

Active perception and perceiving action:  
The Shared Circuits Hypothesis<sup>1</sup>

Susan Hurley

*Abstract:*

*Recently research on imitation and its role in social cognition has been flourishing across various disciplines. After briefly reviewing these developments under the headings of behavior, subpersonal mechanisms, and functions of imitation, I advance the **shared circuits hypothesis**. This hypothesis about subpersonal functional architecture describes a unified framework for the mechanisms that enable control, imitation, and simulation, which makes explicit the relationships among them. It should have heuristic value in sharpening up questions and predictions at both higher, personal and lower, neural levels, while avoiding over-simple or a priori assumptions of isomorphism between the subpersonal and personal levels. A striking aspect of the shared circuits hypothesis is the way it connects a shared information space for action and perception with a shared information space for self and other, while at the same time illustrating how the distinctions between self and other, and between the imagined and the real, can be overlaid on these shared information spaces. In this model information about intentional agents arrives in the first person plural: without distinction or inference between self and other. The shared circuits hypothesis also illustrates a horizontally modular architecture: it avoids the common conception of perception and action as separate and peripheral to central cognition. Rather, it views perception and action as dynamically co-constituted and shows how cognitively significant resources, such as distinctions between self and other and between the imagined and the real, and information for action understanding and planning, might emerge from the information space that action and perception share.*

This article develops the implications of a view of perception as essentially active, which I and others have been developing<sup>2</sup>, for the perception of action.

Some substantive stage setting may be useful. Elsewhere (1998, 1991) I have identified a conception of the mind I dubbed ‘the classical sandwich’, which is widespread across both philosophy and the empirical sciences of the mind. The classical sandwich conception regards perception as input from the world to the mind, action as output from the mind to the world, and cognition as sandwiched in between. Central cognition, on this view, is where all the conceptually structured general purpose thinking happens: perceptual information is assessed in light of standing beliefs and goals, deliberative and inferential processing occurs, action plans are formulated and sent on for execution. I have argued that the mind isn’t *necessarily* structured in this vertically modular way, and that there is

---

<sup>1</sup> For comments on earlier drafts of this material and related discussions, I am grateful to: John Bargh, Jeremy Butterfield, Nancy Cartwright, John Cummins, Chris Frith, Vittorio Gallese, Tamar Gendler, Philip Gerrans, Robert Gordon, Jeffrey Gray, Rick Grush, Celia Heyes, Marco Iacoboni, Andrew Meltzoff, Alva Noë, Hanna Pickard, Joëlle Proust, Nicholas Rawlins, Simon Saunders, Nick Shea, Evan Thompson, and the members of various audiences on occasions when I have delivered this material. I am also grateful to Nick Chater for his ready agreement to my drawing on material I drafted for Hurley and Chater 2004b, and to Tom Stone of MIT Press.

<sup>2</sup> See for example: Hurley 1998, 2001; O’Regan and Noë 2001a, b, c; Hurley and Noë 2003.

growing evidence that it is not *actually* so structured.<sup>3</sup> Instead, perception and action can be understood as sharing information processing resources in specific domains and as dynamically co-constituted in interaction with the environment, rather than as separate buffers around all-purpose central cognition. Cognitive resources and structure can emerge, layer by layer, from the shared information spaces that enable perception and action. Such horizontally modular structure may be able to do significant parts (I don't claim *all*) of the work the classical sandwich conception assigned to central cognition. In this chapter I provide a detailed view of how this promise might be made good in the realm of the perception of action and associated social cognition.

Some methodological stage setting may also be helpful. My arguments are in the style of philosophy of science<sup>4</sup> rather than philosophy of mind, focussing on empirical research and subpersonal functional architecture rather than personal level experience. Subpersonal processing can provide information that enables personal level processes even though there is no simple isomorphism between the levels. This paper examines philosophical issues that emerge 'organically' from empirical work, rather than importing an independent philosophical agenda from traditional philosophy of mind and applying it to empirical work. This style of philosophizing aims to complement rather than to challenge or displace traditional philosophical approaches; it can usefully be labelled 'natural philosophy'.

In this article I do two things. First, I review recent developments in the cognitive and neuro- sciences concerning imitation.<sup>5</sup> Imitation is still often thought of as a cognitively undemanding form of behavior. But since Thorndike (1898) showed that many animals could learn through trial and error but could not imitate, scientists have come to regard imitative learning as more cognitively demanding than individual trial and error learning. Recent work across various sciences argues that imitation is a rare ability fundamentally linked to characteristically human forms of intelligence, in particular to language, culture, and the ability to understand other minds.<sup>6</sup> This burgeoning body of work has important implications for our understanding of ourselves, both individually and socially. Imitation is not just an important factor in human development, but also has pervasive influence throughout adulthood, in ways we are just starting to understand. I here review why imitation is currently a topic of such intense research interest, under three headings: behavior, subpersonal mechanisms, and functions.

Second, I suggest a hypothesis that draws some of these threads of work on imitation together: the *shared circuits hypothesis*.<sup>7</sup> It includes elements suggested by various researchers, contributes further elements, and unifies these in a distinctive way. It

---

<sup>3</sup> As well as my 1998, 1991 and the references cited there, see Brooks 1999 and Noë forthcoming.

<sup>4</sup> Think of my approach as analogous to philosophy of physics rather than to metaphysics. Thanks here to Nancy Cartwright for discussion.

<sup>5</sup> Drawing on Hurley 2004 and Hurley and Chater 2004b, with revisions.

<sup>6</sup> See for example: Arbib in press; Arbib and Rizzolatti 1997; Gallese 2001, 2004; Gallese and Goldman 1998; Gordon 1995b; Jacoboni 2004; Meltzoff 2004; Rizzolatti and Arbib 1998, 1999; Tomasello 1999; Stamenov & Gallese 2002; Whiten et al 2004; Williams et al 2001; etc.

<sup>7</sup> Drawing on Hurley 2004, with revisions.

describes a unified subpersonal functional architecture for control, imitation, and simulation, in five stages or layers. It is put forward as a heuristic hypothesis that raises further empirical and philosophical questions.

## Part I. An Overview of Imitation

### A. BEHAVIOR

Imitation may be presumed to require at least *copying* in a generic sense: the observer's perception of the model's behavior causes the observer's own similar behavior, in some way such that the similarity between the observed behavior and the observer's behavior plays a role, though not necessarily at a conscious level, in generating the observer's behavior. Even copying in this generic sense raises issues about the mechanisms in play; but I postpone discussion of these and focus first on how imitative behavior is identified, and its distribution across animals, children, and adults.

Imitation needs to be distinguished from other forms of social learning that may look superficially similar.<sup>8</sup> The most restrictive understanding of true imitation requires that a novel action be learned by observing another do it, and, in addition to novelty, requires an instrumental or means/ends structure: you copy the others' means of achieving her goal, not just her goal, or just her movements. Imitation in this sense should be contrasted with stimulus enhancement, goal emulation, and response priming. In *stimulus enhancement* another's action draws your attention to a stimulus that triggers an innate or previously learned response; you do not thereby learn a novel action by observing the other. In *emulation*<sup>9</sup>, by contrast, you observe another achieving a goal in a certain way, find that goal attractive and attempt to achieve it yourself by whatever means. Individual trial and error learning may then lead you to the other's means of achieving the goal. A further contrast is with mere *response priming*, as in flocking behavior or contagious yawning, in which bodily movements are copied, but not as a learned means to a goal.

Goal emulation and response priming can be thought of as providing the ends and means components, respectively, of full-fledged imitation. However, the distinction between ends and means is not absolute; ends and means can be more or less distal or proximal, which can make for misunderstandings in discussions of whether ends or means or both are copied and hence of whether imitation or emulation is present (see Voelkl and Huber 2000, 196, 201). A movement may be the proximal means to achieving a bodily posture, which could be regarded as the proximal end of the movement (see Graziano et al 2002, 354-355); but the posture may in turn be regarded as a means to achieving some effect on

---

<sup>8</sup> The concept of '*true imitation*' is contested, owing in part to the different theoretical aims and methodologies of those concerned with imitation. See and cf: Byrne 2004, Heyes 1996, 2001; Rizzolatti 2004. What matters for present purposes is not what deserves this label, but that relevant distinctions be recognized.

<sup>9</sup> 'Emulation' in the sense well established in social learning theory (see Call & Carpenter 2002; Tomasello 1999) should be distinguished from the quite independent sense of the same term as used by Rick Grush (1995 and forthcoming). I use 'emulation' here in the social learning theory sense, not Grush's sense.

an external object or member of the social group, which may be a more distal end. We can understand more complex forms of imitation in terms of a structured sequence and/or hierarchy of means/ends relationships, in which one acquires a goal, learns how to achieve it by achieving several subgoals, learns how to achieve the subgoals by certain means, and so on.

