition, policy, and the possibility of responsibility (Hurley 1989; 2003; 2006a).

Susan died before she could see this project into print. But had she been able to do so, she would have been truly delighted by the wide array of thoughtful, challenging, and constructive commentaries that her shared circuits model (SCM) has elicited. One special source of delight would have been the sheer interdisciplinary diversity of the responses. For Susan believed very strongly that a proper understanding of minds, persons, and reasons would emerge only from tough, cooperative, interdisciplinary work drawing on psychology, philosophy, neuroscience, social science, and cognitive science. One measure of the success of SCM is thus its capacity to stir that larger scientific pot. In that, the treatment (as we see) succeeds wonderfully. Moreover, there seems to be significant agreement concerning many of the finer details of the story. In our response, we try to do three things: First, we briefly clarify the nature of the story on offer; second, we highlight (and where possible respond to) the main critical issues raised by the commentators; and third, we showcase the exciting range of new suggestions (and additional mechanisms) that the commentary phase has uncovered. In suggesting the responses that follow, we are acutely aware of our own shortcomings as surrogate respondents. Some of the issues raised simply exceeded our grasp of the subject area, the target material, or both. Where this has arisen, we have simply remained silent, and beg the readers’ (and the commentators’) forbearance.

R1. The nature of the beast

One of the challenges facing an embodied and situated approach to cognitive science is to map out a path that might have taken humans from the basic kinds of capacities for on-line adaptive response we share with robots and nonhuman animals to distinctively human cognitive abilities, such as the capacity for rational deliberation and the ability to make sense of the purposeful behaviour of other agents. The shared circuits model

(SCM) contributes to meeting this challenge. It describes a set of mechanisms that might have taken humans and their evolutionary ancestors from active perception to imitation, mindreading, and deliberative, strategic thinking.

Each of the model's five layers is given a functional description, which deliberately abstracts away from details of neural implementation. This is not to say that the model is silent about implementation issues – it predicts a common coding for perception and action implemented by the mirror system, for instance (see sect. 2.2 of the target article). However, it leaves the details of how each layer might be neurally implemented as open questions for further investigation. At least one commentator (**Preston**) took the lack of detail at the neuroanatomical/functional level to be a weakness of the model. Notice, however, that any explanation of how the layers of the model are implemented in the brain will itself most likely be a functional explanation. Preston concedes as much in referring to a neuroanatomical/functional level of description. Consider, for instance, a putative explanation of how layer 3 could be implemented in the mirror system. Such an explanation will identify a widely distributed neural system that includes amongst other regions the temporal lobe, the rostral inferior parietal lobule, and the ventral premotor cortex. In virtue of what do these separate neural regions form parts of a single neural system that implements a mirror system? Arguably, disparate neural regions form a part of a single distributed system because of what they do – because the cells at these regions have activation profiles that make a contribution to realising a particular task. It is true that SCM doesn't tell us which parts of the brain coalesce to form the different layers of the model, but it does purport to describe what sub-tasks these parts of the brain must perform if they are to contribute to realising a capacity for action understanding.

In this spirit, we propose interpreting SCM as having two objectives. First, it offers a task-level description of action understanding. Second, it identifies possible mechanisms that could do the work of accomplishing each of these tasks. Action understanding is a multi-faceted capacity, including the abilities to learn novel behaviours by copying, to predict and explain other's behaviour, and to think strategically about one's social interactions. Hurley decomposes each of these complex capacities into more basic sub-tasks. The different layers of SCM describe possible mechanisms that could perform these sub-tasks.

Consider SCM's layer 3 as an example. Layer 3 identifies a mechanism for mirroring – the kind of behavioural priming that occurs in us when we observe an action performed by another person. Our observing the other's action makes us more likely to perform an action of the same type, ourselves. Hurley argues that our capacity for action understanding is facilitated by this tendency to copy behaviours. As the behaviours we can copy become more complex, so also does the repertoire of actions we can potentially understand. Copying behaviour is therefore a sub-task one has to be capable of performing if one is to get into the business of understanding the goal-directed behaviour of others.

Now consider the mechanism SCM introduces to explain an animal's ability to copy behaviour. The mechanism is not new to layer 3 but has already been introduced

at the previous layer to explain amongst other things the motor system's execution of fast, fluent sensorimotor behaviours. Thus, an explanation is being given of one aspect of action understanding in terms of the same mechanisms used to control sensorimotor behaviour. Often, sensorimotor behaviour will require sensory feedback to be made available faster than the sensory systems can supply. A way around this problem would be for the motor system to employ its learned associations between motor outputs and the sensory consequences of those outputs to make predictions about what will happen when a given motor command is executed. In control theory, these predictive models are called *forward models*. At layer 3, forward models are run in reverse, so that a system can, instead of predicting the sensory consequence of an action, work out from an observed action the motor commands that were the cause of this action. Running a forward model in reverse thus produces a motor command. Whether the system then carries out this motor command is a further question. Most of the time, the execution of motor commands arrived at exogenously from observing the actions of others will be inhibited. This makes good evolutionary sense, as **Makino** (and indeed Hurley) notes: A creature that copied the behaviour of an approaching predator would not be long of this world.

How does this mechanism of running forward models in reverse explain an animal's ability to copy behaviour? When a forward model is run in reverse, the outcome of this process will be the production of a motor command. This explains why, when we observe another acting, we are automatically primed to perform the same or similar action ourselves. We have this standing disposition because running a forward model backwards produces in us the same motor plan that initiated the action in the other person.

We can also see now what Hurley means when she claims that a shared information space for perception and action can also function as a shared information space for self and other. Perception and action share a common information space in part because perceiving another agent acting causes our motor system to produce the same or similar motor commands that were the causes of that agent's actions. This information space can also be deployed in action understanding. Say I perceive another person reach for and answer their mobile phone. My motor system will now run a forward model in reverse and produce the same or similar motor plan that led the other person to reach and pick up their ringing handset. This motor plan is now available to be used by other subpersonal systems (layer 4 of SCM) to make sense of the other person's action and its causes. However, the information my motor system makes available doesn't distinguish between motor plans that are my own and have been initiated endogenously, and the motor plans of another person that have been initiated in me exogenously. In this sense, the information space that perception and action share is also an information space that self and other share. This sharing of information makes possible a kind of direct, non-inferential understanding of the action of others. There is no difficulty at this level of processing about how we acquire access to information about the causes of another person's actions. At this level of processing, the

information that is used to make sense of the actions of others is *intersubjective* – it doesn't distinguish between self and other.

So far, we have proposed a way of thinking about each of SCM's layers, but of course a good deal of the model's explanatory work is accomplished by exploring possible interactions between the layers. We have already seen something of how one layer can borrow mechanisms from an earlier layer and exapt this mechanism for a new function. Hurley also describes some ways in which layers might function together to achieve complex tasks that each layer cannot accomplish on its own. Thus, the forward models of layer 2 can be combined with a mechanism for monitored output inhibition introduced at layer 4 with the result that a system can begin to model possible courses of action and assess the consequences of these actions in advance of executing them. A system with layers 2 and 4 can begin to engage in trial-and-error learning "in the head." By combining the mirroring functions made possible by layer 3 with the monitored output inhibition of layer 4, we get a system that can distinguish actions that are its own from actions that belong to another. One difference between motor plans that are endogenously produced and those that are exogenously produced by observing the action of another is that the latter tend to be inhibited. By monitoring inhibited output, a system could thus acquire information that could be used to distinguish its own endogenously produced motor plans from motor plans it finds itself with as a consequence of layer 3. Moreover, such a system could use the information made available by layer 3 to begin to make sense of the other's behaviour. Depending on the system's own repertoire of behaviour and the associations between means and ends that fuel its forward models, a system combining layers 3 and 4 could use its learned associations to understand the means/end structure of the other's behaviour. Layer 5 introduces a capacity for the monitored simulation of inputs. This can be combined with layers 3 and 4 to generate information about the possible actions of others and the causes and effects of such possible actions. A system with this combination of layers can begin to engage in strategic thinking and game-theoretic deliberation about the action of others. (For more details on the role of mindreading in strategic social intelligence, see Hurley 2005a.)

R2. Linking layers

Many of Hurley's commentators raised questions that bear on the interaction among layers we have sketched, and it is to these questions that we now turn. **Chakrabarti & Baron-Cohen's** commentary raises some challenging questions about the developmental progression between layers. **Oberman & Ramachandran** also wonder how shared circuits might develop, but they raise questions about both the ontogeny and the phylogeny of shared circuits. Are shared circuits hard-wired, learned, or a combination of both?

Hurley states that she doesn't take SCM to imply a single account of the development of capacities for imitation, mindreading, and deliberative thinking. The numbering of the layers, she writes, "does not necessarily

represent the order of evolution or development" (sect. 4, para. 9). Rather, the model is intended to provoke hypotheses that map the layers onto "specific phylogenetic or ontogenetic progressions" (sect. 4, para. 9). In other words, Hurley is not committed to a particular answer about how the different layers might feed into a story about the development of mindreading capacities. Nor does she take a firm stand on the question of whether a capacity for imitative learning is culturally acquired or forms a part of our innate inheritance. Hurley believed that questions of this kind would be settled through close collaboration between science and philosophy. Hence, she would have very much welcomed **Oberman & Ramachandran's** constructive suggestions about possible experiments that might answer the nature/nurture question as it arises for mirroring.

This is not to say that Hurley had nothing to say either in her target article or elsewhere about these questions. Indeed, she makes a concrete proposal about how mirroring might have arisen. She begins by telling a Hebb-inspired story about how cells might come to fire both for others' actions that are observed and for actions of one's own that are executed. (The story doesn't originate with Hurley, but can also be found in Goldman [2006, Ch. 6] Heyes [2002; 2005], and Keysers & Perrett [2004]) Suppose the action is one of grasping. Superior temporal sulcus (STS) neurons that respond to observations of grasping behaviour might overlap in time with activity in areas (e.g., PF and F5) that are involved in initiating the grasping behaviour. As a result of Hebbian learning, the connections between STS and the motor areas will be reinforced. The effect of this reinforcement will be that cells in motor areas will fire both when the agent observes his own movements and when he observes the movements of others. Clearly, this sort of account is going to work only for movements that the agent can see himself performing. In order to account for the copying of facial expressions – we can assume that an agent will often be able to copy many facial expressions without being able to see his own face – Hurley introduces a number of different factors. She concedes that there could be a role for innate supra-modal correspondence between observed acts and an observer's similar acts, of the kind suggested by Meltzoff and Moore's (1997) active intermodal mapping (AIM) hypothesis. Hurley also considers a number of other possible explanations for mirroring when one cannot observe one's own behaviour, including one in terms of stimulus enhancement (sect. 3.3, para. 7, 8). We won't repeat the hypothesis. Suffice it to say that Hurley saw a role both for learning and for innate capacities in explaining the emergence of mirroring.

Could a creature understand the action of others but lack a capacity for mirroring? Conversely, could a creature have an unimpaired capacity for mirroring but be incapable of making sense of the behaviour of others? Answering these questions promises to have ramifications for how we think about the relation between mirroring and layers 4 and 5 that do the work of explaining mindreading. **Chakrabarti & Baron-Cohen** suggest that psychopaths may have intact mindreading abilities but deficits in affective empathy. Affective empathy is arguably explained by the mirroring capacities introduced at layer 3, more on which in a moment. Thus, psychopaths may present a case in which we have intact layers 4 and 5 but a

compromised layer 3. Subjects with autism spectrum disorders (ASDs), on the other hand, exhibit impairments in mindreading and affective empathy. Perhaps they provide an example of what can go wrong when layers 3, 4, and 5 are compromised. Chakrabarti & Baron-Cohen present these two cases as an example of double dissociation of mindreading and affective empathy capacities. However, subjects with ASDs do not display the opposite profile to that of psychopaths: they do not have intact capacities for affective empathy but impaired mindreading skills. Baron-Cohen and Wheelwright (2004), for instance, found that subjects with Asperger Syndrome scored significantly lower than normals in a questionnaire testing for empathic skills. Psychopathy certainly establishes the possibility of mindreading without affective empathy, but ASD does not, as far as we can see, establish the possibility of affective empathy without mindreading. This casts doubt on the suggestions that we here confront a clear double dissociation. But leaving this issue to one side, we want to focus on Chakrabarti & Baron-Cohen's interesting claim that this body of evidence challenges the claim that layers 3 and 4 are required for layer 5.

The first response we would make in Hurley's defence is that SCM makes no hypotheses about the development of mindreading abilities, and in particular it does not explicitly claim that layers 3 and 4 are necessary for the emergence of mindreading abilities at layer 5. We have just seen how Hurley left as an open question how the layers of SCM relate to the development and acquisition of mindreading abilities. Nevertheless, it is true that Hurley does offer an explanation as to how a capacity for mindreading might get started, and perhaps it is this story that **Chakrabarti & Baron-Cohen** mean to dispute. We will first consider whether the explanation Hurley offers commits her to the claim that layers 3 and 4 are required for mindreading abilities. Second, we will assess whether the latter hypothesis is really challenged by the disorders discussed by Chakrabarti & Baron-Cohen.

Is Hurley committed to the claim that layers 3 and 4 are necessary if a person is to acquire the sorts of mindreading skills made possible by layer 5? We have already explained how Hurley took mindreading to begin at layer 4. Prior to layer 4, the information a creature has available for making sense of the behaviour of others does not distinguish between self and other. This information can be used to (implicitly) recognise and identify agents that behave in ways similar to me. This recognition of the fundamental similarity between self and other forms the basis for empathy. However, *mindreading* – the interpretation and prediction of other's actions – begins with the acquisition of a grasp of the self/other distinction. Once this distinction is understood, a creature can begin to attribute mental states to the other – to interpret or “read” the other's mind. According to SCM, one acquires an ability to distinguish self and other by acquiring a mechanism for monitoring the inhibition of mirroring. The monitoring of inhibited mirroring allows a creature to identify motor plans that are not its own. With this understanding in place, the creature can begin to populate the world with other perspectives and decentre from its own situation in the here and now to entertain other possible points of view. This capacity becomes more powerful in creatures that possess the representational capacities introduced at layer 5, and can model not just possible courses of action

but can in addition model possible mappings from sensory input to motor output. At first glance, then, it would seem correct to attribute to Hurley the hypothesis that layers 3 and 4 are required for layer 5. Layer 4 looks to be required for layer 5 since the former supplies models of the outputs, which are used at layer 5 in the simulations of complete mappings from inputs to outputs. Layer 3 seems to be required by layer 4 since it makes available the bi-directional simulations that are taken off-line at layer 4. Thus, we might conclude on these grounds that Hurley is indeed committed to the hypothesis attacked by **Chakrabarti & Baron-Cohen**.

Although there are strong grounds for attributing this hypothesis to Hurley, it doesn't seem to us to be strictly entailed by SCM. SCM suggests an explanation of how an animal might come to be able to distinguish itself from others. If we accept the idea that mirroring provides information that does not differentiate self from other, some such explanation is required. However, the cases discussed by **Chakrabarti & Baron-Cohen** involve individuals whose capacity for mirroring is impaired. Such individuals will not need layer 4 to distinguish themselves from others, since they do not have information at their disposal for which the self/other distinction fails. They precisely do not identify with others empathically, nor do they recognise themselves to be similar to others. Although they do not need layer 4 to distinguish themselves from others, layer 4 could nevertheless continue to work in conjunction with layer 2 to provide information about alternative possible courses of action. Layer 4 could continue to supply simulations of possible actions to layer 5. Therefore, it doesn't seem out of the question that a system could exhibit the sort of mindreading abilities made possible by layers 4 and 5 despite having an impaired layer 3.

Suppose we nevertheless concede that a fully intact layer 3 is required for layers 4 and 5. Could an individual with an impaired layer 3 nevertheless exhibit intact mindreading abilities? Possibly. Hurley concedes that even a creature equipped with the sorts of sophisticated representational capacities ushered in by layer 5 might not have what it takes for full-fledged mature mindreading. Mature mindreaders can track many different agents, identifying them in a wide range of different situations. Hurley suggests that language might well be required for an “understanding of multiple others with multiple alternatives and varying beliefs” (sect. 3.5, para. 2). Suppose a person could acquire mastery of a language without the use of layer 3 (a possibility challenged by the claim that imitation is required for the acquisition of language; but for a defence of this hypothesis, see, e.g., Arbib & Rizzolatti 1997; Iaconi 2005). It would then be possible for such a person to exhibit high-level mindreading skills despite lacking a capacity for mirroring. Perhaps such an individual could acquire a theory of mind by learning generalisations relating behaviour, the environment, and mental states in much the same way as scientists generate theories based on observations (for an account of mindreading abilities along these lines, see, e.g., Gopnik & Wellman 1992; 1994). Gallagher (2005) suggests that a high-functioning autistic like Temple-Grandin might deploy exactly this type of theorising strategy to understand the intentions and emotions of others. Gallagher writes of Temple-Grandin that she “reads

about people, and observes them, in an attempt to arrive at the various principles that would explain and predict their actions in what she describes as ‘a strictly logical process’ (Gallagher 2005, p. 236). We can imagine a psychopath acquiring mindreading abilities in much the same way. We conclude, then, that even if we were to suppose that SCM entails the hypothesis that layer 3 is required for layers 4 and 5 (something we are inclined to dispute), SCM can still handle the case of psychopathy.

What about subjects with ASD? These subjects have difficulties taking up perspectives that are not their own, and there is some evidence that this leads to difficulties in imitating (for a balanced assessment of this evidence, see Goldman 2005). Hobson and Lee (1999), for instance, showed that autistic subjects failed to copy behaviours that required perspective switching. In one task, the experimenter took a wooden pipe rack in his left hand and held it against the upper part of his left shoulder. With his right hand, the experimenter took a wooden stick and strummed across the ridges and slots of the pipe rack three times, making a staccato sound. Of the 16 autistic subjects, 15 ran the stick over the pipe rack, but only 2 of the 16 held the pipe rack against their shoulder as the experimenter had demonstrated. In order to copy the action, the autistic subjects had to perform the same action in relation to a different body, their own. First, they had to recognise the relation between the action and the experimenter’s body. This required them to switch from their own perspective to adopt that of the experimenter. Having recognised the relation of the action to the experimenter’s body, they then had to switch back and re-enact this relation from their own perspective. SCM would predict that subjects with impaired mirroring and mindreading abilities would find this sort of perspective switching difficult. Subjects with impaired mirroring abilities will not identify with and recognise others as similar to themselves. Furthermore, when layers 4 and 5 are damaged, subjects will find it difficult to detach from their own perspective. This is exactly what we find in autistic subjects. We conclude that the deficits we find in subjects with ASD may also be consistent with SCM.

Hurley tells us that the non-negotiable parts of the model concern (1) the explanation of mirroring as ‘an exaptive reversal of online prediction’ and (2) ‘the way the actual/possible and self/other distinctions arise as online processes are overlain by monitored inhibition’ (sect. 4, para. 9). We consider next the commentaries that challenge each of these claims beginning with the first of SCM’s non-negotiable claims. Whereas the set of issues we have just considered relate to the interaction between layers, the commentaries to which we now turn question the use that is made of forward models and sensory feedback introduced at layers 1 and 2 to explain the capacities for mindreading that come on the scene with layers 3 to 5.

Goldman worries about the attempt to explain the mirroring of emotion, pain, and other sensations by appeal to lower-level mechanisms of adaptive feedback control and forward models introduced at layers 1 and 2. He challenges a core claim of SCM that there are systematic relationships between the mechanisms used in the control of sensorimotor behaviour and those that underpin our mindreading abilities. Similar concerns are voiced in the commentaries of **Heyes**, **Preston**, and **Chakrabarti & Baron-Cohen**. Heyes argues that the account of

mirroring at layer 3 may not be readily applied to intransitive actions like facial expressions and gesture. Yet she points out that much of the evidence for the mirror system in humans comes from the copying of intransitive actions, rather than the instrumental actions modelled by SCM. Preston claims that there are no good reasons from either phylogeny or ontogeny to claim that control mechanisms like those found at layers 1 and 2 precede the mirroring mechanisms of layer 3. Again the worry seems to be that the appeal to control theory is inadequate when it comes to explaining the sort of mindreading involved in understanding others’ emotional experiences. Chakrabarti & Baron-Cohen wonder how SCM applies to the processing of facial expression. They suggest that the perception of emotions could recruit layers 4 and 5 to different extents, and that SCM makes no provision for such a possibility. Preston can be understood as raising a related concern when she cites evidence in support of the claim that the perception of emotional facial expressions activates semantic-level representations for specific emotions.

We suggest two lines of response. First, it should be recognised that Hurley never claimed to have identified a set of mechanisms that can account for every aspect of social cognition. SCM as it is described in the target article is offered as an account of the understanding of *instrumental actions* – actions that have a means-end structure. Hurley claims that mechanisms from control theory can explain this particular type of social cognition. She claims that there is a systematic relationship between the control and mirroring of instrumental actions. If it should turn out that there is no such systematic relationship between control and the mirroring of intransitive or expressive actions, this would not harm SCM, which claims only that such a relationship holds for the case of instrumental actions. It is certainly an interesting question as to whether SCM might be extended to account not just for our understanding of instrumental actions, but also for what Hurley calls ‘expressive actions.’ Indeed, this is one of the questions Hurley raises in section 4.1.2, and we shall consider this possibility in more detail shortly. It is surely worthwhile to ask how many of our higher cognitive capacities can be explained by appeal to the same basic mechanisms we employ in sensorimotor behaviour. Natural selection often works by taking mechanisms that already exist and tinkering with them. It therefore makes good evolutionary sense to suppose that the very same mechanisms that are used to control sensorimotor behaviour might also serve a very different function in making possible mindreading and social intelligence more generally.

Goldman, **Heyes**, and **Preston** suggest, however, that a different set of mechanisms might be required to explain emotional mirroring from those described at layer 3. Consider the following example of emotional mirroring. I watch a couple arguing and I perceive the woman’s fear at her partner’s anger: I see the fear written on her face. When I see her fear, the same parts of my brain are active as when I myself feel fear. Williams et al. (2001), showed subjects Ekman faces expressing fear, and found that when fearful faces produced increases in skin conductance this response was also accompanied by increased activity in the amygdala. (Ekman faces are photographs of expressive faces used in emotion

recognition experiments.) Adolphs et al. (1994) found that patients with amygdala damage were poor at recognising fear in photographs depicting facial expressions. As in the case of mirroring of instrumental actions, seeing a facial expression of emotion primes us to feel the emotion ourselves. It would seem, however, that SCM's layer 3 cannot explain this type of mirroring. There doesn't seem to be anything analogous to a forward model that could, in the case of emotion mirroring, be run in reverse. When we feel fear, this feeling manifests itself in some change in facial musculature. However, the change in facial musculature we undergo is not a means to achieving any end. We don't intend anything in this case, nor do we act in order to bring about what we intend. Expressive actions like the facial expression of emotions do not have a means-end structure, which is just to say that they are not instrumental actions.

By way of a second response, we want to briefly consider whether an account of emotional mirroring could be given which builds on the sorts of control mechanisms Hurley appeals to explain the mirroring of instrumental actions. In an unpublished review of Goldman's (2006) *Simulating Minds*, Hurley writes that to understand how control and mirroring are related outside the context of instrumental actions and intention reading, an account must be given of "how instrumental and expressive actions are related." She continues, "In my view, mirroring for expressive action builds on the more fundamental, control-related mirroring for instrumental action" (Hurley 2007, p. 11). She doesn't expand on this comment, but we will try to fill out what she might have had in mind, first by contrasting her view of how simulations are involved in mirroring with that of Goldman. On the basis of this contrast one of us (Kiverstein) has developed a somewhat speculative suggestion about how Hurley might have thought about shared circuits as they arise in the context of emotional mirroring. Along the way, we will also have some things to say about how instrumental and expressive actions might be related.

Goldman and Sripada (2005) describe four accounts of emotional mirroring, and plump for what they call *the unmediated resonance model* (also see Goldman (2006, Ch. 6). According to this model, when we see a person's facial expression of emotion, this directly causes in us a similar emotional state: "observation of the target's face 'directly' without any mediation ... triggers (sub-threshold) activation of the same neural substrate associated with the emotion in question" (Goldman & Sripada 2005, p. 207). We come to share the emotion that the other person displays because our observing this display causes in us an emotional experience of the same type. We can recognise the other's emotion on this model because we come to occupy a state that resembles that of the target. Goldman and Sripada are proposing here a simulation-based account of mindreading for the emotions according to which we simulate the other person's emotional state by instantiating a process which, when it functions properly, results in a state that resembles or matches the target's mental state. Simulation is explained here in terms of the products (mental states and their contents) of a mental process of simulation. If this is correct, it is by first producing in ourselves a mental state that matches or resembles the mental state of the target we are seeking to understand that we become able to work out which mental state to attribute to the other.

Hurley suggests in her author meets critics review of Goldman's book (2007), however, that there is a process/product ambiguity in Goldman's account of simulation. One of the lines of evidence Goldman appeals to in support of his account of emotion experience is the finding that the same brain areas are active during the experience of, and the recognition of, emotions. When these areas are damaged, not only do subjects lack a capacity for a certain type of emotion, but they also have difficulties in recognising this emotion on the basis of facial expressions. This evidence suggests a similarity in the neural/functional processes that subserve emotion experience and emotion recognition within an individual. Goldman's account of emotional mirroring takes this similarity in processes to be grounds for inferring a similarity in emotion experience between observer and target. It is this similarity or resemblance that forms the basis for the observer's attribution of an emotion of a particular type to the target. Hurley argues that interpersonal similarity of mental state of the kind Goldman appeals to in his explanation of emotional mirroring is not sufficient for simulation. Interpersonal similarity of states does not count as simulation unless an individual reuses his own mental processes to drive his mindreading. Thus, Hurley proposes what she calls the *re-use conception* of simulation, according to which we come to recognise and understand what the other is feeling by using the processes that take place in us when we undergo an emotion episode of the same type. It is our re-using this same process that explains how we come to understand what the other is experiencing.

Suppose we accept that similarity of emotional state between observer and target is insufficient for mirroring, but that in addition what is required for true simulation is that the observer's emotion *recognition* process must use her own emotion *experiencing* processes for the purpose of simulating the other. Now compare the case of emotion mirroring with mirroring of instrumental actions. The mirroring of instrumental actions is the result of learned associations between movement and the sensory consequences of movement. There will also be learned associations in the case of emotion expression. The changes in facial musculature, for instance, will have sensory consequences. If Adolphs et al. (2000) are right, there will also be somatosensory changes sometimes throughout the body that are associated with a given emotion. So just as in the case of instrumental actions, there will be associations between the execution of an expressive action and reafferent feedback.

There is, however, an important difference in the case of expressive actions that we should mention. In the case of instrumental actions, associations get set up between motor plans and visual experiences of one's own movement, so that when an observer sees similar movements performed by others the same motor plan is evoked. (See the Hebbian account of mirroring sketched earlier and in the target article [sect. 3.3, para. 5].) However, the reafferent feedback that is available in the case of emotion expression won't include visual feedback: We cannot see our own faces when we express an emotion (except when we use a mirror), nor do we see any of the other bodily changes that accompany an emotion experience. Thus, it would seem we run into the correspondence problem in attempting to explain emotion mirroring. The

correspondence problem is of course a quite general problem for explanations of imitation, and one for which various answers have been proposed. One approach invokes general purpose learning mechanisms (see Brass & Heyes 2005). Another introduces innate special purpose mechanisms (see Meltzoff 2002b). We have already seen how SCM suggests that the correspondence problem may be solved by a variety of different mechanisms, but here is not the place to suggest how SCM might tackle this problem as it arises for the case of emotions. We flag this as a problem to be solved in future work that develops an account of shared circuits for emotional expression.

We now have the ingredients in place to sketch a possible way in which an account of emotional mirroring might build on the connection between mirroring and control that SCM describes. We come to recognise the other's emotion experience by re-using our own emotion-experiencing processes. In the instrumental action case, the processes for identifying the intentions that are the causes of an observed behaviour are the same processes we use to act ourselves. The same is true in the emotion case: The processes for recognising emotion are the same as (or better, they overlap considerably with) the processes that cause the expression of an emotion. Are the processes that cause the expression of an emotion the same as the processes that cause intentional action? There may be some crossover, but there will also be important differences. As **Goldman, Heyes, and Preston** note, there is nothing like a forward model and visual feedback in the case of emotional expression. This is why we cannot simply take layer 3 and apply it to the case of emotion mirroring. Let us, however, ignore these differences for the moment and focus on some broad similarities.

Our own emotion-experiencing processes will include motor processes that cause certain facial expressions and the sensory consequences of these motor processes. The actions that constitute the expression of an emotion will be associated with these various kinds of sensory consequence. Hurley argues that perceptual experience (see Hurley 1998) is the result of tracking relationships between sensory flows of information and motor behaviour (O'Regan & Noë [2001a] propose a similar view). A similar model would seem to be applicable to emotional expression. To experience emotion, on this model, is to track certain invariant relationships between an action that expresses the emotion (e.g., a facial expression) and the sensory consequences of this action. This tracking ability forms a part of SCM at layer 2. We can keep track of changes in flows of sensory information because we have learned to associate movements with certain sensory consequences. It is these associations that form the basis for the forward models introduced at layer 2 and that are re-used at layer 3. Our suggestion is that Hurley's account of perceptual experience might be extended to emotion experience so that, when we undergo an emotion episode, what this involves is our tracking the relationships between an action that is the expression of this emotion and the sensory consequences of this action. When we come to recognise the other's emotion, we do so by using the very same tracking abilities. We come to recognise the other's emotion by re-using the same processes that form the basis for our own experiences of emotion. We take this to be one way in which

emotional mirroring might re-use mechanisms introduced at layers 1 and 2 to account for emotion experience. Hence, we tentatively conclude that emotion experience doesn't present an insurmountable problem for SCM. Rather, it presents an opportunity for the future development of the model.

Preston reports a behavioural study (Preston & Stansfield, in press) she interprets as challenging the sort of account of emotion experience we have just sketched. The findings of the study were that perception of an emotional expression not only results in mirroring but also, as Preston writes, "rapidly activates the *semantic-level* representation for the *specific* emotion." Presumably, by "semantic-level" representation, she means that subjects can identify and recognise the emotion. However, on the simulation-based account of mirroring, this is predicted. The idea is that we use the same processes to recognise emotions that we use to experience emotions.

We turn now to Hurley's account of how the self/other and actual/possible distinctions arise out of monitored inhibition, which Hurley describes as the second non-negotiable feature of SCM. **Preston** claims that it might not be monitoring of inhibited output that generates an understanding of the distinction between self and other, and suggests that there are many other mechanisms that might do this work. She does not, however, say exactly what she has in mind. **Furey & Keenan** pursue a similar worry and ask whether an account might be given of an understanding of the self/other distinction in terms of forward models and the work they do in distinguishing self-caused actions from externally generated actions. Furey & Keenan discuss the case of auditory verbal hallucination when subjects claim to hear voices in their heads. They suggest that the misattribution in this case might be explained by a malfunctioning forward model which doesn't perform its normal function of enabling the subject to distinguish self-generated inner speech from externally generated speech. We find this suggestion very plausible. Furthermore, Furey & Keenan's idea of applying forward models and efference copy to the case of inner speech resonates well with **Garrod & Pickering's** compelling account of the role of forward models in interactive dialogue. However, we take this suggestion to show that forward models can help explain how a subject might acquire an understanding of the difference between self and world. Part of this understanding will include an ability to distinguish his own actions from externally caused events, and efference copy will no doubt have an important role to play in such an explanation (see, e.g., Blakemore et al. 1999). Notice, however, that this is not the same problem that the monitoring of inhibited output was introduced to solve. At layer 3, we have a single process that is involved both in the execution of an action by the self and in the observation of actions performed by others. The information space for perception and action is therefore a shared information space for self and other. Given this, the problem is then to explain how a creature using information that doesn't distinguish between self and other could acquire a grasp of such a distinction. The resources Furey & Keenan describe might yield an understanding of the difference between self and world and enable normal subjects to solve the sorts of attribution problems these commentators describe. However, it is not clear that

these resources could help a mirroring system to differentiate information that relates to the self from information relating to others.