## 1. ANIMALS

Studies of social learning in animals often focus on distinguishing true imitation from other superficially similar behaviors. Stimulus enhancement, goal emulation and response priming are certainly found in nonhuman animals. Careful experiments are needed to obtain evidence of imitation in a more restricted sense. For this purpose, the *two-action* experimental paradigm has become the tool of choice.<sup>10</sup> Suppose two models illustrate two different means of obtaining the same attractive result; one group of animals observes one model, the other observes the other model. Will the observer animals tend differentially to copy the specific method they have seen demonstrated? If not – if the animals' choices of method does not reflect the specific method they have seen modelled, say because animals in both groups converge on one method despite the different methods observed--, they may be displaying mere goal emulation plus trial and error learning, or stimulus enhancement, rather than imitative learning.

The difference between copying the ends of action and copying the means of action is important for understanding the phylogeny of the capacities for imitation and for understanding observed actions by others. For example, consider the view that understanding of events as the purposive acts of other intentional agents precedes imitation phylogenetically (Rizzolatti 2004). This view faces an objection: many animals are able to replicate movements, but it would be implausible to attribute action understanding to them all. Recall that in response priming, observing a movement 'primes' the same movement by the animal, independently of any understanding of the goal of the movement (as in the flocking of birds). But in response to this objection it has been suggested that such low level copying of movements could be present without high level mirroring of goals, or vice versa (Rizzolatti 2004); moreover, while imitative learning requires both of these capacities, either can be present without imitation. Action understanding requires the capacity to mirror goals, which is found in monkeys who have not shown genuine imitative learning. Thus action understanding, along with response priming and goal emulation, can precede genuine imitative learning. The latter would require the interplay of copying of both the ends and the means of action that is found in human beings. The phylogenetically rare capacity for imitative learning is arguably linked to the flexible recombinant means/ends structure of intentional action: the ability to use a given movement for different ends and pursue a given end by a variety of means (Tomasello 1999).

---

<sup>10</sup> See for example Nagell et al 1993; Call and Tomasello 1994; see and cf. Voelkl and Huber 2000 for further refinements.

It has proved remarkably difficult to find evidence of true imitation in nonhuman animals.<sup>11</sup> While early fieldwork with chimps appeared to provide evidence of their imitative abilities, critics challenged this interpretation effectively, and many subsequent experimental studies reported a lack of chimp imitation. For a long while sceptics who regarded the capacity for imitation as exclusively human had the upper hand. For example, in 1993, Tomasello, Kruger, and Ratner found no convincing evidence of imitative learning in nonhuman animals, and proposed that the understanding of behavior as goal-directed or intentional distinguishes human social learning from social learning in other species. On this view, while human beings can either imitate observed means or choose other means to emulate observed goals, other animals do not distinguish means and goals in this way. Rather, animals can copy movements without understanding their relevance to goals, or can learn about the affordances of objects by observing action on them. In neither case, the claim was, do other animals learn about the intentional, means/end structure of the observed action.

But a new consensus is now emerging, as a result of painstaking work to show imitation in some great apes and monkeys<sup>12</sup>, dolphins (Herman 2002), and birds such as parrots<sup>13</sup> and others<sup>14</sup>. Continuities are being described along a spectrum from the capacities of other social animals to the characteristically human, interrelated capacities for imitation, mind-reading, and language.<sup>15</sup>

For example, innovative experiments have extended the two-action method by employing 'artificial fruits', which can be opened in various ways to obtain a treat. These reveal that chimps may imitate with respect to one aspect of a modelled task and emulate for another, while children tend to imitate both aspects, even when the method imitated is inefficient. These experiments suggest that chimps imitate selectively, according to their appraisal of the significance of a particular aspect of the modelled task for achieving their goal, and that children are more likely than chimps to imitate obviously futile movements.<sup>16</sup>

## 2. CHILDREN

Indeed, children are 'imitation machines', with strong conventional and conformist tendencies.<sup>17</sup> While children do not always imitate unselectively and there are cases where they emulate goals (Gergely et al 2002), children have a greater tendency than chimps to imitate rather than to emulate when the method imitated is transparently

---

<sup>11</sup> See and cf. Byrne 1995; Galef 1988, 1998, 2004; Heyes and Galef 1996; Tomasello and Call 1997; Voelkl and Huber 2000; Zentall 2001.

<sup>12</sup> See Whiten et al 2004; Voelkl and Huber 2000.

<sup>13</sup> See Pepperberg 1999, 2002, 2004.

<sup>14</sup> Hunt and Gray 2003; see and cf. Weir et al 2002; Akins and Zentall 1996, 1998; Akins et al 2002.

<sup>15</sup> Tomasello 1999; Arbib forthcoming.

<sup>16</sup> See Whiten 2002b, Whiten et al 1996; Whiten et al 2004; see also Nagell et al 1993; Call and Tomasello 1994; Galef 2004; cf. Heyes 1998; Tomasello and Carpenter, 2004; Harris and Want 2004; Gergely et al 2002.

<sup>17</sup> Meltzoff 2004; Tomasello 1999, 159.

inefficient (Tomasello 1999, 29-30). For example, after seeing a demonstrator use a rake inefficiently, prongs down, to pull in a treat, two-year old children do the same; they almost never turn the rake over and use it more efficiently, edge down, instead. By contrast, chimps given a parallel demonstration tend to turn the rake over (Nagell et al 1993).

One explanation of these strong imitative tendencies of children is that in young children the perception of behavior tends to be enacted automatically in imitative behavior (in a broad sense that includes copying of bodily movements), unless actively inhibited. Inhibition is a function of frontal areas of the brain, but babies and very young children do not yet have a well-developed frontal function or capacity to inhibit imitative tendencies (Kinsbourne 2004).

Imitative behavior appears to play important roles in human development (see in particular the section on mind-reading below). Various imitative and related behaviors are acquired at stages throughout human infancy and development (Meltzoff 2004). Infants from 6 weeks to 14 months recognize that they are being imitated<sup>18</sup>, but only older infants act in ways that apparently purposively test whether they are being imitated. Since only people can imitate systematically, an ability to recognize being imitated provides a means of recognizing that an entity is a person. By 14 months, infants imitate a modelled novel act after a week's delay: for example, they turn on a light by touching a touch-sensitive light panel with their foreheads instead of their hands, differentially copying the novel means modelled as well as the result.<sup>19</sup> They do not turn the light on in this odd way unless they have seen the model do it first. By 15 to 18 months, infants recognize the underlying goal of an unsuccessful act they see modelled, and produce it, using various means: after seeing an adult try but fail to pull a dumbbell apart in her hands, they succeed in pulling it apart using knees as well as hands. However, they do not pick up the goals of failed 'attempts' from similar movements by inanimate devices, which suggests that young children perceive or understand agents and non-agents quite differently.<sup>20</sup>

### 3. ADULTS

Adults with damage to certain frontal areas of the brain also imitate uninhibitedly.<sup>21</sup> Such 'imitation syndrome' patients persistently imitate gestures the experimenter makes, although they have not been instructed to do so, and even when these are socially unacceptable or odd, such as putting on eyeglasses when already wearing glasses. But the human tendency to imitate is not confined to the young and the brain-damaged. While normal adults are usually able to inhibit overt imitation selectively (and it is adaptive to do so), overt imitation can be regarded as just the disinhibited tip of the iceberg of continual covert, inhibited imitation. Such covert imitation may reflect a basic motivation

---

<sup>18</sup> With differential activation of right inferior parietal lobe; see Decety and Chaminade 2004.

<sup>19</sup> Meltzoff 1988, 2004; cf. Gergely et al 2002.

<sup>20</sup> See Meltzoff 1996, 1988, 2004; Meltzoff and Moore, 1977, 1999; etc.

<sup>21</sup> Lhermitte et al 1986; Lhermitte 1986.

of human beings, adults as well as children, to interact synchronously or entrain with one another, which is a mechanism of affiliation as well as of social perception and learning (Kinsbourne 2004).