Northoff, meanwhile, asks some hard questions regarding the exact form of the neural coding implicated by Hurley, and suggests that the coding needs to capture the relations between different stimuli and between stimuli and motor actions. It seems to us that Hurley would agree, and that relational coding (insofar as we understand this notion) is indeed apt for many of the purposes of the SCM. Whether Hurley's story further demands, as Northoff suggests, a radically new conception of the self so as to reflect these relational elements, is a matter we cannot resolve definitively. But elsewhere Hurley speaks intriguingly of the self as a "dynamic singularity" itself created out of a system of relations (see Hurley 1998, pp. 206–207).

Makino argues that the self/other distinction arises not when the motor system monitors its inhibited outputs but instead from a monitoring of failed actions. He raises some interesting and important questions about how the motor system works out which outputs to inhibit, and he suggests as an answer that the motor system will tend to inhibit output in cases when it is operating only with partial information. This may be one way of making this sort of decision, but it doesn't seem obvious to us that all inhibited actions are ones that would fail were they to be performed. Furthermore, it wasn't entirely clear to us how to understand the suggestion of monitoring failed actions. If this consists in monitoring actions that the system fails to perform, then it strikes us that this is just another way of talking about monitoring of inhibited input. We understand inhibition as the default principle that the motor system operates on, because inhibiting actions that the system has a tendency to copy is adaptive. This default is overridden in some cases, and the decision as to when this happens will be in the hands of executive systems in the brain. We will return to a related issue shortly in our discussion of the commentaries by **Behrendt** and **Williams**.

Hove provides some nice examples of what he calls "interpersonal synchrony" in which monitoring of inhibited output might fail to generate an understanding of the difference between self and other. In the sorts of cases he has in mind, our actions are synchronised with those of another person. Hove asks how we distinguish our own actions from those of the other in these sorts of cases. There is no motor output inhibition in the cases Hove describes, so it doesn't look like we can appeal to this mechanism to solve the problem. Hove certainly raises an interesting problem here, but it seems to us that the sorts of cases his puzzle arises for are not the ones layer 4 is introduced to explain. His examples of interpersonal synchrony do not involve the copying of behaviour, so they do not meet Hurley's definition of mirroring, where *mirroring* is the process that occurs when observing a behaviour primes the observer to perform the same behaviour himself. We turn next to questions relating to imitation.

R3. Imitation and mirroring

Williams and **Behrendt** both discuss the question of whether mirroring is an automatic process, but arrive at

opposite answers. Behrendt asks how predictive simulation, which he takes to require consciousness, can be related to mirroring that happens automatically. It is not clear why Behrendt thinks predictive simulation must be conscious. Hurley takes layer 2 of her model, which is the layer at which predictive simulation occurs, to describe a subpersonal mechanism. Forward models do produce simulations that may often involve motor imagery; however, this imagery is not always conscious. Behrendt notes how movement plans can be formed automatically upon perception of a salient event. When this event is an observed action, we have just the sort of mirroring that layer 3 is introduced to explain. Behrendt points out that traits and stereotypes can automatically elicit patterns of behaviour, but again it is just this sort of case that SCM's layer 3 was introduced to explain. Behrendt also points out that our motivation for copying in these cases may often be social approval. **Nielsen** provides some experimental results that support this idea. He suggests that 2-year-old infants will commonly be motivated to copy behaviour because they want to share an experience with the other, and he describes experiments in which 2-year-olds were less inclined to imitate when not engaged in social interaction. We think Hurley would have found these results extremely interesting, but that claims about what motivates us to imitate lie somewhat out of the purview of SCM. It is an important and interesting question as to just why we have a tendency to imitate, and it is a striking finding that this tendency changes as we develop. SCM, however, seeks to show that there is a systematic relationship between the control mechanisms we use in sensorimotor behaviour and the understanding of instrumental actions. It is a further question as to what reasons we have for imitating when we do so.

Williams attributes to Hurley the claim that imitation is automatic. He goes on to make a compelling case for the view that imitation is an "intentional," "effortful," and "selective" process. Williams describes how imitation requires continuous modifying of action plans with the goal of getting the agent's actions to match the actions of the model. Hurley is certainly committed to the claim that the tendency to imitate is automatic, and that the performance of imitative behaviour must be inhibited. However, this seems to be distinct from the claim that the performance of imitative behaviour is automatic. This is precisely not the case. Normally, we copy behaviour covertly, not overtly. It is this covert copying which Hurley claims forms the basis for the simulation routines we use to understand the instrumental actions of others. Mature adult humans can sometimes fail to keep imitation covert when they have suffered damage to their prefrontal cortex or when their caudate nucleus is overactive (Kinsbourne 2005, p. 165). Kinsbourne notes:

Echopraxics do not walk through the world twitching in response to every movement around them. They do not imitate the rustling of the leaves and they do not imitate cars screeching to a halt. One elicits echopraxia by being a doctor, facing a patient, looking somber and purposeful, and giving the patient tasks. (Kinsbourne 2005, p. 166)

So, even in this case, imitation is not wholly uninhibited. Williams makes an interesting suggestion about echopraxia. As already explained, he rightly stresses the role of social interaction in imitative learning. He suggests that echopraxia may be understood as the result of an

“impaired capacity for social judgement and flexible rule learning.” Both capacities are required for imitative learning according to Williams. Thus, when either of these capacities is damaged, the result is that subjects can no longer imitate. We think that the executive areas of the brain responsible for inhibiting imitative behaviour are also very likely involved in regulating behaviour in accordance with social norms and in flexible rule learning. Thus, we wonder whether the difference between Hurley and Williams on this point might not be so great.

Whiten and **Heyes** both put pressure on the conception of imitation Hurley assumes in her target article. Whiten objects to Hurley’s account of the phylogeny of imitation where emulation comes first and imitation only rarely follows. He rejects what he describes as the “dichotomy of imitation or emulation,” arguing that emulation comes in a variety of different forms, some of which overlap with imitation. In his ingenious and important “ghost experiments,” chimpanzees witness only the environmental effects of the complex use of a tool. If the chimpanzees were able to learn through emulation, it ought to be sufficient for them to just observe the goal-directed action. However, Whiten found that chimpanzees could learn the complex use of a tool only by perceiving another chimpanzee (not a “ghost”) use the tool. The suggestion seems to be that, at least for complex techniques, chimpanzees cannot figure out their own means to achieving an end. To learn a complex technique, they must copy both ends and means.

Before we explore some ways in which Hurley might have thought about the relation between emulation and imitation, we should briefly note that she was certainly no sceptic about imitative behaviour in animals. She was keen to stress just how complex imitative behaviour is, requiring as it does that an animal execute movements from its behavioural repertoire in a new way to achieve some desired result. This requires what she describes as the “flexible interplay of copying ends and copying means; a given movement can be used for different ends and a given end pursued by various means” (sect. 2.1, para. 5). Hurley doesn’t deny that some animals are capable of this kind of complex behaviour. She notes, for instance, that in the artificial fruits experiments chimpanzees will imitate selectively only when the method for opening the fruit is the most efficient. Elsewhere (Hurley & Chater 2005b), she discusses and seems to endorse Byrne and Russon’s (1998) finding of program-level imitation in gorillas and orangutans. Program-level imitation occurs when animals learn to copy a specific sequence of behaviours for the performance of a task, such as the preparation of a particular type of plant for eating. Thus, although Hurley certainly insists on a distinction between imitation and emulation, and insists that imitation is phylogenetically rare, she was no sceptic about nonhuman imitation.

Hurley did, however, insist that the capacity for social learning varied across species. She identifies two factors that contribute to this variation (sect. 3.3, para. 9):

- (1) The grain and complexity of instrumental control capacities
- (2) Considerations concerning which of the many control capacities have associated mirroring functions and how richly and flexibly these mirroring circuits can be linked

Some animals will be capable of performing instrumental actions that are more complex than others, where complexity of behaviour is a function of “means/ends chains of differing grains and lengths” (sect. 3.3, para. 11). An animal that combines multiple behaviours in ways that are appropriate to achieving a given end will be capable of forming predictive models that are much richer in structure than will an animal that can perform only simple behaviours to achieve its ends. Mirroring takes the instrumental associations between means (a motor program) and ends (the consequences of performing an action) that provide the information for predictive simulations and uses these associations to mirror the cause of another’s movement. Animals that employ predictive models that are rich in structure will be capable of mirroring instrumental actions that are equally rich in structure. The potential for social learning in such an animal will be much greater than that in animals that are only capable of simple behaviours and predicting the consequences of those simple behaviours. Hence, Hurley concludes that, “Mirroring and simulation might provide information about the goals of certain observed movements, given fine-grained, complex means/end associations but not given coarser control capacities” (sect. 3.3, para. 13). Animals whose behaviour has a rich means/end structure will be capable of complex forms of mirroring, and it is this mirroring that forms the basis for social learning. Animals whose behaviour lacks this structure will be capable only of movement priming or perhaps of goal emulation.

Heyes questions what she describes as Hurley’s *conjunctive conception* of imitation. According to the conjunctive conception, imitation requires both (1) observational learning, as when an agent learns an instrumental relationship between a bodily movement and its effect, and (2) a capacity for copying, where this involves the ability to perform the observed body movement. Heyes suggests that observational learning should be distinguished from copying. It is not clear to us whether she thinks copying is sufficient for imitation. We would question such a claim. Copying seems to be a type of behaviour that could be manifested by creatures that are only capable of what Hurley calls stimulus enhancement: The action of another animal draws the observing animal’s attention to a stimulus, and the stimulus then triggers an innate or previously learned response. Copying also seems to occur in cases of movement priming, when observing a bodily movement primes an animal to perform a similar movement, but not as a means to an end. It seems to us that copying has to go together with observational learning if the animal is to be correctly described as having learned through *imitation* – that is to say, by performing some sequence of movements from its behavioural repertoire in a new way so as to achieve a desired result. We would not describe an animal as having learned through imitation if the animal doesn’t understand the instrumental relationship between performing some bodily movements and achieving its ends or goals.

Longo & Bertenthal describe experiments that seem to challenge the connection SCM describes between mirroring and imitation. They argue that in some contexts the motor system will just copy movements and in other contexts the motor system will copy goals. They describe a paradigm in which subjects are shown a computer-generated

hand performing movements, some of which are possible and others of which are impossible. They found that subjects attempt to copy both the possible and impossible movements unless their attention is explicitly drawn to the manner in which the movements are being performed. It is not clear to us what the instrumental action is in this experiment. What are the means and ends? What is the computer-generated hand moving its hand to do? Given that it is unclear what the goal of the movement is in this case, it is hard to assess whether subjects were just copying the movements in the first set-up and the goals in the second (once their attention was drawn to the manner in which the movement was performed). We think it would be interesting to run the same experiment again, but this time to have the computer-generated hand explicitly perform movements with the end of achieving some goal. In one case, the end could be performed by means of biomechanical movements that are impossible for the human hand, and, in the second case, by movements the human hand could perform.

R4. Beyond shared circuits

Apart from the large raft of questions concerning the details of the inter-layer transitions, relations, and interactions, a number of commentators raised questions of scope. How much can the shared circuits model, with its strong commitment to a single kind of model and mechanism (a control-theoretic account of simulation and mirroring, pursued through a cascade of stepwise refinements) really explain? As the advertising would have it, the model aims to reach and illuminate “imitation, deliberation, and mindreading.” But does it really have the resources to do so? More accurately, just *how much* of our understanding of the minds, goals, and intentions of other agents can be explained using the *kinds* of resources Hurley so ably displays?

In much this vein, **Carpendale & Lewis** worry that the model “fails to reach action understanding because it relies on mirroring as a driving force.” Their key charge is that mechanisms of mirroring are simply too unintelligent to yield much in the way of action understanding. I may see you point at something, and that may automatically activate a pointing tendency in me (even if it is inhibited), but what does that tell me about why you are pointing? What is missing, they suggest, is “experience in shared routines.” Carpendale & Lewis are right, we think, to identify these kinds of limits in the direct application of the model. For mirroring circuits do, indeed, only deliver information about others’ goals and intentions for classes of actions whose purpose is already appreciated: the act of using a tool to get food, for example. Even to intelligently combine the meanings of already-understood actions (to understand, for example, someone’s pointing to a tool that is being used to get food) would be a cognitive task whose successful undertaking plausibly requires more than the kinds of circuitry discussed in the target article alone.

This kind of worry is also prominent in the commentary by **Preston**, who, while agreeing that many basic perception-action mechanisms are preserved in our higher-level understandings, notes that such mechanisms require the agent to already command an understanding of (or at

least, some form of representation of) the type of action or state at issue. Closely related issues are raised by **van Rooij, Haselager, & Bekkering** [van Rooij et al.], who note that direct simulation (used to mirror the means-end structure of observed actions) is often inadequate to reveal the goal of (or the intentions behind) an action. This is because our actual understanding of the operative goals and intentions is often profoundly affected by the context in which the action occurs. As a result, there is no one-to-one relation between actions (conceived as sets of motor signals) and goals. **Chakrabarti & Baron-Cohen**, in their incisive discussion of some possible shortcomings of the SCM, point out that in many cases one needs to understand that the intention of the other is that you should do something different to (but complementary to) their own action. For example, two people can carry a heavy log, but not by copying each other’s actions, which will be log-end-specific. Here, the automatic activity of the mirroring system seems more of a liability than an asset (but see our remarks later on the commentary by **Hove**, for one possible solution, consistent with the spirit of the SCM).

Yet another way of raising the same kind of issue is usefully displayed by **Pagliari & Castelfranchi**, who suggest that SCM is hamstrung by its failure to consider some additional roles of goal-states, namely, their role in (not just control but also) “evaluation and motivation”: that is, in deciding upon appropriate goals for our own and others’ actions, and in evaluating the actions in terms of those newly arrived-at goals. These elements are clearly central both to the understanding and the generation of intentional action. But they do not seem to be naturally captured, at least in any of the more advanced flavours we have just been discussing, by the bedrock story about mirroring and shared information spaces for perception and action.

All this, we feel, is exactly as it should be. It was not Hurley’s aim to offer a single kind of mechanism as a cognitive panacea, capable (all on its own) of explaining all aspects of human intelligent performance. Rather, the story is better seen as an attempt to display one key ingredient in such stories – one that has the virtue of first appearing (in basic form) at quite low levels of cognitive sophistication, and then making a contribution at many later stages. But intelligent human performance is not to be understood as flowing solely from the operation of that key ingredient alone. Rather, the ingredient is one enabling element in a larger story, whose full shape has yet to be determined.

Several commentators made helpful suggestions concerning such additional mechanisms. The basic story, **Whiten** suggests, might well need to be combined with an account that displays a developing capacity for “secondary representation” (Perner 1991), allowing us to maintain multiple perspectives simultaneously on a single physical event. **Iacoboni** reports the exciting discovery, in human frontal cortex, of what Iacoboni and Dapretto (2006) dub “super mirror neurons” that seem to modulate the activity of standard mirror neurons in various ways. Such modulatory effects look to offer one possible mechanism by means of which an increasingly intelligent use of the mirror system itself may be enabled. **Longo & Bertenthal**, while also noting the general limitation that mirroring requires the presence of the target action in one’s own

repertoire, report work (Longo et al., in press) that puts this fact to use as a way of demonstrating the existence of mirroring at many levels of grain or abstraction. Longo & Bertenthal also report developmental studies showing the “progression of inhibitory control over mirroring responses.” Therefore, although there is clearly an important issue to be resolved concerning the increasingly intelligent use of shared circuits, there seems no reason to doubt their role as a functional element in a wide variety of sophisticated forms of reason and understanding. We agree with **Goldman**, however, that it is not clear that control theory alone will provide a sufficient framework within which to accommodate and understand the full gamut of mechanisms active in our understanding of self and of others.

One specific area where many commentators felt that the SCM fell short was in accounting for various forms of joint action and joint attention. Thus, **Hove** notes that interpersonal synchrony raises issues that do not arise in most cases of imitation and mirroring. To synchronize our actions with those of another (as in the log-carrying case suggested by **Chakrabarti & Baron-Cohen**), we may need to predict what the other will do and to act accordingly.

Semin & Cacioppo go further, arguing that despite presenting itself as a model of social cognition, SCM as it stands is an individualistic model that fails to account for the co-regulation of action or the distributed nature of real social cognition. Thus, imitative behaviour, for example, is not just about the reproduction of behaviour but additionally helps establish connections between individuals that support the co-regulation of action. The point about establishing connections is expanded by **Nielsen**, who notes that some forms of imitation and interpersonal synchrony may be best understood as what Freeman (2000) calls “technologies of social bonding.” The point about co-regulation, if we understand it correctly, is that the resources that guide and explain the behaviours of the collective (which may be as small as two) are themselves distributed across the agents and (perhaps) aspects of the situation. Examples of co-regulated behaviours include cases of mutual entrainment, such as rhythmic clapping, and cases of complementary action of the kind previously described (where a complex common task requires different but matching actions by multiple agents).

Despite laying out this missing territory in compelling detail, **Semin & Cacioppo** remain silent on just how such phenomena may best be accommodated. A promising suggestion is made by **Hove**, who notes that one key may be the use of the kinds of predictive simulation stressed by the SCM, but with some of the predictions targeting the actions of others and the joint effects of the actions of self and other. A concrete example of how shared circuits may contribute to one specific form of joint action is given in the thoughtful contribution by **Garrod & Pickering**, who focus on the potential role of such circuits in dialogue. Here, there is emerging evidence that agents use their own production systems to generate predictions about the other person’s speech output, in a way that aids their own comprehension. This is a neat example of one way in which the kinds of action/perception found in layer 3 of the SCM may contribute to what are intuitively much “higher” cognitive capacities.

Several commentators note a *prima facie* challenge to Hurley’s heavy use of control theory as a framework for SCM. Thus, **Goldman** notes that although there is strong evidence for the role of efferent copy and reafferent input in the domain of perception and action, no such body of evidence exists for many of the other domains (such as that of pain, feelings, and emotions – for the latter case, see also the contribution by **Preston** and our own comments in section R2) where various shared circuits also seem to enable mirroring and simulation to occur. This calls into question, Goldman suggests, the guiding idea that a control-theoretic perspective is apt as a general framework for all “shared circuit”-style phenomena. In its place, he proposes a Hebbian learning paradigm in which associative learning binds together various forms of neural activation. **Oberman & Ramachandran** reject the Hebbian alternative as a sufficient account of the development of F5 mirror neurons themselves, on the grounds that one still needs to explain why some F5 neurons end up having mirror properties while others do not. Goldman might (indeed, probably would) accept the existence of what Oberman & Ramachandran call “specialized mechanisms and hardwired constraints” for this special population, while still rejecting (but again, see our comments in section R2) any generalization of the control-theoretic explanatory apparatus to other domains in which mirroring and mindreading also seem to occur. At least one of us (Clark) is inclined to the view that the choice of a single control-theoretic perspective to address all such phenomena would indeed be premature. It seems unlikely that all the work required can be achieved by any single kind of mechanism. Nonetheless, the attempt to display a wide variety of mirror-system phenomena from a control-theoretic perspective strikes us as eminently worthwhile. Subsequent departures from that perspective, and the exploration of additional kinds of mechanism and explanatory framework, can then be motivated and described on a case-by-case basis.

We would like to end by flagging, once again, what we take to be the central contribution of the SCM, which is the suggestion that social cognition is continuous with more basic cases in which we perceive the actions of others by means that involve (and not merely as collateral effects or learnt associations) one’s own capacities for similar actions. In this way, Hurley posits a “shared information space” as a starting point for our explorations of interpersonal space and as a lever for our coordinated action. The problem facing the intelligent agent is then, not so much how to learn about the minds of others, as how to separate her own mind from the minds of others. Insofar as this is correct, it turns much of the standard discussion inside out. “Mindreading” becomes the norm, though at the cost (*Star Trek* fans will recognize) of a Borg-like threat of mutual cognitive dissolution. By monitored inhibition of output, we nonetheless end up extruding a genuine (but perhaps fragile?) self/other distinction in the face of (and without ever disabling) those basic tendencies of automatic copying and simulation. Such a story, if it is true, matters in ways that go far beyond the immediate concerns of the cognitive scientific community. It matters for policy, for education, for psychiatry, and for our own self-understanding as a species. Hurley herself was keenly aware of this larger picture, and we would recommend that interested readers consult her powerful

paper, Hurley (2006a), revealingly entitled “Bypassing Conscious Control: Media Violence, Unconscious Imitation, and Freedom of Speech.”

Much, to be sure, remains unresolved. Hurley’s story, at least in its broadest outlines, is compatible (as many commentators rightly observed) with a wide variety of ways of “filling in the mechanisms” and of linking (or even identifying) the putative layers. But whatever the details, there seems something deeply right about the guiding spirit. That spirit is a vision of the human mind as fundamentally social, as an evolved organ not of solipsistic individual cognizing, but of social and communal *co-cognizing*. That kind of talk is not unfamiliar (especially to those working in developmental science), but it has not yet informed the shape of the cognitive scientific mainstream. Hurley’s great achievement is to place this kind of model center stage, and to do so in a way that – as we have seen – is both concrete enough to raise questions of detail, scope, and adequacy, yet general enough to invite constructive elaboration for many years to come.

ACKNOWLEDGMENTS

This Response was prepared thanks to support for both authors from the AHRC, under the ESF Eurocores CNCC scheme, for the CONTACT (Consciousness in Interaction) project AH/E511139/1. We thank Till Vierkant and members of the Philosophy, Psychology and Informatics reading group at Edinburgh for helpful discussion of the target article.

References

[Letters “a” and “r” appearing before authors’ initials refer to target article and response, respectively.]

- Adolphs, R. (2002) Recognizing emotion from facial expressions: Psychological and neurological mechanisms. *Behavioral and Cognitive Neuroscience Reviews* 1:21–61. [aSH]
- Adolphs, R., Damasio, H., Tranel, D., Cooper, G. & Damasio, A. R. (2000) A role for somatosensory cortices in the visual recognition of emotions as revealed by three-dimensional lesion mapping. *Journal of Neuroscience* 20:2683–90. [rJK]
- Adolphs, R., Tranel, D., Damasio, H. & Damasio, A. (1994) Impaired recognition of emotion in facial expressions following bilateral damage to the human amygdala. *Nature* 372:669–72. [rJK]
- Agre, P. E. (1997) *Computation and human experience*. Cambridge University Press. [GRS]
- Aicken, M. D., Wilson, A. D., Williams, J. H. & Mon-Williams, M. (2007) Methodological issues in measures of imitative reaction times. *Brain and Cognition* 63:304–308. [JHGW]
- Akins, C., Klein, E. & Zentall, T. (2002) Imitative learning in Japanese quail (*Conturnix japonica*) using the bidirectional control procedure. *Animal Learning and Behavior* 30:275–81. [aSH]
- Akins, C. & Zentall, T. (1996) Imitative learning in male Japanese quail (*Conturnix japonica*) using the two-action method. *Journal of Comparative Psychology* 110:316–20. [aSH]
- (1998) Imitation in Japanese quail: The role of reinforcement of demonstrator responding. *Psychonomic Bulletin and Review* 5:694–97. [aSH]
- Amodio, D. M. & Frith, C. D. (2006) Meeting of minds: The medial frontal cortex and social cognition. *Nature Reviews: Neuroscience* 7:268–77. [JHGW]
- Anisfeld, M. (1979) Interpreting “imitative” responses in early infancy. *Science* 205:214–15. [aSH]
- (1984) *Language development from birth to three*. Erlbaum. [aSH]
- (1991) Neonatal imitation. *Developmental Review* 11:60–97. [aSH]
- (1996) Only tongue protrusion modeling is matched by neonates. *Developmental Review* 16:149–61. [aSH]
- (2005) No compelling evidence to dispute Piaget’s timetable of the development of representational imitation in infancy. In: *Perspectives on imitation: From neuroscience to social science, vol. 2*, ed. S. Hurley & N. Chater, pp. 107–31. MIT Press. [aSH]

- Anisfeld, M., Turkewitz, G., Rose, S., Rosenberg, F., Sheiber, F., Couturier-Fagan, D., Ger, J. & Sommer, I. (2001) No compelling evidence that newborns imitate oral gestures. *Infancy* 2:111–22. [aSH]
- Arbib, M. (2005) From monkey-like action recognition to human language: An evolutionary framework for neurolinguistics. *Behavioral and Brain Sciences* 28(2):105–21. [aSH]
- Arbib, M., Billard, A., Iacoboni, M. & Oztop, E. (2000) Synthetic brain imaging: Grasping, mirror neurons and imitation. *Neural Networks* 13:975–97. [aSH]
- Arbib, M. & Rizzolatti, G. (1997) Neural expectations: A possible evolutionary path from manual skills to language. *Communication and Cognition* 29:393–424. (Reprinted in *The nature of concepts: Evolution, structure, and representation*, ed. P. van Looche, pp. 128–54. Routledge. [aSH, rJK])
- Baldwin, D. (1995) Understanding the link between joint attention and language. In: *Joint Attention: Its origin and role in development*, ed. C. Moore & P. Dunham, pp. 131–58. Erlbaum. [aSH]
- Baldwin, J. (1896) A new factor in evolution. *American Naturalist* 30:441–51, 536–53. [aSH]
- Bargh, J. (1999) The most powerful manipulative messages are hiding in plain sight. *The Chronicle of Higher Education* 45(January 29):B6. [aSH]
- (2005) Bypassing the will: Towards demystifying the nonconscious control of social behavior. In: *The new unconscious*, ed. R. Hassin, J. Uleman & J. Bargh. Oxford University Press. [aSH]
- Bargh, J. & Chartrand, T. (1999) The unbearable automaticity of being. *American Psychologist* 54:462–79. [aSH]
- Bargh, J., Chen, M. & Burrows, L. (1996) The automaticity of social behavior: Direct effects of trait concept and stereotype activation on action. *Journal of Personality and Social Psychology* 71:230–44. [aSH]
- Bargh, J., Gollwitzer, P., Lee-Chai, A., Barndollar, K. & Trötschel, R. (2001) The automated will: Nonconscious activation and pursuit of behavioral goals. *Journal of Personality and Social Psychology* 81:1014–27. [aSH]
- Barkley, R. A. (2001) The executive functions and self-regulation: An evolutionary neuropsychological perspective. *Neuropsychology Review* 11(1):1–29. [aSH]
- Baron-Cohen, S. (1995a) *Mindblindness: An essay on autism and theory of mind*. MIT Press/Bradford Books. [BC]
- (1995b) The eye direction detector (EDD) and the shared attention mechanism (SAM): Two cases for evolutionary psychology. In: *Joint attention: Its role in development*, ed. C. Moore & P. Dunham, pp. 41–60. Erlbaum. [BC]
- Baron-Cohen, S. & Wheelwright, S. (2004) The empathy quotient (EQ): An investigation of adults with Asperger syndrome or high functioning autism, and normal sex differences. *Journal of Autism and Developmental Disorders* 34:163–75. [BC, rJK]
- Behrendt, R. P. (2005) Passivity phenomena: Implications for the concept of self. *Neuro-Psychoanalysis* 7:185–207. [R-PB]
- Bekkering, H. & Wohlschläger, A. (2002) Action perception and imitation: A tutorial. In: *Attention and performance XIX. Common mechanisms in perception and action*, ed. W. Prinz & B. Hommel, pp. 294–314. Oxford University Press. [aSH]
- Bermúdez, J. (2000) Personal and subpersonal: A difference without a distinction. *Philosophical Explorations* (Special Issue) 2:63–82. [aSH]
- (2003) Nonconceptual mental content. In: *The Stanford encyclopedia of philosophy*, Spring 2003 edition, ed. E. N. Zalta. Stanford University Press. Available at: <http://plato.stanford.edu/archives/spr2003/entries/content-nonconceptual/>. [aSH]
- Bertenthal, B. I., Longo, M. R. & Kosobud, A. (2006) Imitative response tendencies following observation of intransitive actions. *Journal of Experimental Psychology: Human Perception and Performance* 32:210–25. [MRL]
- Blackmore, S. (1999) *The meme machine*. Oxford University Press. [aSH]
- (2000) The meme’s eye view. In: *Darwinizing culture: The status of memetics as a science*, ed. R. Aunger, pp. 25–42. Oxford University Press. [aSH]
- (2001) Evolution and memes: The human brain as a selective imitation device. *Cybernetics and Systems* 32:225–55. [aSH]
- Blair, J. & Perschardt, K. (2003) Empathy: A unitary circuit or a set of dissociable neuro-cognitive systems? *Behavioral and Brain Sciences* 25(1):27–28. [BC]
- Blair, R. J., Jones, L., Clark, F. & Smith, M. (1997) The psychopathic individual: A lack of responsiveness to distress cues? *Psychophysiology* 34:192–98. [BC]
- Blakemore, S. (2003) Deluding the motor system. *Consciousness and Cognition* 12:647–55. [LF]
- Blakemore, S. & Decety, J. (2001) From the perception of action to the understanding of intention. *Nature Reviews: Neuroscience* 2:561–67. [aSH]
- Blakemore, S.-J., Frith, C. D. & Wolpert, D. W. (1999) Spatiotemporal prediction modulates the perception of self-produced stimuli. *Journal of Cognitive Neuroscience* 11:551–59. [rJK]
- Bock, L. J. (1986) Syntactic priming in language production. *Cognitive Psychology* 18:355–87. [GRS]
- (1989) Close class immanence in sentence production. *Cognition* 31:163–89. [GRS]