Despite inhibition, the underlying tendency to imitate remains and can readily be revealed or released. A number of experiments show how action can be spontaneously induced or modulated by the perception of similar action. Similarity between stimulus and response affects responses. For example, normal adult subjects, instructed to point to their nose when they heard the word 'nose!' or to point to a lamp when they heard the word 'lamp', performed perfectly while watching the experimenter model the required performance, but made mistakes when they observed the experimenter doing something else: they tended to copy what they saw done rather than to follow the instruction heard. (See Prinz 1990; Eidelberg 1929). In imitative interference paradigms, both the initiation and selection of gestures are faster when participants are primed by perception of similar gestures or of their results or goals, even if such primes are logically irrelevant to their task. Induction paradigms examine when spontaneous movements are induced by actions you actually perceive (*perceptual induction*, or involuntary imitation) or by actions you would like to perceive (*intentional induction*, as when moviegoers or sports fans in their seats make gestures they'd like to see made). Both types of induction are found, and are modulated by various contextual factors. Interestingly, perceptual induction can in some cases depend on background beliefs about whether what is perceived is the result of agency or rather generated by a computer, or on whether one's action is understood as part of a coordinated collective action by more than one person. (See W. Prinz 2004). Thus, perception has effects on action that, even when they are automatic, can also have cognitive depth, in that they depend on the way participants understand what they are perceiving and doing.

In considering the way observed action or external visual or verbal representations of action can control action, it is useful to distinguish between the imitation of *specific observed behaviors*, which may be facilitated by shared representations of our own acts and observed acts, and what has been called *the chameleon effect*, where *complex patterns of behavior* are induced. Demonstrations of the chameleon effect show that we automatically tend to assimilate our behavior to our social environment: modelled or represented personality traits and stereotypes tend automatically to activate corresponding behavior in us.<sup>22</sup> This tendency arguably acts as 'social glue', leading people to coordinate actions, to interact more smoothly, and to like each other. (See Dijksterhuis 2004; Chartrand & Bargh 1999).

In an experiment involving a specific observed behavior, normal adult subjects who interact in an unrelated task with someone who rubs her foot rub their own feet significantly more; transferred to another partner who touches his face, subjects start to touch their faces instead. Other experiments involve complex patterns of behavior, traits, or stereotypes, which may be primed by representations rather than perceptions of behavior. For example, normal adult participants have been primed by exposure to stimuli associated with traits (such as hostility, rudeness, politeness) or with stereotypes

---

<sup>22</sup> Bargh (forthcoming) comments on the striking similarities between imitation syndrome patients and normal college students who display the chameleon effect.

(such as elderly persons, college professors, soccer hooligans). Hostility-primed participants are found to deliver more intense 'shocks' than control participants in subsequent, ostensibly unrelated experiments based on Milgram's (1963) classic shock experiments. Rudeness-primed participants spontaneously behave more rudely, and politeness-primed participants more politely, than control participants, in subsequent, ostensibly unrelated interactions with experimenters. Youthful participants subliminally primed with words associated with the elderly, such as 'grey', 'bingo', 'sentimental', subsequently walk more slowly, perform worse on memory tasks, and express more conservative attitudes than similar-aged control participants. Perhaps most strikingly of all, college professor-primed participants perform better and soccer hooligan-primed participants perform worse than control participants on a subsequent, ostensibly unrelated, general knowledge quiz. That is, some participants were asked before doing the multiple choice test to do some ostensibly unrelated exercises about college professors, while a control group was not. The participants primed by thinking about this stereotype, generally associated with intelligence, got significantly higher scores. In another session different participants were given an unrelated exercise about soccer hooligans, while a control group was not. The participants primed by thinking about this stereotype, generally associated with lack of intelligence, got significantly lower scores.<sup>23</sup> (See Dijksterhuis and van Knippenberg 1998; Dijksterhuis 2004)

Such priming results are very robust: they hold across a wide range of verbal and visual primes and induced behavior, across dozens of different stereotypes and general traits, and using a variety of different priming methods, when primes are presented subliminally as well as when participants are conscious of them. Whether the subjects are conscious of the primes or not, they are unaware of any influence or correlation between the primes and their behavior. These results show that exposure to traits and stereotypes elicits general patterns of behavior and attitudes and influences the ways in which behavior is done. These influences are rapid, automatic, and unconscious, apply both to ends and to means, and do not depend on the subjects' volition or on their having a relevant independent goal that would rationalize their primed behavior. Imitation in this broad sense is our default social behavior, which needs to be specifically inhibited or overridden. Just thinking about or perceiving a certain kind of action automatically increases, in ways of which participants are unaware, the likelihood of engaging in that general type of behavior oneself (Dijksterhuis 2004). It is difficult for subjects in these experiments to accept that these broad imitative influences apply to themselves: both because they are unconscious and automatic, so people are not aware of them, and because such external influences threaten their conception of themselves as being in conscious control of their own behavior (Bargh 1999). Nevertheless, these influences are often inhibited, for example, by goals that make conflicting demands: elderly-primed participants do not walk more slowly if they have an independent need to hurry.

## B. MECHANISMS

---

<sup>23</sup> Priming with specific exemplars (e.g. Albert Einstein) produces contrast or comparison effects, while priming with a generic category (e.g. scientists) produces assimilation effects.

I now turn to examining the subpersonal mechanisms that might enable and explain imitative behavior.

## 1. THE CORRESPONDENCE PROBLEM

The generic idea of copying perceived behavior appears to require the solution to a difficult *correspondence problem* (Nehaniv & Dautenhahn 2002): how is perceived action by another agent translated into similar performance by the observer? When I imitate your hand movements at least I can see my own hands, even though my visual perspective on the two actions is different. But when I imitate your facial gestures, I cannot see my own face. How is the mapping from perception to similar behavior achieved? What information and mechanisms are needed to solve this problem?

Evidence that newborns and infants under 1 month old imitate facial gestures, even though they cannot see their own faces, suggests that there are at least some innate, supramodal correspondences between action and perception of similar action.<sup>24</sup> However, this would leave room for acquisition of further correspondences as imitative abilities develop. And sceptics about newborn imitation may also be skeptical about innate correspondences (Heyes 2004).

It has been argued by Celia Heyes that imitation does not require innate correspondences between perception and similar action. Rather, these could be acquired, in the right environment, through general-purpose associative learning mechanisms whereby neurons that fire together, wire together. The needed sensorimotor associations could be acquired through both direct and indirect routes. Direct sensorimotor associations could be formed when someone watches her own hand gesture, for example. But this won't work when the agent cannot perceive her own actions, as in facial gestures. Here, the association could be mediated by a third item, such as a mirror, an action word, or a stimulus that evokes the same behavior in the actor and in other agents the actor observes. Moreover, adult imitation of infants is common, and can perform the associative function of a mirror. In effect, the associative mechanism that enables imitation can extend into the cultural environment. Novel acts can be learned by observing another agent perform an unfamiliar sequence of familiar act elements, where each perception of an act element already has a motor association, resulting in a new sequence of motor elements that become linked through repetition to give rise to a novel act. On this view, imitation can emerge from the interaction between organisms with general purpose associative learning mechanisms and certain cultural environments. (See Heyes 2002, 2004.)

## 2. COMMON CODING

---

<sup>24</sup> Meltzoff 1988 through 2004; Meltzoff and Moore 1977 through 2000; etc. Developmental psychologist Moshe Anisfeld (1979 through 2004, Anisfeld et al 2001) represents a minority, including also Celia Heyes and Susan Jones, who remain skeptical of evidence for very early and newborn imitation.

Evidence of infant imitation has been explained in terms of a shared information space for perception and action: proprioceptive feedback from the infant's own acts are compared and matched to an observed target act, where these are coded in common, supramodal terms (Meltzoff and Moore 1977). An innate common code could initially code for relations among bodily organs such as lips and tongue, and develop through experience of body babbling toward more dynamic, complex and abstract coding (Meltzoff and Moore 1997). But a common code might also be acquired, along the lines suggested by Heyes (see her 2004, which discusses how the innateness and common coding issues are related; see also Part II below).

Evidence of the normal adult tendency to imitate and the reaction time advantage of imitative tasks has also been explained in terms of common coding of perception and action (see W. Prinz 2004). This would enable and facilitate imitation, by avoiding the correspondence problem and the need for translation between unrelated input and output codes to solve it. The common coding of perception and action has been associated with what William James called *ideomotor theory*, according to which every representation of movement awakes in some degree the movement that it represents. In particular, a representation of observed movement by another agent tends inherently to produce a similar movement by the observer, and has priming effects even when movements do not break through overtly. The regular concurrence of action with perceived effects enables the prediction of the effects of an action<sup>25</sup> and the selection of action, given an intention to produce certain effects<sup>26</sup> (Greenwald 1970, 1972). As a result, representation of a regular effect of action, whether proximal or distal, acquires the power to evoke similar action, if not inhibited.