- Bock, L. J. & Loebell, H. (1990) Framing sentences. *Cognition* 35:1–39. [GRS]
- Bonnie, K. E., Horner, V., Whiten, A. & de Waal, F. B. M. (2007) Spread of arbitrary customs among chimpanzees: A controlled experiment. *Proceedings of the Royal Society of London, Series B* 274:367–72. [AW]
- Boyd, R. & Richerson, P. (1982) Cultural transmission and the evolution of cooperative behavior. *Human Ecology* 10:325–51. [aSH]
- (1985) *Culture and the evolutionary process*. University of Chicago Press. [aSH]
- Branigan, H. P., Pickering, M. J. & Cleland, A. A. (2000) Syntactic coordination in dialogue. *Cognition* 75:B13–B25. [SG]
- Brass, M. (1999) Imitation and ideomotor compatibility. Unpublished doctoral dissertation, University of Munich, Germany. [aSH]
- Brass, M., Bekkering, H. & Prinz, W. (2001) Movement observation affects movement execution in a simple response task. *Acta Psychologica* 106:3–22. [aSH]
- Brass, M., Bekkering, H., Wohlschläger, A. & Prinz, W. (2000) Compatibility between observed and executed finger movements: Comparing symbolic, spatial, and imitative cues. *Brain and Cognition* 44:124–43. [JHGW]
- Brass, M., Derrfuss, J., Matthev-von Cramon, G. & von Cramon, D. Y. (2003) Imitative response tendencies in patients with frontal brain lesions. *Neuropsychology* 17(2):265–71. [aSH]
- Brass, M., Derrfuss, J. & von Cramon, D. Y. (2005) The inhibition of imitative and overlearned responses: A functional double dissociation. *Neuropsychologia* 43:89–98. [JHGW]
- Brass, M. & Heyes, C. M. (2005) Imitation: Is cognitive neuroscience solving the correspondence problem? *Trends in Cognitive Sciences* 9:489–95. [CH, r]K, JHGW]
- Brooks, R. (1999) *Cambrian intelligence*. MIT Press. [aSH, GRS]
- Bruner, J. S. (1969) Eye, hand, and mind. In: *Studies in cognitive development: Essays in honor of Jean Piaget*, ed. D. Elkind & J. H. Flavell, pp. 223–35. Oxford University Press. [MRL]
- Buccino, G., Binkofski, F., Fink, G. R., Fadiga, L., Gallese, V., Seitz, R. J., Zilles, K., Rizzolatti, G. & Freund, H.-J. (2001) Action observation activates premotor and parietal areas in a somatotopic manner: An fMRI study. *European Journal of Neuroscience* 13:400–404. [aSH]
- Bylander, T., Allemand, D., Tanner, M. C. & Josephson, J. R. (1991) The computational complexity of abduction. *Artificial Intelligence* 49:25–60. [IVR]
- Byrne, R. (1995) *The thinking ape: Evolutionary origins of intelligence*. Oxford University Press. [aSH]
- (1998) Imitation: The contributions of priming and program-level copying. In: *Intersubjective communication and emotion in early ontogeny*, ed. S. Braten, pp. 228–44. Cambridge University Press. [aSH]
- (1999) Imitation without intentionality: Using string parsing to copy the organization of behavior. *Animal Cognition* 2:63–72. [aSH]
- (2002a) Imitation of complex novel actions: What does the evidence from animals mean? *Advances in the Study of Behavior* 31:77–105. [aSH]
- (2002b) Seeing actions as hierarchically organized structures: Great ape manual skills. In: *The imitative mind*, ed. A. Meltzoff & W. Prinz, pp. 122–40. Cambridge University Press. [aSH]
- (2005) Detecting, understanding, and explaining animal imitation. In: *Perspectives on imitation: From neuroscience to social science, vol. 1*, ed. S. Hurley & N. Chater, pp. 225–42. MIT Press. [aSH]
- Byrne, R. & Russon, A. (1998) Learning by imitation: A hierarchical approach. *Behavioral and Brain Sciences* 21:667–721. [aSH, r]K]
- Byrne, R. & Whiten, A., eds. (1988) *Machiavellian intelligence: Social expertise and the evolution of intellect in monkeys, apes and humans*. Oxford University Press. [aSH]
- Call, J., Agnetta, B. & Tomasello, M. (2000) Social cues that chimpanzees do and do not use to find hidden objects. *Animal Cognition* 3:23–34. [aSH]
- Call, J. & Carpenter, M. (2002) Three sources of information in social learning. In: *Imitation in animals and artifacts*, ed. K. Dautenhahn & C. Nehaniv, pp. 211–28. MIT Press. [aSH]
- Call, J. & Tomasello, M. (1994) The social learning of tool use by orangutans (*Pongo pygmaeus*). *Human Evolution* 9:297–313. [aSH]
- (1999) A nonverbal theory of mind test: The performance of children and apes. *Child Development* 70:381–95. [aSH]
- Calvo-Merino, B., Glaser, D. E., Grèzes, J., Passingham, R. E. & Haggard, P. (2005) Action observation and acquired motor skills: An fMRI study with expert dancers. *Cerebral Cortex* 15:1243–49. [MRL]
- Caporael, L. (1997) The evolution of truly social cognition: The core configurations model. *Personality and Social Psychology Review* 1:276–98. [GRS]
- Carpendale, J. I. M. & Lewis, C. (2004) Constructing an understanding of mind: The development of children's social understanding within social interaction. *Behavioral and Brain Sciences* 27(2):79–151. [JIMC]
- (2006) *How children develop social understanding*. Blackwell. [JIMC]
- Carpenter, M., Akhtar, N. & Tomasello, M. (1998) Fourteen- through 18-month-old infants differentially imitate intentional and accidental actions. *Infant Behavior and Development* 21:315–30. [aSH]
- Carpenter, M., Tomasello, M. & Striano, T. (2005) Role reversal imitation and language in typically developing infants and children with autism. *Infancy* 8(3):253–78. [BC]
- Castelfranchi, C. (1998) Modelling social action for AI agents. *Artificial Intelligence* 103:157–82. [FP]
- Castelfranchi, C. & Paglieri, F. (2007) The role of beliefs in goal dynamics: Prolegomena to a constructive theory of intentions. *Synthese* 155:237–63. [FP]
- Catmur, C., Walsh, V. & Heyes, C. M. (2007) Sensorimotor learning configures the human mirror system. *Current Biology* 17:1527–31. [CH]
- Chakrabarti, B. & Baron-Cohen, S. (2006) Empathizing: Neurocognitive developmental mechanisms and individual differences. *Progress in Brain Research* 156:406–17. (Special issue on “Understanding Emotions”.) [BC]
- Chakrabarti, B., Bullmore, E. T. & Baron-Cohen, S. (2006) Empathizing with basic emotions: Common and discrete neural substrates. *Social Neuroscience* 1(3–4):364–84. [BC]
- Chartrand, T. & Bargh, J. (1999) The chameleon effect: The perception-behavior link and social interaction. *Journal of Personality and Social Psychology* 76:893–910. [SG, CH, MJH, aSH, MRL, GRS]
- Chartrand, T., Maddux, W. & Lakin, J. (2005) Beyond the perception-behavior link: The ubiquitous utility and motivational moderators of nonconscious mimicry. In: *The new unconscious*, ed. R. Hassin, J. Uleman & J. Bargh. Oxford University Press. [aSH]
- Christiansen, M. (1994) Infinite languages, finite minds: Connectionism, learning and linguistic structure. Unpublished doctoral dissertation, University of Edinburgh. [aSH]
- (2005) On the relation between language and (mimetic) culture. In: *Perspectives on imitation: From neuroscience to social science, vol. 2*, ed. S. Hurley & N. Chater, pp. 391–96. MIT Press. [aSH]
- Christiansen, M. H. & Kirby, S. (2003) Language evolution: Consensus and controversies. *Trends in Cognitive Science* 7(7):300–307. [aSH]
- Clark, A. (1997) *Being there: Putting brain, body, and world together again*. MIT Press. [GN]
- (1999) An embodied cognitive science? *Trends in Cognitive Sciences* 3(9):345–51. [GN]
- Colby, C. L. & Goldberg, M. E. (1999) Space and attention in parietal cortex. *Annual Review of Neuroscience* 22:319–49. [R-PB]
- Cooper, G. F. (1990) The computational complexity of probabilistic inference using Bayesian belief networks. *Artificial Intelligence* 42(2–3):393–405. [IVR]
- Craighero, L., Buccino, G. & Rizzolatti, G. (2002) Speech listening specifically modulates the excitability of tongue muscles: A TMS study. *European Journal of Neuroscience* 15:399–402. [aSH]
- Csibra, G. (2005) Mirror neurons and action observation: Is simulation involved? In: *What do mirror neurons mean?* Interdisciplines Web Forum, available at: <http://www.interdisciplines.org/mirror/papers/4>. [aSH]
- Csibra, G. & Gergely, G. (2007) “Obsessed with goals”: Functions and mechanisms of teleological interpretation of actions in humans. *Acta Psychologica* 124:60–78. [FP]
- Cuijpers, R. H., van Schie, H. T., Koppen, M., Erhagen, W. & Bekkering, H. (2006) Goals and means in action observation: A computational approach. *Neural Networks* 19:311–22. [IVR]
- Damasio, A. R. (2003) *Looking for Spinoza*. Harcourt. [AIG]
- Danielson, P. (1991) Closing the compliance dilemma: How it's rational to be moral in a Lamarekian world. In: *Contractarianism and rational choice*, ed. P. Vallentyne, pp. 291–322. Cambridge University Press. [aSH]
- (1992) *Artificial morality: Virtuous robots for virtual games*. Routledge. [aSH]
- Davies, M. & Stone, T. (1995a) *Folk psychology*. Blackwell. [aSH]
- (1995b) *Mental simulation*. Blackwell. [aSH]
- Dawkins, R. (1976/1989) *The selfish gene*. Oxford University Press (Original publication 1976; second edition 1989.) [aSH]
- (1982) *The extended phenotype*. Oxford University Press. [aSH]
- de Lange, F. P., Spronk, M., Willems, R. M., Toni, I. & Bekkering, H. (submitted) Complementary systems for understanding action intentions. [IVR]
- de Vignemont, F. & Foucheret, P. (2004) The sense of agency: A philosophical and empirical review of the “Who” system. *Consciousness and Cognition* 13:1–19. [LF]
- Deacon, T. (1997) *The symbolic species: The coevolution of language and the human brain*. Penguin Books/Norton. [aSH]
- Decety, J. & Chaminade, T. (2003) Neural correlates of feeling sympathy. *Neuropsychologia* 41(2):127–38. [aSH]
- (2005) The neurophysiology of imitation and intersubjectivity. In: *Perspectives on imitation: From neuroscience to social science, vol. 1*, ed. S. Hurley & N. Chater, pp. 119–40. MIT Press. [aSH]
- Decety, J., Grèzes, J., Costes, N., Perani, D., Jeannerod, M., Procyk, E., Grassi, F. & Fazio, F. (1997) Brain activity during observation of action: Influence of action content and subject's strategy. *Brain* 120:1763–77. [aSH]
- deCharms, R. C. & Zador, A. (2000) Neural representation and the cortical code [Review] *Annual Review of Neuroscience* 23:613–47. [GN]

- Dennett, D. (1969) *Content and consciousness*. Routledge & Kegan Paul. [aSH]
- (1987) *The intentional stance*. MIT Press. [FP]
- (1991) *Consciousness explained*. Little, Brown. [aSH]
- (1995) *Darwin's dangerous idea: Evolution and the meanings of life*. Simon & Schuster/Penguin. [aSH]
- Dijksterhuis, A. (2005) Why we are social animals: The high road to imitation as social glue. In: *Perspectives on imitation: From neuroscience to social science, vol. 2: Imitation, human development, and culture*, ed. S. Hurley & N. Chater, pp. 207–20. MIT Press. [aSH, MN]
- Dijksterhuis, A. & Bargh, J. (2001) The perception-behavior expressway: Automatic effects of social perception on social behavior. *Advances in Experimental Social Psychology* 33:1–40. [aSH]
- Dijksterhuis, A. & van Knippenberg, A. (1998) The relation between perception and behavior or how to win a game of Trivial Pursuit. *Journal of Personality and Social Psychology* 74:865–77. [aSH]
- Di Pellegrino, G., Fadiga, L., Fogassi, L., Gallese, V. & Rizzolatti, G. (1992) Understanding motor events: A neurophysiological study. *Experimental Brain Research* 91:176–80. [aSH]
- Dimberg, U. & Oehman, A. (1996) Behold the wrath: Psychophysiological responses to facial stimuli. *Motivation and Emotion* 20:149–82. [SDP]
- Dimberg, U., Thunberg, M. & Elmehed, K. (2000) Unconscious facial reactions to emotional facial expressions. *Psychological Science* 11:86–89. [GRS]
- Dindo, M., Thierry, B. & Whiten, A. (2007) Social diffusion of novel foraging methods in brown capuchin monkeys (*Cebus apella*). *Proceedings of the Royal Society of London, Series B*. DOI:10.1098/rspb.2007.1318. [Online publication.] [AW]
- Eidelberg, L. (1929) Experimenteller Beitrag zum Mechanismus der Imitationsbewegung. *Jahresbucher fur Psychiatrie und Neurologie* 45:170–73. [aSH]
- Elsner, B. & Hommel, B. (2001) Effect anticipation and action control. *Journal of Experimental Psychology: Human Perception and Performance* 27:229–40. [R-PB]
- Elton, M. (2000) Consciousness: Only at the personal level. *Philosophical Explorations* (Special Issue) 3(1):25–42. [aSH]
- Fadiga, L., Craighero, L., Buccino, G. & Rizzolatti, G. (2002) Speech listening specifically modulates the excitability of tongue muscles: A TMS study. *European Journal of Neuroscience* 15:399–402. [SG, aSH]
- Fadiga, L., Fogassi, L., Pavesi, G. & Rizzolatti, G. (1995) Motor facilitation during action observation: A magnetic stimulation study. *Journal of Neurophysiology* 73:2608–11. [CH, aSH]
- Feinberg, T. & Keenan, J. (2005) Where in the brain is the self? *Consciousness and Cognition* 14:661–78. [LF]
- Ferguson, M. J. & Bargh, J. A. (2004) How social perception can automatically influence behavior. *Trends in Cognitive Sciences* 8:33–39. [R-PB]
- Fernyhough, C. (2004) Alien voices and inner dialogue: Towards a developmental account of auditory verbal hallucinations. *New Ideas in Psychology* 22:49–68. [LF]
- Ferrari, P. F., Visalberghi, E., Paukner, A., Fogassi, L., Ruggiero, A. & Suomi, S. J. (2006) Neonatal imitation in rhesus macaques. *PLoS Biology* 4:e302. (Online journal.) [MRL]
- Fiske, A. P. (1992) *Structures of social life*. Free Press. [GRS]
- Flanagan, J., Vetter, P., Johansson, R. & Wolpert, D. (2003) Prediction precedes control in motor learning. *Current Biology* 13:146–50. [aSH]
- Fogassi, L., Ferrari, P. F., Gesierich, B., Rozzi, S., Chersi, F. & Rizzolatti, G. (2005) Parietal lobe: From action organization to intention understanding. *Science* 308:662–67. [aSH]
- Ford, K. M. & Pylyshyn, Z. W., eds. (1996) *The robot's dilemma revisited: The frame problem in artificial intelligence*. Ablex. [IvR]
- Freeman, W. (2000) A neurobiological role of music in social bonding. In: *The origins of music*, ed. N. L. Wallin, B. Merker & S. Brown, pp. 411–24. MIT Press. [MJH, rJK]
- Frijda, N. H. (1986) *The emotions*. Cambridge University Press. [FP]
- Friston, K. J. (1997) Another neural code? *NeuroImage* 5(3):213–20. [GN]
- Frith, C. (1992) *The cognitive neuropsychology of schizophrenia*. Erlbaum/Taylor & Francis. [aSH]
- Frith, C., Blakemore, S. & Wolpert, D. (2000a) Abnormalities in the awareness and control of action. *Philosophical Transactions of the Royal Society, Series B: Biological Sciences* 355:1771–88. [aSH]
- (2000b) Explaining the symptoms of schizophrenia: Abnormalities in the awareness of action. *Brain Research: Brain Research Reviews* 31:357–63. [LF]
- Frith, C. & Wolpert, D. (2004) *The neuroscience of social interaction*. Oxford University Press. [aSH]
- Frith, U. (2001) Mind blindness and the brain in autism. *Neuron* 32:969–79. [BC]
- Furuyama, N., Hayashi, K. & Mishima, H. (2005) Interpersonal coordination among articulations, gesticulations, and breathing movements: A case of articulation of /a/ and flexion of the wrist. In: *Studies in perception and action*, ed. H. Heft & K. L. Marsh, pp. 45–48. Erlbaum. [GRS]
- Galef, B. (1988) Imitation in animals: History, definition, and interpretation of data from the psychological laboratory. In: *Social learning: Psychological and biological perspectives*, ed. T. Zentall & B. Galef, pp. 3–28. Erlbaum. [aSH]
- (1998) Recent progress in the study of imitation and social learning in animals. In: *Advances in psychological science, vol. 2: Biological and cognitive aspects*, ed. M. Sabourin, F. Craik & M. Roberts, pp. 275–79. Psychological Press. [aSH]
- (2005) Breathing new life into the study of animal imitation: What and when do chimpanzees imitate? In: *Perspectives on imitation: From neuroscience to social science, vol. 1*, ed. S. Hurley & N. Chater, pp. 295–97. MIT Press. [aSH]
- Gallagher, S. (2005) *How the body shapes the mind*. Oxford University Press. [rJK]
- Gallese, V. (2000) The inner sense of action: Agency and motor representations. *Journal of Consciousness Studies* 7(10):23–40. [aSH]
- (2001) The “shared manifold” hypothesis: From mirror neurons to empathy. *Journal of Consciousness Studies* 8:33–50. [MJH, aSH]
- (2003) The manifold nature of interpersonal relations: The quest for a common mechanism. *Philosophical Transactions of the Royal Society of London, Series B: Biological Sciences* 358:517–28. [aSH, AIG]
- (2005) “Being like me”: Self-other identity, mirror neurons and empathy. In: *Perspectives on imitation: From neuroscience to social science, vol. 1*, ed. S. Hurley & N. Chater, pp. 101–18. MIT Press. [aSH]
- Gallese, V., Fadiga, L., Fogassi, L. & Rizzolatti, G. (1996) Action recognition in the premotor cortex. *Brain* 119:593–609. [FP]
- Gallese, V. & Goldman, A. (1998) Mirror neurons and the simulation theory of mindreading. *Trends in Cognitive Sciences* 2:493–501. [aSH, FP]
- Gallese, V., Keysers, C. & Rizzolatti, G. (2004) A unifying view of the basis of social cognition. *Trends in Cognitive Sciences* 8(9):396–403. [aSH]
- Garrod, S. & Anderson, A. (1987) Saying what you mean in dialogue: A study in conceptual and semantic co-ordination. *Cognition* 27:181–218. [SG]
- Garrod, S. & Pickering, M. J. (2004) Why is conversation so easy? *Trends in Cognitive Sciences* 8:8–11. [R-PB, SG]
- Gazzola, V., Aziz-Zadeh, L. & Keysers, C. (2006) Empathy and the somatotopic auditory mirror system in humans. *Current Biology* 16:1824–29. [JHGW]
- Gergely, G. (2003) What should a robot learn from an infant? Mechanisms of action interpretation and observational learning in infancy. *Connection Science* 15:191–209. [MN]
- Gergely, G., Bekkering, H. & Király, I. (2002) Rational imitation in preverbal infants. *Nature* 415:755. [aSH, MN, IvR, JHGW]
- Gil-White, F. (2005) Common misunderstandings of memes (and genes): The promise and the limits of the genetic analogy to cultural transmission processes. In: *Perspectives on imitation: From neuroscience to social science, vol. 2*, ed. S. Hurley & N. Chater, pp. 317–38. MIT Press. [aSH]
- Goldenberg, G. (2003) Apraxia and beyond: Life and work of Hugo Liepmann. *Cortex* 39:509–24. [JHGW]
- Goldman, A. (1989) Interpretation psychologized. *Mind and Language* 4:161–85. [aSH]
- (1992) In defense of the simulation theory. *Mind and Language* 7:104–19. [aSH]
- (2005) Imitation, mindreading, and simulation. In: *Perspectives on imitation: From neuroscience to social science, vol. 2*, ed. S. Hurley & N. Chater, pp. 79–93. MIT Press. [aSH, rJK]
- (2006) *Simulating minds: The philosophy, psychology and neuroscience of mindreading*. Oxford University Press. [AIG, rJK]
- (in press) Mirroring, mindreading, and simulation. In: *Mirror neuron systems: The role of mirroring processes in social cognition*, ed. J. Pineda. Humana Press. [AIG]
- Goldman, A. I. & Sripada, C. (2005) Simulationist models of face-based emotion recognition. *Cognition* 94:193–213. [AIG, rJK]
- Gopnik, A. & Wellman, H. M. (1992) Why the child's theory of mind really is a theory. *Mind and Language* 7(1–2):145–71. [rJK]
- (1994) The theory-theory. In: *Mapping the mind: Domain specificity in cognition and culture*, ed. L. Hirschfield & S. Gelman, pp. 257–93. Cambridge University Press. [rJK]
- Gordon, R. (1986) Folk psychology as simulation. *Mind and Language* 1:159–71. [aSH]
- (1995a) Simulation without introspection or inference from me to you. In: *Folk psychology*, ed. M. Davies & T. Stone, pp. 53–67. Blackwell. [aSH]
- (1995b) Sympathy, simulation, and the impartial spectator. *Ethics* 105:727–42. [aSH]
- (1996) “Radical” simulationism. In: *Theories of theories of mind*, ed. P. Carruthers & P. Smith, pp. 11–21. Cambridge University Press. [aSH]
- (2002) Simulation and reason explanation: The radical view. *Philosophical Topics* (Special Issue) 29:175–92. [aSH]
- (2005) Intentional agents like myself. In: *Perspectives on imitation: From neuroscience to social science, vol. 2*, ed. S. Hurley & N. Chater, pp. 95–106. MIT Press. [aSH]
- Gray, J. (1991) The neuropsychology of schizophrenia. *Behavioral and Brain Sciences* 14(1):1–84. [aSH]