Support for these ideas can be found from various further sources. Neurophysiologists have shown that observing a particular action primes precisely the muscles that would be needed to perform the same action (Fadiga et al 1995, 2002). Watching an action sequence speeds up your own performance of the same sequence; merely imagining a skilled performance, in sport or music, improves your own performance: constitutes a kind of practicing, as many athletes and musicians know<sup>27</sup>. Similar ideas have been applied to the perception and experience of emotion. Simulation theorist Robert Gordon argues that it takes a special containing mechanism to keep the emotion-recognition process from reverting to emotional contagion; this mechanism is not fail-safe. If simulation theory is right, he holds, only a thin line separates one's own mental life from one's representation of the mental life of another; off-line representations of other people have an inherent tendency to go on-line.<sup>28</sup>

### 3. MIRROR NEURONS

Common coding theories characterize the subpersonal architecture that enables imitation in functional terms. But how might this architecture be realized neurally? What kind of

---

<sup>25</sup> As in a forward model or efference copy; see below.

<sup>26</sup> As in an inverse model; see below.

<sup>27</sup> Pascuale-Leone 2001, Jeannerod 1997, 117, 119-122.

<sup>28</sup> Gordon 1995b; see also Adolphs 2002.

neural architecture and processing would enable and explain imitative and related behaviors? The recent discovery of mirror neurons seems relevant here.

Certain neurons appear to constitute a direct link between perception and action in that their firing correlates with specific perceptions as well as specific actions. Some of these, *canonical neurons*, can be thought of as reflecting affordances (in Gibson's sense, 1986): they fire when a certain type of action is performed, but are also triggered when objects that afford such actions are perceived. Others, *mirror neurons*, fire when a certain type of action is performed, but also when another agent is observed doing the same type of action.<sup>29</sup> That is, mirror neurons are sensitive both to others' actions and to equivalent actions of one's own; they do not register the difference between one's own action and similar actions by others. They can be very specifically tuned. For example, certain cells fire when a monkey sees the experimenter bring food to her own mouth with her own hand *or* when the monkey brings food to his own mouth (even in the dark, so that the monkey cannot see his hand).

---

<sup>29</sup> For surveys, see Rizzolatti 2004; Iacoboni 2004; Stamenov & Gallese 2002.

The function of mirror neurons in relation to imitation is a matter of intense current interest. When mirror neurons were discovered, it was tempting to suggest that they enable imitation by avoiding the correspondence problem: if the same neurons code for perceived action and matching performance, it may seem that no neural translation is needed. But things are not quite that simple. Giacomo Rizzolatti, one of the discoverers of mirror neurons, holds that imitation requires both the ability to understand another agent's action and the ability to replicate it. On his view, recall, action understanding precedes imitation phylogenetically and is subserved by mirror systems, which are necessary but not sufficient for imitation. Rizzolatti suggests that the motor resonance set up by mirror neurons makes action observation meaningful by linking it to the observer's own potential actions. (See Rizzolatti 2004.)<sup>30</sup>

Mirror neurons were discovered in monkeys by single-cell recording; evidence for human mirror systems includes brain imaging work, as well as demonstrations that observing another agent act primes the muscles the observer would need to do the same thing. Rizzolatti describes mirror neurons in frontal areas of monkey brain as part of a neural circuit that also includes parietal and visual areas. A similar circuit in human beings, he suggests, constitutes a control system, in which an intended imitative movement is compared to an observed target movement, enabling imitative learning.<sup>31</sup> In monkeys, mirror neurons appear to code for the goals of performed or observed actions. For example, a monkey mirror neuron may fire then the monkey reaches for an apple or sees another reach for it, even if a screen has come down to hide the apple. However, if the monkey has first seen that there is nothing behind the screen, observation of the same reaching movement toward a spot behind the same screen will not activate that neuron (Rizzolatti 2004). By contrast, the human mirror system extends also to the specific movements that are the means to achieving goals. As we saw, this difference between

---

<sup>30</sup> Psychologist Paul Harris (in discussion, Royaumont conference, 2002) has suggested an experimental assessment of the extent to which mirror neurons subserve action understanding in monkeys. Monkey mirror neurons fire when the monkey reaches for an apple, or when she sees the experimenter reach for the apple. The same mirror neurons also fire when the monkey sees a screen come down in front of the apple, so that it is no longer visible, and then sees the experimenter's hand reach behind the screen to where the apple is hidden. But they do not fire when the monkey first sees that there is no apple, and then the screen comes down and the monkey sees the experimenter's hand reach behind the screen in the same way. The mirror neurons, that is, appear to code for the goal of the action. Harris suggests a variant that would address how insightfully the monkey attributes goals to others. Suppose the monkey and experimenter look at a nut, and see the screen come down in front of it. Then the experimenter leaves the room. The monkey is permitted to remove the nut. Now 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? If so, this would suggest that the monkey attributes the goal of reaching for the nut to the experimenter, who 'doesn't know' that the nut is no longer there. Or will it not fire, because the nut is not there? Does the mirror neuron, that is, code for the *intended goal* of the observed action, or merely its *result*? Note that even chimps fail nonverbal false belief tests; see and cf. Call & Tomasello 1999; Call et al 2000; cf. Hare et al 2000, 2001. For discussion see Hurley 2003.

<sup>31</sup> Rizzolatti 2004. Others postulate similar control systems, though they differ on details: e.g., Rizzolatti locates the comparator site in PF, while Marco Iacoboni (2004) locates it in STS.

mirroring the ends of action and mirroring the means of action is important for the argument 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 associated with the same movements in the observer), that could provide information about the goal-directness of the observed action, but would not necessarily provide information for imitative learning, that is, about how to achieve the goal by means of the observed movements.

Human brain imaging studies suggest a division of labor within the mirror system: that frontal areas of the mirror system code for the ends or goals of action, and parietal areas for movements and means. To enable imitation, both areas generate motor signals relating to a planned imitative action for comparison to the observed action; the motor plan is then adjusted until a match is obtained. This neural architecture for imitation has been compared to current ideas about functional architectures for instrumental motor control, which combine inverse and forward models. *Inverse models* estimate what motor plan is needed to reach a certain goal from a given state of affairs. They can be adjusted by comparison with real feedback from motor activity, but this is slow. It is often more efficient to use real feedback to train *forward models*, which take copies of motor plans as input and simulate or predict their consequences. Forward models can then be used with inverse models to control goal-directed behavior more efficiently. In particular, forward models can predict the consequences of a planned imitative action for comparison to the observed action, so that the motor plan can be adjusted until a match is obtained. (See Iacoboni 2004.)

In sum, mirror neurons are arguably necessary, though not sufficient, for full-fledged imitation. They were discovered in macaque monkeys, but while these monkeys can emulate, they have not been shown to be able to imitate in the strict sense.<sup>32</sup> In these monkeys, the mirror system appears to code for the ends rather than the means of action. In human beings, by contrast, the mirror system has instrumental structure: some parts of it code for the goals of actions, others for specific movements that are the means used to achieve goals. It has been suggested that the human mirror system can be used to imitate and not just to emulate because it codes for means as well as ends, unlike the macaque's. Various suggestions have been made about the way in which mirror neurons may contribute to subpersonal comparator control circuits.

### C. FUNCTIONS

There is also intense current interest in whether mirror neurons can shed light on the possible functions of imitation in relation to other distinctively human capacities, such as the capacities for language, and for identifying with others and understanding the mental states that motivate the actions of others. The greatest differences between chimp and human brains are precisely in the significant expansion of the areas around the Sylvian fissure that subserve imitation, language, and the understanding of action (Iacoboni 2004). This is indeed where mirror neurons are concentrated. The relationships among this trio

---

<sup>32</sup> But see Voelkl and Huber 2000 for evidence of imitation in marmosets.

of capacities--for language, mind-reading, and imitation--are of fundamental importance for understanding the transition of human infants to adult persons. Does the development of either language or mind-reading 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? The answers are controversial, and may of course differ for language and mind-reading. A further controversy, about whether mind-reading is best understood as theorizing about other minds or as simulating them, is also relevant here. How does the theory-simulation controversy bear on the relationships between imitation and mind-reading, or vice versa?

It is arguable that imitation has other important functions as well, in human life and culture. I shall survey various ideas about the possible functions of imitation under four headings: language, the ratchet effect and cultural evolution, cooperation, and mind-reading.<sup>33</sup> My treatment of the first three possible functions of imitation will be very brief; they are important context for the shared circuits hypothesis I go on to develop and help to suggest some of its broader relevance in understanding what is distinctive about human minds. The mind-reading function will be discussed at greater length, since it is more directly relevant to the shared circuits hypothesis.