- (2004) *Consciousness: Creeping up on the hard problem*. Oxford University Press. [aSH]
- Craiano, M., Taylor, C., Moore, T. & Cooke, D. (2002) The cortical control of movement revisited. *Neuron* 36:349–62. [aSH]
- Greenwald, A. (1970) Sensory feedback mechanisms in performance control: With special reference to the ideomotor mechanism. *Psychological Review* 77:73–99. [aSH]
- (1972) On doing two things at once: Time sharing as a function of ideomotor compatibility. *Journal of Experimental Psychology* 94:52–57. [aSH]
- Grèzes, J. & Decety, J. (2001) Functional anatomy of execution, mental simulation, observation, and verb generation of actions: A meta-analysis. *Human Brain Mapping* 12:1–19. [R-PB]
- Grush, R. (1995) *Emulation and cognition*. Doctoral dissertation, Department of Philosophy, University of California at San Diego. [aSH]
- (2004) The emulation theory of representation: Motor control, imagery, and perception. *Behavioral and Brain Sciences* 27:377–442. [aSH]
- Hare, B., Call, J., Agnetta, B. & Tomasello, M. (2000) Chimpanzees know what conspecifics do and do not see. *Animal Behaviour* 59:771–85. [aSH]
- Hare, B., Call, J. & Tomasello, M. (2001) Do chimpanzees know what conspecifics know and do not know? *Animal Behaviour* 61:139–51. [aSH]
- Hare, B. & Tomasello, M. (2004) Chimpanzees are more skillful at competitive than cooperative tasks. *Animal Behaviour* 68:571–81. [JIMC]
- Hari, R., Forss, N., Avikainen, S., Kirveskari, E., Salenius, S. & Rizzolatti, G. (1998) Activation of human primary motor cortex during action observation: A neuro-magnetic study. *Proceedings of the National Academy of Sciences USA* 95:15061–65. [aSH]
- Harris, P. & Want, S. (2005) On learning what not to do: The emergence of selective imitation in young children's tool use. In: *Perspectives on imitation: From neuroscience to social science*, vol. 2, ed. S. Hurley & N. Chater, pp. 148–62. MIT Press. [aSH]
- Haruno, M., Wolpert, D. & Kawato, M. (2001) Mosaic model for sensorimotor learning and control. *Neural Computation* 13:2201–20. [aSH]
- Haselager, W. F. G. (1997) *Cognitive science and folk psychology: The right frame of mind*. Sage. [IvR]
- Heiser, M., Iacoboni, M., Maeda, F., Marcus, J. & Mazziotta, J. C. (2003) The essential role of Broca's area in imitation. *European Journal of Neuroscience* 17:1123–28. [aSH]
- Henrich, J. & Boyd, R. (1998) The evolution of conformist transmission and the emergence of between-group differences. *Evolution and Human Behavior* 19:215–41. [aSH]
- Henrich, J. & Gil-White, F. (2001) The evolution of prestige: Freely conferred status as a mechanism for enhancing the benefits of cultural transmission. *Evolution and Human Behavior* 22:165–96. [aSH]
- Herman, L. (2002) Vocal, social, and self-imitation by bottlenosed dolphins. In: *Imitation in animals and artifacts*, ed. K. Dautenhahn & C. Nehaniv, pp. 63–106. MIT Press. [aSH]
- Hesslow, G. (2002) Conscious thought as simulation of behaviour and perception. *Trends in Cognitive Sciences* 6:242–47. [aSH]
- Heyes, C. (1996) Genuine imitation? In: *Social learning in animals: The roots of culture*, ed. C. Heyes & B. Galef Jr., pp. 371–89. Academic Press. [aSH]
- (1998) Theory of mind in nonhuman primates. *Behavioral and Brain Sciences* 21:101–14. [aSH]
- (2001) Causes and consequences of imitation. *Trends in Cognitive Sciences* 5:253–61. [CH, aSH]
- (2002) Transformational and associative theories of imitation. In: *Imitation in animals and artifacts*, ed. K. Dautenhahn & C. Nehaniv, pp. 501–24. MIT Press. [rJK]
- (2005) Imitation by association. In: *Perspectives on imitation: From neuroscience to social science*, vol. 1: *Mechanisms of imitation and imitation in animals*, ed. S. Hurley & N. Chater, pp. 157–76. MIT Press. [AIG, aSH, rJK]
- Heyes, C. & Dickinson, A. (1993) The intentionality of animal action. In: *Consciousness*, ed. M. Davies & G. Humphreys. Blackwell. [aSH]
- Heyes, C. & Galef, B., eds. (1996) *Social learning in animals: The roots of culture*. Academic. [aSH]
- Hobson, P. & Lee, A. (1999) Imitation and identification in autism. *Journal of Child Psychology and Psychiatry* 40:464–59. [rJK]
- Hommel, B. (2004) Event files: Feature binding in and across perception and action. *Trends in Cognitive Sciences* 11:494–500. [GN]
- Hommel, B., Musseler, J., Aschersleben, G. & Prinz, W. (2001) The Theory of Event Coding (TEC): A framework for perception and action planning. *Behavioral and Brain Sciences* 24(5):849–78; discussion, 878–937. [GN]
- Hopper, L. M., Lambeth, S. P., Schapiro, S. J. & Whiten, A. (in press) Observational learning in chimpanzees and children studied through “ghost” conditions. *Proceedings of the Royal Society of London, B: Biological Sciences*. [AW]
- Hopper, L. M., Spiteri, A., Lambeth, S. P., Schapiro, S. J., Horner, V. & Whiten, A. (2007) Experimental studies of traditions and underlying transmission processes in chimpanzees. *Animal Behaviour* 73:1021–32. [AW]
- Horner, V. & Whiten, A. (2005) Causal knowledge and imitation/emulation switching in chimpanzees (*Pan troglodytes*) and children (*Homo sapiens*). *Animal Cognition* 8:164–81. [JHGW]
- Horner, V., Whiten, A., Flynn, E. & de Waal, F. B. M. (2006) Faithful replication of foraging techniques along cultural transmission chains by chimpanzees and children. *Proceedings of the National Academy of Sciences, USA* 103:13878–83. [AW]
- Hornsby, J. (2000) Personal and sub-personal: A defence of Dennett's early distinction. *Philosophical Explorations* (Special Issue) 3:6–24. [aSH]
- Hove, M. J. & Risen, J. L. (submitted) It's all in the timing: Interpersonal synchrony increases affiliation. [MJH]
- Howard, J. (1988) Co-operation in the prisoner's dilemma. *Theory and Decision* 24:203–13. [aSH]
- Hunt, G. & Gray, R. (2003) Diversification and cumulative evolution in New Caledonian crow tool manufacture. *Proceedings of the Royal Society London, Series B: Biological Sciences* 270:867–74. [aSH]
- Hurley, S. (2001) Perception and action: Alternative views. *Synthese* 291:3–40. [aSH]
- (2003) *Justice, luck, and knowledge*. Harvard University Press. [rJK]
- (2005a) Social heuristics that make us smarter. *Philosophical Psychology* 18(5):585–611. [aSH, rJK]
- (2005b) The shared circuits model: How control, mirroring and simulation can enable imitation and mindreading. In: *What do mirror neurons mean?* Inter-disciplinary Web Forum. Available at: <http://www.interdisciplines.org/mirror/papers/5>. [aSH]
- (2006a) Bypassing conscious control: Media violence, unconscious imitation, and freedom of speech. In: *Does consciousness cause behavior? An investigation of the nature of volition*, ed. S. Pockett, W. Banks & S. Gallagher, pp. 301–39. MIT Press. [rJK]
- (2006b) Making sense of animals. In: *Rational animals?* ed. S. Hurley & M. Nudds. Oxford University Press. (Revised version of Hurley, S. [2003] Animal action in the space of reasons. *Mind and Language* 18:231–56.) [aSH]
- (2007) Commentary on Alvin Goldman's *Simulating Minds*. Paper read at the American Philosophical Association Pacific Division Meeting 2007 Author meets Critics. (Unpublished manuscript). [rJK]
- (in press) Varieties of externalism. In: *The extended mind*, ed. R. Menary. Ashgate. [aSH]
- Hurley, S. & Chater, N., eds. (2005a) *Perspectives on imitation: From neuroscience to social science*, vols 1 & 2. MIT Press. [aSH]
- (2005b) Introduction: The importance of imitation. In: *Perspectives on imitation: From neuroscience to social science*, vol. 1: *Mechanisms of imitation and imitation in animals*, ed. S. Hurley & N. Chater, pp. 1–52. MIT Press. [aSH, rJK]
- Hurley, S. & Noë, A. (2003) Neural plasticity and consciousness. *Biology and Philosophy* 18:131–68. [aSH]
- Hurley, S. L. (1989) *Natural reasons: Personality and polity*. Oxford University Press. [aSH, rJK]
- (1998) *Consciousness in action*. Harvard University Press. [aSH, rJK]
- Hutchins, E. (1995) *Cognition in the wild*. MIT Press. [GRS]
- Iacoboni, M. (2005) Understanding others: Imitation, language, empathy. In: *Perspectives on imitation: From neuroscience to social science*, vol. 1: *Mechanisms of imitation and imitation in animals*, ed. S. Hurley & N. Chater, pp. 77–99. MIT Press. [aSH, rJK]
- (2008) *Mirroring people*. Farrar, Straus & Giroux. [MI]
- Iacoboni, M. & Dapretto, M. (2006) The mirror neuron system and the consequences of its dysfunction. *Nature Reviews Neuroscience* 7(12):942–51. [MI, rJK]
- Iacoboni, M., Molnar-Szakacs, I., Gallese, V., Buccino, G., Mazziotta, J. & Rizzolatti, G. (2005) Grasping the intentions of others with one's own mirror neuron system. *PLoS Biology* 3(3):529–35 [e79]. [aSH]
- Iacoboni, M., Woods, R. P., Brass, M., Bekkering, H., Mazziotta, J. C. & Rizzolatti, G. (1999) Cortical mechanisms of human imitation. *Science* 286:2526–28. [MJH]
- Jackson, P. L., Brunet, E., Meltzoff, A. N. & Decety, J. (2006) Empathy examined through the neural mechanisms involved in imagining how I feel versus how you feel pain. *Neuropsychologia* 44:752–61. [SDP]
- Jackson, P. L., Meltzoff, A. N. & Decety, J. (2004) How do we perceive the pain of others? A window into the neural processes involved in empathy. *NeuroImage* 24:771–79. [AIG]
- James, W. (1890) *The principles of psychology*, vol. 1. Dover/Macmillan. [R-PB, JHGW]
- Jansson, E., Wilson, A. D., Williams, J. H. & Mon-Williams, M. (2007) Methodological problems undermine tests of the ideomotor conjecture. *Experimental Brain Research* 182:549–58. [JHGW]
- Jeannerod, M. (1997) *The cognitive neuroscience of action*. Blackwell. [aSH]
- (2001) Neural simulation of action: A unifying mechanism for motor cognition. *Neuroimage* 14:S103–S109. [aSH]
- Jones, S. & Fernyhough, C. (2007a) Neural correlates of inner speech and auditory verbal hallucinations: A critical review and theoretical integration. *Clinical Psychology Review* 27:140–54. [LF]

- (2007b) Thought as action: Inner speech, self-monitoring, and auditory verbal hallucinations. *Consciousness and Cognition* 16:391–99. [LF]
- Keller, P. E., Knoblich, G. & Repp, B. H. (2007) Pianists duet better when they play with themselves: On the possible role of action simulation in synchronization. *Consciousness and Cognition* 16:102–11. [MJH]
- Keysers, C. & Gazzola, V. (2006) Towards a unifying neural theory of social cognition. *Progress in Brain Research* 156:383–406. [AIG]
- Keysers, C. & Perrett, D. (2004) Demystifying social cognition: A Hebbian perspective. *Trends in Cognitive Sciences* 8:501–507. [AIG, rJK]
- Keysers, C., Wicker, B., Gazzola, V., Anton, J.-L., Fogassi, L. & Gallese, V. (2004) A touching sight: SII/PV activation during the observation of touch. *Neuron* 42:335–46. [AIG]
- Kilner, J. M., Paulignan, Y. & Blakemore, S. J. (2003) An interference effect of observed biological movement on action. *Current Biology* 13:522–25. [JHGW]
- Kinsbourne, M. (2005) Imitation as entrainment: Brain mechanisms and social consequences. In: *Perspectives on imitation: From neuroscience to social science, vol. 2: Imitation, human development, and culture*, ed. S. Hurley & N. Chater, pp. 163–72. MIT Press. [aSH, rJK]
- Kirsch, D. (1995) The intelligent use of space. *Artificial Intelligence* 73:31–68. [GRS]
- Knoblich, G. & Jordan, J. S. (2003) Action coordination in groups and individuals: Learning anticipatory control. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 29:1006–16. [MJH]
- Krebs, J. & Dawkins, R. (1984) Animal signals: Mindreading and manipulation. In: *Behavioural ecology: An evolutionary approach*, 2nd edition, ed. J. Krebs & N. Davies, pp. 380–402. Blackwell. [aSH]
- Lambie, J. A. & Marcel, A. J. (2002) Consciousness and the varieties of emotion experience: A theoretical framework. *Psychological Review* 109(2):219–59. [GN]
- Lamm, C., Batson, C. D. & Decety, J. (2007) The neural substrate of human empathy: Effects of perspective-taking and cognitive appraisal. *Journal of Cognitive Neuroscience* 19:42–58. [SDP]
- Lee, T.-W., Josephs, O., Dolan, R. J. & Critchley, H. D. (2006) Imitating expressions: Emotion-specific neural substrates in facial mimicry. *Social Cognitive and Affective Neuroscience* 1(2):122–35. [BC]
- Leslie, A. M. (1987) Pretense and representation in infancy: The origins of “theory of mind.” *Psychological Review* 94:412–26. [AW]
- Leube, D., Knoblich, G., Erb, M., Grodd, W., Bartels, M. & Kircher, T. (2003) The neural correlates of perceiving one’s own movements. *NeuroImage* 20:2084–90. [LF]
- Lhermitte, F. (1983) “Utilization behaviour” and its relation to lesions of the frontal lobes. *Brain* 106:237–55. [aSH]
- (1986) Human autonomy and the frontal lobes: Part II. *Annals of Neurology* 19:335–43. [aSH]
- Lhermitte, F., Pillon, B. & Serdaru, M. (1986) Human autonomy and the frontal lobes. Part I: Imitation and utilization behavior: A neuropsychological study of 75 patients. *Annals of Neurology* 19:326–34. [aSH, JHGW]
- Longo, M. R. & Bertenthal, B. I. (2006) Common coding of observation and execution of action in 9-month-old infants. *Infancy* 10:43–59. [MRL]
- Longo, M. R., Kosobud, A. & Bertenthal B. I. (in press) Automatic imitation of biomechanically possible and impossible actions: Effects of priming movements vs. goals. *Journal of Experimental Psychology: Human Perception and Performance*. [MRL, rJK]
- Makino, T. & Aihara, K. (2003) Self-observation principle for estimating the other’s internal state: A new computational theory of communication. Mathematical Engineering Technical Reports METR 2003-36, Department of Mathematical Informatics, Graduate School of Information Science and Technology, University of Tokyo. [TM]
- (2006) Multi-agent reinforcement learning algorithm to handle beliefs of other agents’ policies and embedded beliefs. In: *Proceedings of the Fifth International Joint Conference on Autonomous Agents and Multiagent Systems (AAMAS ’06), Hakodate, Japan*, pp. 789–791. ACM. [TM]
- Makino, T., Hirayama, K. & Aihara, K. (2005) Understanding others: Possible links among parity, mirror neurons, and communication. *Behavioral and Brain Sciences* 28(2). Available at: <http://www.bbsonline.org/Preprints/Arbib-05012002/Supplemental/Makino.html>. (Online publication) [TM]
- Marken, R. (2002) *More mindreadings: Methods and models in the study of purpose*. New View. [aSH]
- McDowell, J. (1994) The content of perceptual experience. *Philosophical Quarterly* 44:190–205. [aSH]
- Meltzoff, A. (1988a) Infant imitation after a 1-week delay: Long-term memory for novel acts and multiple stimuli. *Developmental Psychology* 24:470–76. [aSH]
- (1988b) Infant imitation and memory: Nine-month-olds in immediate and deferred tests. *Child Development* 59:217–25. [MJH]
- (1990) Foundations for developing a concept of self: The role of imitation in relating self to other and the value of social mirroring, social modeling, and self practice in infancy. In: *The self in transition: Infancy to childhood*, ed. D. Cicchetti & M. Beeghly, pp. 139–64. University of Chicago Press. [aSH]
- (1995) Understanding of the intentions of others: Re-enactment of intended acts by 18-month-old children. *Developmental Psychology* 31:838–50. [aSH]
- (1996) The human infant as imitative generalist: A 20-year progress report on infant imitation with implications for comparative psychology. In: *Social learning in animals: The roots of culture*, ed. C. Heyes & B. Galef, Jr., pp. 347–70. Academic Press. [aSH]
- (2002a) Elements of a developmental theory of imitation. In: *The imitative mind*, ed. A. Meltzoff & W. Prinz, pp. 19–41. Cambridge University Press. [aSH]
- (2002b) Imitation as a mechanism of social cognition: Origins of empathy, theory of mind, and the representation of action. In: *Handbook of childhood cognitive development*, ed. U. Goswami, pp. 6–25. Blackwell. [aSH, rJK]
- (2005) Imitation and other minds: The “like me” hypothesis. In: *Perspectives on imitation: From neuroscience to social science, vol. 2*, ed. S. Hurley & N. Chater, pp. 55–77. MIT Press. [aSH]
- Meltzoff, A. & Moore, M. (1977) Imitation of facial and manual gestures by human neonates. *Science* 198:75–78. [aSH]
- (1983a) Imitation of facial and manual gestures by human neonates. *Science* 198:75–78. [aSH]
- (1983b) Newborn infants imitate adult facial gestures. *Child Development* 54(3):702–709. [BC, aSH]
- (1989) Imitation in newborn infants: Exploring the range of gestures imitated and the underlying mechanisms. *Developmental Psychology* 25(6):954–62. [BC, aSH]
- (1997) Explaining facial imitation: A theoretical model. *Early Development and Parenting* 6:179–92. [aSH, rJK, LMO]
- (1999) Persons and representations: Why infant imitation is important for theories of human development. In: *Imitation in infancy*, ed. J. Nadel & G. Butterworth, pp. 9–35. Cambridge Studies in Cognitive and Perceptual Development. Cambridge University Press. [aSH]
- (2000) Resolving the debate about early imitation. In: *Infant development: The essential readings*, ed. D. Muir, pp. 167–81. Blackwell. [aSH]
- Miall, R. C. (2003) Connecting mirror neurons and forward models. *Neuroreport* 14(16):1–3. [LF, aSH]
- Milgram, S. (1963) Behavioral study of obedience. *Journal of Abnormal and Social Psychology* 67:371–78. [aSH]
- Miller, G., Galanter, E. & Pribram, K. H. (1960) *Plans and the structure of behaviour*. Holt, Rinehart & Winston. [FP]
- Millikan, R. (1991) Perceptual content and Fregean myth. *Mind* 100(4):439–59. [aSH]
- (1993) Content and vehicle. In: *Spatial representation*, ed. N. Eilan, R. McCarthy & B. Brewer, pp. 256–68. Blackwell. [aSH]
- (2005) Some reflections on the simulation theory–theory debate. In: *Perspectives on imitation: From neuroscience to social science, vol. 2*, ed. S. Hurley & N. Chater, pp. 182–88. MIT Press. [aSH]
- (2006) Styles of rationality. In: *Rational animals?* ed. S. Hurley & M. Nudds, pp. 117–26. Oxford University Press. [aSH, FP]
- Milner, A. D. & Goodale, M. (1995) *The visual brain in action*. Oxford University Press. [aSH]
- Mithen, S. (1999) Imitation and cultural change: A view from the Stone Age, with specific reference to the manufacture of handaxes. In: *Mammalian social learning: Comparative and ecological perspectives*, ed. H. O. Box & K. R. Gibson, pp. 389–400. Cambridge University Press. [AW]
- Mukamel, R., Ekstrom, E., Kaplan, J. T., Iacoboni, M. & Fried, I. (2007) Mirror properties of single cells in human medial frontal cortex. Program No. 127.4 2007 Abstract Viewer/Itinerary Planner. Society for Neuroscience, San Diego, CA, 2007. Online publication. [MI]
- Murphy, S. T. & Zajonc, R. B. (1993) Affect, cognition, and awareness: Affective priming with optimal and suboptimal stimulus exposures. *Journal of Personality and Social Psychology* 64(5):723–39. [SDP]
- Myova-Yamakoshi, M., Tomonaga, M., Tanaka, M. & Matsuzawa, T. (2004) Imitation in neonatal chimpanzees (*Pan troglodytes*). *Developmental Science* 7:437–42. [MRL]
- Nagell, K., Olguin, R. & Tomasello, M. (1993) Processes of social learning in the tool use of chimpanzees (*Pan troglodytes*) and human children (*Homo sapiens*). *Journal of Comparative Psychology* 107:174–86. [aSH]
- Nakahara, K. & Miyashita, Y. (2005) Understanding intentions: Through the looking glass. *Science* 308:644–45. [aSH]
- Neda, Z., Ravasz, E., Brechet, Y., Vicsek, T. & Barabasi, A. L. (2000) The sound of many hands clapping: Tumultuous applause can transform itself into waves of synchronized clapping. *Nature* 403:849–50. [GRS]
- Nehaniv, C. & Dautenhahn, K. (2002) The correspondence problem. In: *Imitation in animals and artifacts*, ed. K. Dautenhahn & C. Nehaniv, pp. 42–61. MIT Press. [aSH]
- Neumann, R. & Strack, F. (2000) “Mood contagion”: The automatic transfer of mood between persons. *Journal of Personality and Social Psychology* 79:211–23. [GRS]

- Newman-Norlund, R. D., van Schie, H. T., van Zijl, A. M. J. & Bekkering, H. (2007) The mirror neuron system is more active during complementary compared with imitative action. *Nature Neuroscience* 10(7):817–18. [IVR]
- Nielsen, M. (2006) Copying actions and copying outcomes: Social learning through the second year. *Developmental Psychology* 42:555–65. [MN]
- Nielsen, M., Simcock, G. & Jenkins, L. (in press) The effect of social engagement on 24-month-olds' imitation from live and televised models. *Developmental Science*. [MN]
- Noë, A. (2004) *Action in perception*. MIT Press. [aSH, GN]
- Northoff, G. (2004) *Philosophy of the brain: The brain problem*. John Benjamins. [GN]
- Northoff, G., Heinzel, A., de Greck, M., Bermpohl, F., Dobrowolny, H. & Panksepp, J. (2006) Self-referential processing in our brain: A meta-analysis of imaging studies on the self. *NeuroImage* 31:440–57. [GN]
- Northoff, G., Schneider, F., Rotte, M., Matthiae, C., Tempelmann, C., Wiebking, C., Bermpohl, F., Heinzel, A., Danos, P., Heinze, H. J., Bogerts, B., Walter, M. & Panksepp, J. (in press) Differential parametric modulation of self-relatedness and emotions in different brain regions. *Human Brain Mapping*. [GN]
- O'Regan, J. K. & Noë, A. (2001a) A sensorimotor account of vision and visual consciousness. *Behavioral and Brain Sciences* 24:883–917. [aSH, rJK]
- (2001b) Acting out our sensory experience. *Behavioral and Brain Sciences* 24:955–75. [aSH]
- (2001c) What it is like to see: A sensorimotor theory of perceptual experience. *Synthese* 129:79–103. [aSH]
- Ortony, A., Clore, G. L. & Collins, A. (1988) *The cognitive structure of emotions*. Cambridge University Press. [FP]
- Oztop, E., Wolpert, D. & Kawato, M. (2005) Mental state inference using visual control parameters. *Cognitive Brain Research* 29:129–51. [aSH]
- Pardo, J. S. (2006) On phonetic convergence during conversational interaction. *Journal of the Acoustical Society of America* 119:2382–93. [SG]
- Pascual-Leone, A. (2001) The brain that plays music and is changed by it. *Annals of the New York Academy of Sciences* 930:315–29. [aSH]
- Pepperberg, I. (1999) *The Alex studies: Cognitive and communicative studies on grey parrots*. Harvard University Press. [aSH]
- (2002) Allospecific referential speech acquisition in Grey parrots (*Psittacus erithacus*): Evidence for multiple levels of avian vocal imitation. In: *Imitation in animals and artifacts*, ed. K. Dautenhahn & C. Nehaniv, pp. 109–31. MIT Press. [aSH]
- (2005) Insights into vocal imitation in Grey parrots (*Psittacus erithacus*). In: *Perspectives on imitation: From neuroscience to social science, vol. 1*, ed. S. Hurley & N. Chater, pp. 243–62. MIT Press. [aSH]
- Perner, J. (1991) *Understanding the representational mind*. Bradford Books/MIT Press. [rJK, AW]
- Peterson, G. & Trapold, M. (1982) Expectancy mediation of concurrent conditional discriminations. *American Journal of Psychology* 95:571–80. [aSH]
- Pezzulo, G. & Castelfranchi, C. (2007) The symbol detachment problem. *Cognitive Processing* 8:115–31. [FP]
- Phillips, W., Baron-Cohen, S. & Rutter, M. (1995) To what extent do children with autism understand desires? *Development and Psychopathology* 7:151–70. [BC]
- Pickering, M. J. & Garrod, S. (2004) Toward a mechanistic psychology of dialogue. *Behavioral and Brain Sciences* 27:169–225. [SG]
- (2007) Do people use language production to make predictions during comprehension? *Trends in Cognitive Sciences* 11:105–10. [SG]
- Povinelli, D. (1996) Chimpanzee theory of mind? In: *Theories of theories of mind*, ed. P. Carruthers & P. Smith, pp. 293–329. Cambridge University Press. [aSH]
- (2000) *Folk physics for apes*. Oxford University Press. [aSH]
- Povinelli, D. & Vonk, J. (2006) We don't need a microscope to explore the chimpanzee's mind. In: *Rational animals?* ed. S. Hurley & M. NuDDS. Oxford University Press. [aSH]
- Powers, W. T. (1973) *Behavior: The control of perception*. Aldine. [aSH]
- Press, C., Bird, G., Flach, R. & Heyes, C. (2005) Robotic movement elicits automatic imitation. *Brain research: Cognitive Brain Research* 25:632–40. [JHGW]
- Preston, S. D., Bechara, A., Damasio, H., Grabowski, T. J., Stansfield, R. B., Mehta, S. & Damasio, A. R. (2007) The neural substrates of cognitive empathy. *Social Neuroscience* 2(3–4):254–75. [SDP]
- Preston, S. D. & de Waal, F. B. M. (2002) Empathy: Its ultimate and proximate bases. *Behavioral and Brain Sciences* 25(1):1–72. [aSH, SDP]
- Preston, S. D., Hall, J., Berman, M. & Damasio, H. (in preparation a) Empathy from another perspective: Comparing empathy for hospital patients by subjects with and without depression. [SDP]
- Preston, S. D., Polk, T. A., Grabowski, T. J., Magnotta, V. A., Stansfield, R. B., Damasio, A. R. & Damasio, H. (in preparation b) The neural correlates of empathy and helping: Responses to patient accounts of serious illness. [SDP]
- Preston, S. D. & Stansfield, R. B. (in press) The Emostroop effect: Task-irrelevant facial emotions are processed spontaneously, rapidly and at the level of the specific emotion. *Cognitive, Affective, and Behavioral Neuroscience*. [rJK, SDP]
- Prinz, J. (2005) Imitation and moral development. In: *Perspectives on imitation: From neuroscience to social science, vol. 2*, ed. S. Hurley & N. Chater, pp. 267–82. MIT Press. [BC, aSH]
- Prinz, W. (1984) Modes of linkage between perception and action. In: *Cognition and motor processes*, ed. W. Prinz & A. F. Sanders, pp. 185–93. Springer. [aSH]
- (1987) Ideomotor action. In: *Perspectives on perception and action*, ed. H. Heuer & A. Sanders, pp. 47–76. Erlbaum. [aSH]
- (1990) A common-coding approach to perception and action. In: *Relationships between perception and action: Current approaches*, ed. O. Neumann & W. Prinz, pp. 167–201. Springer. [aSH]
- (2002) Experimental approaches to imitation. In: *The imitative mind*, ed. A. Meltzoff & W. Prinz, pp. 143–62. Cambridge University Press. [aSH]
- (2005) An ideomotor approach to imitation. In: *Perspectives on imitation: From neuroscience to social science, vol. 1*, ed. S. Hurley & N. Chater, pp. 141–56. MIT Press. [aSH]
- Pylyshyn, Z. W., ed. (1987) *The robot's dilemma: The frame problem in artificial intelligence*. Ablex. [IVR]
- Rakoczy, H., Warneken, F. & Tomasello, M. (2007) “This way!”, “No, that way!” – 3-year olds know that two people can have mutually incompatible desires. *Cognitive Development* 22:47–68. [aSH]
- Ramachandran, V. S. & Oberman, L. M. (2007) *Broken mirrors: A theory of autism*. Scientific American, June 2007, special edition. [LMO]
- Rasch, R. A. (1988) Timing and synchronization in ensemble performance. In: *Generative processes in music: The psychology of performance, improvisation, and composition*, ed. J. A. Sloboda, pp. 70–90. Clarendon Press. [MJH]
- Richell, M., Newman, L., Baron-Cohen, S., Wheelwright, S. & Blair (2003) Theory of mind and psychopathy: Can psychopathic individuals read the “language of the eyes”? *Neuropsychologia* 41:523–26. [BC]
- Rietveld, E. (in preparation) Affordance selection and monitoring. Paper presented at Philosophy of Psychology, Neuroscience and Biology Graduate Conference, All-Souls, Oxford, April 2005. [aSH]
- Rizzolatti, G. (2005) The mirror neuron system and imitation. In: *Perspectives on imitation: From neuroscience to social science, vol. 1*, ed. S. Hurley & N. Chater, pp. 55–76. MIT Press. [aSH]
- Rizzolatti, G. & Arbib, M. (1998) Language within our grasp. *Trends in Neuroscience* 21:188–94. [aSH]
- (1999) From grasping to speech: Imitation might provide a missing link [Reply]. *Trends in Neuroscience* 22:152. [aSH]
- Rizzolatti, G., Camarda, R., Fogassi, M., Gentilucci, M., Luppino, G. & Matelli, M. (1988) Functional organization of inferior area 6 in the macaque monkey: II. Area F5 and the control of distal movements. *Experimental Brain Research* 71:491–507. [aSH]
- Rizzolatti, G. & Craighero, L. (2004) The mirror-neuron system. *Annual Review of Neuroscience* 27:169–92. [MI]
- Rizzolatti, G., Fadiga, L., Fogassi, L. & Gallese, V. (2002) From mirror neurons to imitation: Facts and speculations. In: *The imitative mind: Development, evolution and brain bases*, ed. A. Meltzoff & W. Prinz, pp. 247–66. Cambridge University Press. [aSH, MRL]
- Rizzolatti, G., Fadiga, L., Gallese, V. & Fogassi, L. (1995) Premotor cortex and the recognition of motor actions. *Cognitive Brain Research* 3:131–41. [aSH]
- Rizzolatti, G., Fadiga, L., Matelli, M., Bettinardi, V., Paulesu, E., Perani, D. & Fazio, F. (1996) Localization of grasp representation in humans by PET: Observation versus execution. *Experimental Brain Research* 111:246–52. [aSH]
- Rizzolatti, G., Fogassi, L. & Gallese, V. (2001) Neurophysiological mechanisms underlying the understanding and imitation of action. *Nature Reviews: Neuroscience* 2:661–70. [R-PB]
- Rizzolatti, G. & Luppino, G. (2001) The cortical motor system. *Neuron* 31(6):889–901. [MI]
- Rosenblueth, A., Wiener, N. & Bigelow, J. (1968) Behaviour, purpose, and teleology. In: *Modern systems research for the behavioural scientist*, ed. W. Buckley, pp. 368–72. Aldine. [FP]
- Ross, H. & Lollis, S. (1987) Communication within infant social games. *Developmental Psychology* 23(2):241–48. [BC]
- Ruby, P. & Decety, J. (2001) Effect of subjective perspective taking during simulation of actions: A PET investigation of agency. *Nature Neuroscience* 4:546–50. [aSH]
- Sachs, H. A., Schegloff, E. A. & Jefferson, G. (1974) A simplest systematics for the organization of turn-taking for conversation. *Language* 50:696–735. [GRS]
- Sato, A. & Yasuda, A. (2005) Illusion of self-agency: Discrepancy between the predicted and actual sensory consequences of actions modulates the sense of self-agency, but not the sense of self-ownership. *Cognition* 94:241–55. [MJH]
- Scaife, M. & Bruner, J. (1975) The capacity for joint visual attention in the infant. *Nature* 253:265–66. [BC]
- Schmidt, R. C., Carello, C. & Turvey, M. T. (1990) Phase transitions and critical fluctuations in the visual coordination of rhythmic movements between people. *Journal of Experimental Psychology: Human Perception and Performance* 16:227–47. [MJH]