## 1. LANGUAGE

Intriguingly, mirror neurons coding for the goals of action in human beings are concentrated in Broca's area<sup>34</sup>, one of the primary language areas of the brain. This area is among those activated in humans when imitative tasks are performed. Transient virtual 'lesions' to Broca's area created by transcranial magnetic stimulation interfere with imitative tasks.<sup>35</sup> Now a broadly nativist view of language could motivate a kind of protectiveness about Broca's area as the best candidate for an innate language module in the brain.<sup>36</sup> However, the discovery that Broca's area is occupied by the mirror system and has an essential role in imitation has underscored questions about how language acquisition might exploit imitative learning mechanisms rather than (or in addition to) expressing innate linguistic knowledge<sup>37</sup> and has generated new arguments about how language might depend on the capacity for imitation, either in evolutionary or developmental timeframes.<sup>38</sup>

---

<sup>33</sup> This discussion of possible functions of imitation is not intended to be exhaustive. For discussion of the function of imitation in moral development, see J. Prinz 2004.

<sup>34</sup> Or the homologue thereof in monkeys.

<sup>35</sup> Iacoboni 2004; see also Heiser et al 2003.

<sup>36</sup> A remark to this effect was made by Marco Iacoboni at the Royaumont conference (2002) on imitation; see his 2004.

<sup>37</sup> See also Heyes 2004 on how the mirror properties of neurons might be acquired.

<sup>38</sup> Arbib 2002 and in press; Arbib and Rizzolatti 1997; Iacoboni 2004; Rizzolatti and Arbib 1998, 1999; Stamenov & Gallese 2002. On language and imitation/social learning more generally, see: Baldwin 1995; Christiansen 1994, 2004; Christiansen et al 2002; Deacon 1997; Tomasello 1999 on establishing shared reference to an object through joint attention, established by gaze following and role-reversal imitation.

What are the key features of imitation and the human mirror system that language might build on or exploit? First, I suggest, flexible, articulated relations between means and ends in imitative learning could be an evolutionary precursor of the arbitrary relations between symbols and what they refer to. Second, as Iacoboni (2004) and Arbib (2002, and in press) argue in different ways, the mirror system provides a common code for the actions of self and other, and thus for language production and perception; by enabling intersubjective action understanding, the mirror system may be the basis for the intersubjective ‘parity’ or sharing of meaning that is essential to language. Third, as Arbib suggests, the flexible recombinant structure of ends and means in imitation may be a precursor of recombinant grammatical structure in language.<sup>39</sup> Fourth, as Richard Byrne suggests, the problem of finding recombinant units of action in apparently smooth streams of bodily movement has many parallels with the problem of finding linguistic units such as words in the apparently continuous acoustic stream of speech. Skilled action has a modular structure that facilitates flexible recombinant function. Byrne describes how patterns of organization of action might be learned imitatively (*program level imitation*), despite variation in implementational details, by means of a mirror mechanism combined with a behavior parsing mechanism that would recognize the boundaries of behavior modules. Behavior parsing capacities may be an important precursor to more sophisticated human abilities for high-level perception of an underlying structure of intentions and causes in the surface flux of experience. (See Byrne 2004.) It is tempting also to regard behavior parsing and the recombinant structure of program level imitation as precursors of syntactic parsing and the recombinant structure of language.

## 2. THE RATCHET EFFECT AND CULTURAL EVOLUTION

Why might evolution favor neural structures that enable or facilitate response priming, emulation, and imitation? Suppose variations in the behavioral traits of adults that are not genetically heritable slightly favor some members of a given generation over others, so that some reproduce successfully and others do not. Their offspring may benefit if they can acquire the behavioral traits of their successful parents through response priming, emulation, and imitation as well as through genetic inheritance. A young creature that tends to copy its parents will tend to pick up the nonheritable behaviors of creatures that have survived long enough to reproduce, and tend to form associations between such behaviors and the environmental conditions in which they are appropriate. Depending on how costly or error-prone individual learning is, imitation may contribute more to genetic fitness.

If full-fledged imitation, as opposed to response priming and goal emulation, requires mirror circuits for means and ends to be wired together in the right way, it may be a more difficult trick for evolution to pull off, and so rarer – as indeed it is. But wouldn’t this rare development actually be a maladaptive handicap? Recall that children seem to be at a disadvantage (at least in the short run) compared to chimps in two-action paradigms, because children have a greater tendency to imitate even inefficient models while chimps have a greater tendency to emulate and find a more efficient means to achieve an

---

<sup>39</sup> Cf. Iacoboni (2004) for a different view of how evolution leads from action recognition through imitation to language.

attractive goal.<sup>40</sup> Despite the short-term disadvantages of imitating with such determination, could it be adaptive in the long run?

*The ratchet effect* (Tomasello 1999) explains how it might be. Occasionally, gifted or lucky individuals may have rare insights or discover novel means of achieving goals, which would not be rediscovered readily by independent trial and error learning. Without imitation, they would be lost; imitation preserves and spreads these discoveries around, making them available to all as a platform for further developments. Once the capacity for imitation has evolved genetically, imitation provides, via the ratchet effect, a mechanism that drives cultural and technological transmission, accumulation and evolution. A similar point is made by meme theory. *Memes* are units of cultural evolution analogous in some respects to genes; both are replicators that evolve through a process of imperfect copying under selective pressure. Memes are understood to be whatever is copied by imitation, the transmission mechanism that makes memetic evolution possible, so imitation is fundamental to meme theory. While genetic adaptations may explain the emergence of the capacity for imitation in the first place, once imitation is on the scene, meme theory explains cultural evolution in terms of the comparative reproductive success of memes themselves rather than the comparative reproductive success of genes. Various accounts have been given of the ways in which culture and life co-evolve, and of how cultural evolution can drive genetic evolution, as well as vice versa.<sup>41</sup>

#### 4. COOPERATION

As well as being subject to automatic imitative influences, human beings often deliberately select a pattern of behavior to imitate because it is associated with certain traits and stereotypes, even if they do not themselves actually partake of the relevant traits or stereotypes. This can be benign and contribute to moral development (see J. Prinz 2004); perhaps I can become virtuous, as Aristotle suggested, by behaving like a virtuous person. But like automatic imitation, deliberate selective imitation does not always operate benignly.

The ability to turn imitation on and off selectively can be a Machiavellian social advantage<sup>42</sup>: for example, by imitating the behavioral signs used by a group of cooperators to identify members, you may be able to obtain the benefits of cooperation from others, but then inhibit your own cooperative behavior before it comes time to

---

<sup>40</sup> Nagell et al, 2003; Whiten et al 2004.

<sup>41</sup> See and cf. Baldwin (1896); Dawkins (1976/1989); Boyd and Richerson (1982, 1985), Blackmore (1999, 2000, 2001); Dennett (1995); Deacon (1997); Hurley and Chater (2004b), part 4. The reproductive success of memes can depend on feedback effects that operate via their information content, or on content-independent tendencies, such as a tendency to copy the most frequent memes, or those associated with the highest-status people; see Boyd & Richerson (1985); Henrich & Boyd (1998); Henrich & Gil-White (2001). Harris and Want (2004) argue it is capacity for selective imitation rather than for imitation per se that underwrites the ratchet effect and the explosive development of complex tool forms in the upper Paleolithic.

<sup>42</sup> On Machiavellian intelligence, see Byrne and Whiten 1988; Whiten and Byrne 1997.

reciprocate. That is, a group of cooperators may develop shared behaviors by means of which members identify one another as cooperators and exclude noncooperators from free riding. Noncooperators may then selectively imitate such behaviors in order to induce cooperative behavior from group members, and then fail to return cooperative behavior, thus deceptively obtaining the benefits of cooperation without paying the costs. So-called *greenbeard genes* could produce genetically determined analogues of such imitative free riding (see Dawkins 1982, 149). But the evolution of a general capacity for flexible selective imitation would make it possible to get the advantages of free riding without the need to evolve genes for specific behaviors.

How can the potential benefits of cooperation be achieved despite the ubiquitous threat of free riding? Certain solutions to collective action problems in effect require being able to recognize and identify with others' mental states. For example, John Howard's (1988) Mirror Strategy for one-off Prisoners' Dilemmas provides a simple if self-referential rule: cooperate with any others you encounter who act on this very same rule. This rule has been implemented computationally and shown to outperform defection even where the same players do not meet repeatedly (when players do play one another repeatedly, it is well known that various strategies, such as Tit-for-Tat, can outperform Defection).

Notice a striking property<sup>43</sup>: mirror strategists need to detect the way one another think—that is, their methods of choice or intentions or mental states or programs—, not just to observe one another's behavior. Which choices the cooperators should make are not determined until it is known whether relevant other agents are cooperators, so cooperation must be conditioned not on the choices of others but their methods of choice. Of course, it may be difficult to detect methods of choice with perfect reliability, especially given incentives to obtain the benefits of cooperation without paying the costs by deceptive mimicry of the signals by which cooperators identify one another. Such imitative free riding could be expected in turn to prompt more sophisticated and insightful methods of detecting the true methods of choice behind possibly deceptive behavioral signals. An arms race would result between insightful recognition of true methods of choice and deceptive mimicry of behavior patterns associated with methods of choice.

These observations suggest a hypothesis about a functional difference between 'mind-reading'<sup>44</sup> and mere smart behavior-reading in solving collective action problems. Psychologists ask: what is the functional difference between genuine mind-reading and smart behavior-reading (Whiten 1996)? Many of the social problems animals face can be solved merely in terms of behavior-circumstance correlations and corresponding

---

<sup>43</sup> Which also applies to Regan's (1980) proposal for solving coordination problems.

<sup>44</sup> I prefer the term 'mind-reading' to 'theory of mind', when used generically to indicate abilities to discern, understand and/or identify with the mental states of others. The two major substantive theories of how mind-reading is done are theory theory and simulation theory; 'theory of mind' used generically is liable to be confused with theory theory or to imply that theory theory occupies the whole field.

behavioral predictions, without the need to postulate mediating mental states.<sup>45</sup> And after all, it might be said, all we ever 'really observe' is behavior in environments; we infer mental states from this. However, mind-readers do not merely keep track of the behavior of other agents, but also understand other agents in terms of their mental states. Mind-readers can attribute intentions to others even when their acts do not carry out their intentions; they can attribute beliefs to others even when those beliefs are false. Mind-reading is something that human children only learn to do gradually; for example, children under four do not generally attribute false beliefs to others. Moreover, the capacity for mind-reading is characteristically human; evidence for mind-reading in nonhuman animals is scarce and controversial.<sup>46</sup> What problem solving pressures are addressed by going beyond the mere tracking of behavior-circumstance correlations to attribute mental states to explain observed behavior?

I suggest that this question is illuminated by being put in the context of the problem of how to obtain the benefits of cooperation, and in particular the way identifying others as cooperators may require recognizing and identifying with their mental states. As indicated, at least some solutions to problems of cooperation and coordination require more than merely tracking the behavior of others, in part because of the way behavior is subject to deceptive imitation. If cooperators need to know whether others have the mental processes of a cooperator before they can determine what cooperators will do, as in Howard's Mirror Strategy, they must rely on more than unmediated associations between circumstances and behavior. Such solutions require the understanding of others in terms of the mental states that generate their behavior, such as intentions to cooperate with other cooperators. Moreover, participants in such collective action would have not just to be mind-readers, but also to be able to identify, more or less reliably, other mind-readers. In order to counter the invasion of collective cooperative units by ever more sophisticated deceptive imitation that obtains the benefits of cooperation without paying the cost, recognition of other cooperators would be driven progressively further away from merely detecting surface behavior-circumstance correlations toward the detection of underlying mental variables, or mind-reading. Capacities for mimicry might in turn be expected to become more subtle and mentalistic. Mind-reading and intersubjective identification may result from such an arms race.<sup>47</sup>

## 5. MIND-READING

Human beings are distinctive among animals in both their strong tendency to imitate and their skill at understanding other minds. I've explained how, in the presence of imitation, mind-reading might emerge from an arms race between cooperation and deception. What more can be said about the possible functions of imitation in relation to mind-reading?

---

<sup>45</sup> See and cf. Heyes 1998; Povinelli 1996; forthcoming; Heyes and Dickinson 1993; Call and Tomasello 1999.

<sup>46</sup> See and cf. Davies and Stone 1995a, b; Carruthers and Smith; Heyes 1998; Hare et al 2000, 2001; Povinelli forthcoming; Tomasello and Call, forthcoming.

<sup>47</sup> See Hurley forthcoming for an extended version of the argument of this section; see also Hurley 2003b.

As already indicated, a hypothesis gaining support among scientists and philosophers is that mirror systems are part of the mechanism for understanding observed actions by intersubjective identification with others, for empathy and for simulation.<sup>48</sup> When you see someone do something, your own motor system is primed to imitate, even if imitation is inhibited or taken 'off-line': simulation can be regarded as off-line copying. This enables you to regard yourself and others as similar, to identify with others, and to understand the motivation of others' actions in a means/ends structured way. Extensions of the human mirror system provide a plausible neural basis for emotional understanding and empathy.<sup>49</sup> Within this broad intellectual development, however, it is useful to distinguish several different positions. I here compare those of neuroscientist Gallese, developmental psychologist Meltzoff, philosopher Gordon, and developmental psychologist/primatologist Tomasello.

Gallese's *shared manifold hypothesis* proposes that the mirror system has a general role in enabling empathy. The mirror system develops out of the way biological control systems model interactions between organisms and their environments. Mirror systems are the neural basis of a primitive intersubjective information space or shared manifold, which is prior to self-other distinctions both phylogenetically and ontogenetically but is preserved in human adults. This view softens the contrast between adult human mind-reading and mere responses to others' behavior as found in other social animals. The shared manifold underwrites automatic intersubjective identifications not just across different perceptual modalities and action, but also for sensations and emotions. There is evidence, for example, of mirror mechanisms for pain and disgust; and hearing an expression of anger increases the activation of muscles used to express anger. (See Gallese 2001, 2004)

Meltzoff (2004) draws on his work on early imitation to argue that early imitation and its enabling mechanisms beget the understanding of other agents, rather than the other way around. In Meltzoff's view, the ability to understand other minds has innate foundations, but develops in stages. Imitation plays a critical role in his arguments for a middle ground between Fodorian nativism and Piagetian theory. Infants have a primitive ability to recognize being imitated and to imitate, hence to recognize people as different from other things and to recognize equivalences between the acts of self and other. The initial bridge between self and other provides a basis for privileged access to people that we do not have to other things, which is developed in an early three-stage process. First, own acts are linked to others' similar acts supramodally, as evidenced by newborn imitation of others' facial acts. Second, own acts of certain kinds are linked bi-directionally to own mental states of certain kinds, through learning. Third, others' similar acts are linked to others' similar mental states. This process gets mind-reading started on understanding agency and the mental states most directly associated with it: desires, intentions, perceptions,

---

<sup>48</sup> See e.g. Gallese 2001, 2004; Gallese & Goldman 1998; Gordon 1995a,b; Goldman 2004, discussing autistic subjects' deficiency at imitation in early years; Williams et al 2001 on mirror neurons as basis for imitation and mind-reading and autism as a deficit in an ontogenetic cascade. Consider also the finding of Chartrand and Bargh 1999, experiment 3, that those high in dispositional empathy imitated more than those low in empathy. Thanks here to John Bargh.

<sup>49</sup> As argued in various ways by Gallese 2004 and Meltzoff 2004, Rizzolatti 2004, Iacoboni 2004, Decety and Chaminade 2004; see also Adolphs 2002.

emotions. Meltzoff emphasizes that the ability to understand other minds is not all or nothing.<sup>50</sup> Understanding mental states further from action, such as false beliefs, comes later, with further development.

This early three-stage process is not conceived as a matter of formal reasoning, but rather of processing the other as 'like me'. On the one hand, Meltzoff is often interpreted as viewing mind-reading in terms of theoretical inferences from first-person mind-behavior links to similar third-person links, in an updating of classical arguments from analogy.<sup>51</sup> There are indeed clear elements of first-to-third-person inference in his view of how mind-reading develops. As he expresses it, 'the crux of the idea is that infants may use their own intentional actions as a framework for interpreting the intentional actions of others' (2004, XXX). For example, 12-month-olds infants follow the 'gaze' of a model significantly less when the model's eyes are closed than when they are open, but do not similarly refrain from following the 'gaze' of blindfolded models until they are given first person experience with blindfolds. But on the other hand, the initial self-other linkage Meltzoff postulates, expressed in newborn imitation, is via a supramodal common code for observed and observer's acts, which is direct and noninferential (see Meltzoff and Moore 1997). On a graded view of mind-reading such as this, the role of theoretical inference from the first- to the third-person in mind-reading can enter at later stages and increase significantly with development.