- Schmitt, A. & Grammer, K. (1997) Social intelligence and success: Don't be too clever in order to be smart. In: *Machiavellian intelligence II: Extensions and evaluations*, ed. A. Whiten & R. Byrne, pp. 86–111. Cambridge University Press. [aSH]
- Sebanz, N., Bekkering, H. & Knoblich, G. (2006) Joint action: Bodies and minds moving together. *Trends in Cognitive Sciences* 10:70–76. [MJH]
- Semin, G. R. & Cacioppo, J. T. (in press a) Grounding social cognition: Synchronization, entrainment, and coordination. In: *Embodied grounding: Social, cognitive, affective, and neuroscientific approaches*, ed. G. R. Semin & E. R. Smith. Cambridge University Press. [GRS]
- (in press b) From embodied representation to co-regulation. In: *Mirror neuron systems: The role of mirroring processes in social cognition*, ed. J. A. Pineda. Humana. [GRS]
- Shergill, S., Bullmore, E., Simmons, A., Murray, R. & McGuire, P. (2000) Functional anatomy of auditory verbal imagery in schizophrenic patients with auditory hallucinations. *American Journal of Psychiatry* 157:1691–93. [LF]
- Singer, T., Seymour, B., O'Doherty, J., Kaube, H., Dolan, R. & Frith, C. (2004) Empathy for pain involves the affective but not sensory components of pain. *Science* 303:1157–62. [AIG, SDP]
- Singer, T., Seymour, B., O'Doherty, J., Klaas, S., Dolan, R. & Frith, C. (2006) Empathic neural responses are modulated by the perceived fairness of others. *Nature* 439:466–69. [AIG, SDP]
- Smith, E. R. & Semin, G. R. (2004) Socially situated cognition: Cognition in its social context. *Advances in Experimental Social Psychology* 36:53–117. [GRS]
- Sommerville, J. A., Woodward, A. L. & Needham, A. (2005) Action experience alters 3-month-old infants' perception of others' actions. *Cognition* 96:B1–B11. [MRL]
- Spence, S., Brooks, D., Hirsch, S., Liddle, P., Meehan, J. & Grasby, P. (1997) A PET study of voluntary movement in schizophrenic patients experiencing passivity phenomena (delusions of alien control). *Brain* 120:1997–2011. [LF]
- Stamenov, M. & Gallese, V., eds. (2002) *Mirror neurons and the evolution of brain and language*. John Benjamins. [aSH]
- Stengel, E. (1947) A clinical and psychological study of echo-reactions. *Journal of Mental Science* 93:598–612. [MRL]
- Sterelny, K. (2003) *Thought in a hostile world*. Blackwell. [aSH]
- Suddendorf, T. & Whiten, A. (2001) Mental evolution and development: Evidence for secondary representation in children, great apes and other animals. *Psychological Bulletin* 127:629–50. [AW]
- Tennie, C., Call, J. & Tomasello, M. (2006) Push or pull: Imitation versus emulation in human children and great apes. *Ethology* 112:1159–69. [MN]
- Thagard, P. (2000) *Coherence in thought and action*. MIT Press. [IvR]
- Thorndike, E. (1898) Animal intelligence: An experimental study of the associative process in animals. *Psychological Review Monograph Supplement* 2(4):551–53. [aSH]
- Tognoli, E., Lagarde, J., DeGuzman, G. C. & Kelso, J. A. S. (2007) The phi complex as a neuromarker of human social coordination. *Proceedings of the National Academy of Science USA* 104:8190–95. [MJH]
- Tomasello, M. (1996) Do apes ape? In: *Social learning in animals: The roots of culture*, ed. C. Heyes & B. Galef, Jr., pp. 319–46. Academic Press. [aSH]
- (1998) Emulation learning and cultural learning. *Behavioral and Brain Sciences* 21:703–704. [aSH]
- (1999) *The cultural origins of human cognition*. Harvard University Press. [aSH, FP]
- Tomasello, M. & Call, J. (1997) *Primate cognition*. Oxford University Press. [aSH]
- (2006) Do chimpanzees know what others see – or only what they are looking at? In: *Rational animals?* ed. S. Hurley & M. Nudds. Oxford University Press. [aSH]
- Tomasello, M. & Carpenter, M. (2005) Intention reading and imitative learning. In: *Perspectives on imitation: From neuroscience to social science*, vol. 2, ed. S. Hurley & N. Chater, pp. 133–48. MIT Press. [aSH]
- Tomasello, M., Carpenter, M., Call, J., Behne, T. & Moll, H. (2005) Understanding and sharing intentions: The origins of cultural cognition. *Behavioral and Brain Sciences* 28(5):675–91. [BC]
- Tomasello, M., Kruger, A. & Ratner, H. (1993) Cultural learning. *Behavioral and Brain Sciences* 16:495–552. [aSH]
- Uddin, L. Q., Iacoboni, M., Lange, C. & Keenan, J. P. (2007) The self and social cognition: The role of cortical midline structures and mirror neurons. *Trends in Cognitive Sciences* 11(4):153–57. [MI]
- Umla, M., Kohler, E., Gallese, V., Fogassi, L., Fadiga, L., Keysers, C. & Rizzolatti, G. (2001) I know what you are doing: A neurophysiological study. *Neuron* 31:155–65. [aSH]
- Uzgris, I. (1981) Two functions of imitation during infancy. *International Journal of Behavioral Development* 4:1–12. [MN]
- van Dijk, J., Kerkhofs, R., van Rooij, I. & Haselager, P. (in press) Can there be such a thing as embodied embedded cognitive neuroscience? *Theory and Psychology*. [IvR]
- van Rooij, I., Bongers, R. M. & Haselager, W. F. G. (2002) A non-representational approach to imagined action. *Cognitive Science* 26(3):345–75. [IvR]
- van Rooij, I. & Wareham, T. (in press) Parameterized complexity in cognitive modeling: Foundations, applications and opportunities. *Computer Journal*. [IvR]
- Voelkl, B. & Huber, L. (2000) True imitation in marmosets. *Animal Behaviour* 60:195–202. [aSH]
- Vogeley, K. & Fink, G. R. (2003) Neural correlates of the first-person perspective. *Trends in Cognitive Sciences* 7(1):38–42. [MI]
- Vygotsky, L. (1987) *Thinking and speech: The collected works of L. S. Vygotsky, vol. 1*. Plenum. (Original work published 1934.) [LF]
- Watkins, K., Strafella, A. P. & Paus, T. (2003) Seeing and hearing speech excites the motor system involved in speech production. *Neuropsychologia* 41:989–94. [SG]
- Weir, A., Chappell, J. & Kacelnik, A. (2002) Shaping of hooks in New Caledonian crows. *Science* 297:981. [aSH]
- Whiten, A. (1996) When does smart behaviour-reading become mindreading? In: *Theories of theories of mind*, ed. P. Carruthers & P. Smith, pp. 277–92. Cambridge University Press. [aSH]
- (1997) The Machiavellian mindreader. In: *Machiavellian intelligence II: Extensions and evaluations*, ed. A. Whiten & R. Byrne, pp. 144–73. Cambridge University Press. [aSH]
- (2002) Imitation of sequential and hierarchical structure in action: Experimental studies with children and chimpanzees. In: *Imitation in animals and artifacts*, ed. K. Dautenhahn & C. Nehaniv, pp. 191–209. MIT Press. [aSH]
- (2006) The dissection of imitation and its “cognitive kin” in comparative and developmental psychology. In: *Imitation and the development of the social mind: Lessons from typical development and autism*, ed. S. Rogers & J. H. G. Williams, pp. 227–50. Guilford Press. [AW]
- Whiten, A. & Byrne, R., eds. (1997) *Machiavellian intelligence II: Extensions and evaluations*. Cambridge University Press. [aSH]
- Whiten, A., Custance, D., Gomez, J., Teixidor, P. & Bard, K. (1996) Imitative learning of artificial fruit processing in children (*Homo sapiens*) and chimpanzees (*Pan troglodytes*). *Journal of Comparative Psychology* 110:3–14. [aSH]
- Whiten, A., Horner, V. & de Waal, F. B. M. (2005a) Conformity to cultural norms of tool use in chimpanzees. *Nature* 437:737–40. [AW]
- Whiten, A., Horner, V. & Marshall-Pescini, S. (2005b) Selective imitation in child and chimpanzee: A window on the construal of others' actions. In: *Perspectives on imitation: From neuroscience to social science*, vol. 1, ed. S. Hurley & N. Chater, pp. 263–83. MIT Press. [aSH]
- Whiten, A., Spiteri, A., Horner, V., Bonnie, K. E., Lambeth, S. P., Schapiro, S. J. & de Waal, F. B. M. (2007) Transmission of multiple traditions within and between chimpanzee groups. *Current Biology* 17:1038–43. [AW]
- 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* 40:655–64. [AIG]
- Williams, J., Whiten, A., Suddendorf, T. & Perrett, D. (2001) Imitation, mirror neurons and autism. *Neuroscience and Biobehavioral Reviews* 25:287–95. [aSH]
- Williams, J. H. G., Whiten, A. & Singh, T. (2004) A systematic review of action imitation in autistic spectrum disorder. *Journal of Autism and Developmental Disorders* 34:285–99. [JHGW]
- Williams, J. H. G., Whiten, A., Waiter, G. D., Pechey, S. & Perrett, D. I. (2007) Cortical and subcortical mechanisms at the core of imitation. *Social Neuroscience* 2:66–78. [JHGW]
- Williams, L. M., Phillips, M. L., Brammer, M. J., Skerrett, D., Lagopoulos, J., Rennie, C., Bahramali, H., Olivieri, G., David, A. S., Peduto, A. & Gordon, E. (2001) Arousal dissociates amygdala and hippocampal fear responses: Evidence from simultaneous fMRI and skin conductance recording. *NeuroImage* 14:1070–79. [rJK]
- Wolpert, D. (1997) Computational approaches to motor control. *Trends in Cognitive Sciences* 1:209–16. [aSH]
- Wolpert, D., Doya, K. & Kawato, M. (2003) A unifying computational framework for motor control and social interaction. *Philosophical Transactions of the Royal Society of London, Series B* 358:593–602. [aSH] Also in: *The neuroscience of social interaction*, ed. C. Frith & D. Wolpert, pp. 305–22. Oxford University Press. [AIG]
- Wolpert, D. & Kawato, M. (1998) Multiple paired forward and inverse models for motor control. *Neural Networks* 11:1317–29. [aSH]
- Wolpert, D. M., Ghahramani, Z. & Flanagan, J. R. (2001) Perspectives and problems in motor learning. *Trends in Cognitive Sciences* 5:487–94. [JHGW]
- Wolpert, D. M., Miall, R. C. & Kawato, M. (1998) Internal models in the cerebellum. *Trends in Cognitive Sciences* 2(9):338–47. [SDP]
- Woodworth, R. S. (1926) *Psychology: A study of mental life*, 6th edition. Methuen. [R-PB]
- Zentall, T. (2001) Imitation and other forms of social learning in animals: Evidence, function, and mechanisms. *Cybernetics and Systems* 32:53–96. [aSH]