Thus, while Meltzoff's approach to mind-reading is usually viewed as a version of the theory theory rather than the simulation theory of mind-reading, it has elements of both approaches. The '*theory theory*' approach to mind-reading regards commonsense psychology as a kind of proto-scientific theory, in which knowledge is represented in the form of laws about mental states and behavior; to the degree that these are not known innately, they are discovered by testing hypotheses against evidence. People's specific mental states and behaviors are inferred from other mental states and behaviors by means of such laws. No copying is involved. By contrast, *simulation theories* understand mind-reading to start with the mind-reader taking someone else's perspective and generating pretend mental or behavioral states that match the other person's. These are not made the object of theoretical inference, but rather are used as inputs to the simulator's own psychological processes, including decision-making processes, while these are held 'off-line', producing simulated mental states and behavior as output. The simulated outputs are then assigned to the other person; these may be predicted behaviors by the other, or mental states of the other that explain his observed behaviors. Such simulation is an extension of practical abilities rather than a theoretical exercise: the simulator copies the states of the other and uses the copies in her own decision-making equipment, instead of making inferences about the other based on laws.<sup>52</sup>

---

<sup>50</sup> The same point can be made for other animals; see Tomasello 1999 on levels of mind-reading ability.

<sup>51</sup> Goldman's important version of simulation theory, which space does not permit me to discuss here, shares this first-to-third-person aspect. See his 1989, 1992, 2004.

<sup>52</sup> Within this broad theory/simulation contrast, many finer distinctions have rightly been drawn, among various versions, levels of description and aims within each category. For some of these, and challenges to the distinction, see Davies and Stone 1995a,b; Carruthers and Smith 1996. See

Meltzoff's three-stage process can be restated in explicitly theory-theory terms, as follows. At stage one, the innate equivalence between my own acts and others' acts (exploited by early imitation and recognition of being imitated) makes it possible to recognize that some acts (by myself) are similar to other acts (by another). At stage two, first-person experience provides laws that link one's own acts and own mental states. At stage three, it is inferred that another's act that is similar to mine is lawfully linked to the other's mental states in the same way that my act is lawfully linked to my mental states. As Meltzoff points out (personal communication), there is no inference from the first-person to the third-person at stage one of this account; the initial bridge between self and other expressed in imitation and recognition of being imitated is bi-directional. However, an inference from the first-person to the third-person does enter as we proceed through stages two and three of this account: it resembles traditional arguments from analogy in inferring laws linking third-person acts and mental states from laws linking first-person acts and mental states.

By contrast, Gordon's radical version of simulation theory explicitly rejects the first-to-third-person direction of explanation in understanding other minds and offers a different view of the link between imitation and mind-reading.<sup>53</sup> In what he calls 'constitutive mirroring', a copied motor pattern is part of the very perception of the other person's action in the first place, although the motor pattern may be inhibited and thus not produce overt movement. Gordon finds constitutive mirroring in Gallese's account of the primitive intersubjective 'we'-space, the basis of empathy that implicitly expresses the *similarity* of self and other rather than their *distinctness*.<sup>54</sup> Gordon proposes that when constitutive mirroring imposes first-person phenomena, a process of analysis-by-synthesis occurs in which the other'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. For example, 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 part of the fundamental commonality of persons. Thus, as he puts it, I don't infer from the first to the third person, but rather multiply the first person (2004).

Gordon argues that the first-to-third-person inference in Meltzoff's account is problematic, 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 the first stage of Meltzoff's account, the similarity between acts of self and other is supposed to be established by their innate

---

also Millikan 2004, who distinguishes ontological, ontogenetic, and epistemological questions about thoughts of other minds, on which theory-theory and simulation theory may differ.

<sup>53</sup> See Gordon 1986, 1995a, b, 1996, 2002, 2004. Note that it is a mistake to associate simulation theories too closely with the first-to-third-person arguments from analogy (Gordon's view is a counterexample) or theory theories with rejection of this type of argument (Meltzoff's view is a counterexample). The theory vs. simulation distinction cuts across acceptance or rejection of the first-to-third-person direction of explanation.

<sup>54</sup> Gallese understands empathy to involve not the recognition of others as bodies endowed with minds, but rather the assumption of a common scheme of reasons by reference to which persons, self and others alike, are intelligible. Gallese, 2004; see and cf. Strawson 1959, 1966.

equivalence, which is exploited by early imitation; this stage may involve constitutive mirroring, as in Gallese's primitive shared manifold. But the second and third stages of Meltzoff's account, where the analogical inference occurs, requires that self and other also be distinguished: if this kind of act *of mine* is linked to *my* mental states of a certain kind, then a similar (as established in stage one) kind of act *by another person* is also linked to *her* mental states of a similar kind. But, Gordon objects, if I cannot distinguish *a* from *b*, I cannot make an analogical inference from *a* to *b*. While such an inference may sometimes be a feature of mature imitative mirroring, Gordon regards it as beyond the capacities of infants (2004).

However, a standard charge against pure simulation theories of mind-reading has been that they lack the resources to explain how mature mind-readers distinguish and identify different people and keep track of which actions and mental states are whose. Gordon suggests that multiple first persons are distinguished and tracked in the process of making them intelligible as persons, to avoid incoherence and disunity under the common scheme of reasons.<sup>55</sup> Mental states that do not make sense together are assigned to different persons. But can this be done in pure simulation mode, with no overlay of theory and inference? Simulation is supposed to be the off-line use of practical abilities, in contrast to theorizing about actions. But what exactly is the difference between interpreting an action to make sense of it and theorizing about it? When I use practical reason off-line in mind-reading, I don't formulate normative laws from which I make inferences; rather, I activate my own normative and deliberative dispositions. As Millikan might put it (see her 2004), my thought about another's action is not wholly separate from my entertaining that action.

A suggestion worth considering here is this: the fundamental *similarity between self and other* may best be understood not in terms of theorizing, but rather in terms of simulation (as in Gordon's constitutive mirroring, Gallese's shared manifold, the innate self-other equivalence exploited by early imitation, on Meltzoff's view, and my layer 3, below). Such primitive intersubjectivity may persist into adulthood and remain an essential aspect of mature empathy and mind-reading, as Gallese suggests. But as mind-reading develops it also employs a *self/other distinction*, as when the older child attributes to the other false beliefs, different from her own, or distinguishes imitating from being imitated.<sup>56</sup> More generally, mature mind-reading requires the ability to distinguish, identify, and track multiple other persons and to assign acts and mental states to them in a process of interpretation. The full range of distinctions and identifications mature mind-reading

---

<sup>55</sup> See also and cf. Hurley 1998, part 1; 1989.

<sup>56</sup> Decety and Chaminade (2004) invoke single-cell, imaging, and behavioral evidence in support of the shared neural coding of action and the perception of action in a mirror system. They regard such automatic motor resonance as a necessary basis for intersubjectivity in action understanding and emotional empathy. However, they focus not on the intersubjective space that precedes the self-other distinction, but rather on the characteristically human self-other distinction, and the way it is imposed on what is common to the representation of self and other. They report imaging experiments that probe the neural bases of the self-other distinction and reveal the relevance of left-right lateralization.

requires may indeed require theoretical and inferential resources, even while the simulative foundation remains essential.

Gallese, Meltzoff and Gordon stress the contribution of imitation to understanding other agents.<sup>57</sup> By contrast, Tomasello and Carpenter, like Rizzolatti, emphasize the contribution of action understanding to imitation and the ways in which imitative learning depends on intention reading. Tomasello and Carpenter (2004) argue that new paradigms developed with children have made a clearer distinction between imitative learning and other forms of social learning than the two-action method does. In these paradigms, the modelled action is unsuccessful or accidental. If the observer copies what was intended even though it was not shown, as opposed to only the observed movements or the observed though unintended result, that suggests the observer does understand the intentional structure of the observed action. Tomasello and Carpenter argue that, in the new demonstrations of imitative learning in which the modelled behavior is the same and only the modelled intention varies across conditions, the ability to read intentions is needed to explain what is copied. Given the results from the various imitation paradigms, they regard it as most parsimonious to assume that children use their understanding of intentions to imitate.

How then should we view the relationship between imitation and mind-reading? On some views, action understanding precedes imitation and full-fledged imitative learning, with intentional, means/ends structure, depends on intention reading.<sup>58</sup> On other views, imitation underwrites early mind-reading abilities.<sup>59</sup> Are these views in tension?

Not necessarily. In order to appreciate their potential compatibility, however, it is important to distinguish various stages or levels in both imitation and mind-reading and the ways these could built on one another dynamically, in evolutionary and developmental processes. Recall Rizzolatti's argument that action understanding precedes imitation in evolution, which distinguishes the mirroring of movements (in response priming), from the mirroring of goals (in emulation) and from genuine imitative learning with flexible intentional structure relating observed means to observed results. He suggests that the capacity to copy observed results via mirror systems may underwrite phylogenetically early understanding of action in terms of goals and intentions, which in turn is needed for phylogenetically later imitative learning with intentional structure, in which the mirroring of means and of ends are linked flexibly in the larger mirror circuit characteristic of human beings (2004). Recall also earlier suggestions that recognition of a fundamental self-other similarity via simulation (as in Gallese's primitive shared manifold, Meltzoff's innate self-other equivalence, Gordon's constitutive mirroring) may precede developmentally the registration of a self-other distinction, and more generally precede the inferential, interpretative abilities drawn on in more advanced mind-reading to identify and distinguish multiple persons and to keep track of which mental states go with which persons. Thus, very early imitation may express registration of a fundamental self-other similarity, while the distinctive human capacity for imitative learning with its

---

<sup>57</sup> As does Goldman (2004).

<sup>58</sup> Rizzolatti 2004; Tomasello & Carpenter 2004.

<sup>59</sup> Gallese 2004; Meltzoff 2004; Gordon 2004; Goldman 2004.

flexible means/ends structure in turn contributes to the development of the self-other distinction and of more advanced mind-reading skills.

## Part II. The Shared Circuits Hypothesis

I've now reached the second half of this essay, in which I want to draw together some of the threads of work on imitation surveyed above, by putting forward unified functional architecture for control, imitation, and simulation. Various researchers at the currently buzzing intersection of work on these processes have suggested that they are closely connected or even co-constituted.<sup>60</sup> There is something intuitively right and important here, yet the suggested relationships are often partial or expressed in one of several overlapping technical jargons that may be inaccessible to those in other disciplines who are interested in essentially the same issues. At this point it is worth exposing the substantive issues clearly in a way that cuts across disciplinary boundaries. I here put forward in plain terms a unified framework that makes the relationships among the mechanisms that enable control, imitation, and simulation explicit. Many of the phenomena discussed above can be located and understood by reference to this framework. I call it the *shared circuits hypothesis*. It can be regarded as a descendant of the common coding hypothesis about perception and action, though it describes commonality in terms of the functional dynamics rather than the coding of perception and action. It is also a relative of Gallese's shared manifold hypothesis, though it situates elements of Gallese's views explicitly within an overall framework.

The shared circuits hypothesis is a mid-level hypothesis about subpersonal functional architecture, cast at a level of description between those of neural implementation and of the personal level of conscious perception and intentional action.<sup>61</sup> While it may be too early to claim definitive empirical support for this particular specification, it may nevertheless have heuristic value in sharpening up questions at both higher and lower levels, while avoiding over-simple or *a priori* assumptions of isomorphism between subpersonal and personal level descriptions. Some of the work I've surveyed does not keep clear track of distinctions between neural, functional subpersonal, and personal levels of description. While the boundaries between these levels are not wholly opaque, it will conduce to clarity and progress to recognize distinctions between levels, and to frame issues about interlevel relations, more explicitly. Looking downward from the functional shared circuits architecture, we can ask whether there is evidence that particular neural circuits implement parts of it. Looking upward, we can ask what its behavioral and cognitive implications are. If information about self and other is processed subpersonally along the lines suggested by the shared circuits hypothesis, what implications might that have for the role and uses of such information at the personal level? For example, if

---

<sup>60</sup> See and cf. Frith et al., 2000; Gallese, 2000, 2004; Gerrans, forthcoming; Gordon 2002; Grush 1995, forthcoming; Iacoboni 2004; Jeannerod, 1997, 2001; Meltzoff 2004; Oztup and Arbib 2002; Arbib et al 2000; Proust forthcoming; Rizzolatti 2004; Wolpert et al 2003; Gallese & Goldman 1998; Blakemore & Decety 2001; etc.

<sup>61</sup> Read 'animal level' for 'personal level' where appropriate; for a defense of this move, see Hurley 2003.

intersubjective information is prior, at the subpersonal level, to information that differentiates self and other, does this have any implications about the basis of our personal-level knowledge of other minds? Unfulfilled predictions, implausible consequences, or incompatible neural circuitry could lead either away from the general idea of shared circuits for control, imitation, and simulation, or to a better specification of those shared circuits.

I draw attention as I proceed to some striking aspects of the shared circuits hypothesis. In particular, this hypothesis connects a shared information space for action and perception with a shared information space for self and other, while at the same time illustrating how the distinctions between self and other, and between the imagined and the real, can be imposed on these shared information spaces. In this model, information about intentional agents arrives in the subpersonal version of the first-person plural: without distinction or inference between self and other. Moreover, the shared circuits hypothesis avoids the classical sandwich and illustrates a horizontally modular architecture.<sup>62</sup> It conceives of perception and action as dynamically co-constituted and shows how cognitively significant resources, such as the self/other and imagined/real distinctions, and information for action understanding and planning, can emerge from the information space that perception and action share.

The shared circuits hypothesis is a theoretical model that describes a subpersonal functional architecture in five major stages or – better – layers.<sup>63</sup> Some of these could be further expanded into different sub-layers. Multiple instances of the shared circuits structure could be linked together into a network of such shared circuits, for hierarchical yet flexible control permitting decomposition and recombination of elements. Further questions arise about how the specific layers might map onto phylogenetic or ontogenetic stages. The order of the layers is intended to be logically intuitive and to reflect increasing complexity, but not necessarily to represent the order of evolution, development, or learning. In particular, the order of layers 1 and 2, and of layers 4 and 5, is heuristic, as I shall explain below.

What is essential to the shared circuits model is this subpersonal progression:

- predictive simulation of the observable effects of movement, which enables improvement in instrumental motor control
- ‘mirroring’ of observed action
- simulation of the internal causes of observed movement, which enables understanding of the observed actions of others as intentional.

This progression allows us to understand how the shared space for perception and action can be the basis of a shared intersubjective space and of the self/other and imagined/real

---

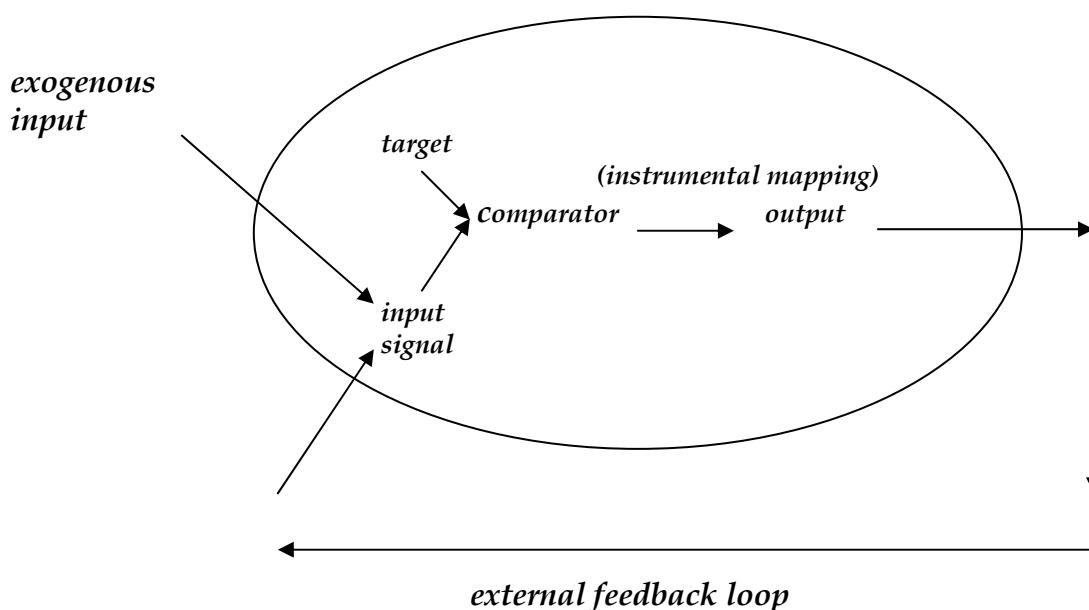
<sup>62</sup> See Hurley 1998, 2001 for arguments against the input-output picture of perception as input from the world to the mind and action as output from the mind to the world, with cognition sandwiched in between.

<sup>63</sup> The allusion to Brooksian subsumption architecture is intentional – another expression of what I call horizontal modularity. See Brooks 1999.

distinctions. Whether this theoretical model describes paths of evolution, development, and/or learning is a further question. In work in progress I consider how it can be adapted to the perception of emotion and emotional understanding.

First layer: basic adaptive feedback control. The first layer constitutes a simple adaptive control system for general purpose motor control, which can usefully be compared to a thermostat. The elements of this are (see Figure 1):

- (1) a target or reference signal (such as desired room temperature, in the case of the thermostat);
- (2) an input signal (such as actual room temperature), which is the joint result of (3) and (5);
- (3) exogenous events in the environment (such as nightfall);
- (4) a comparator, which determines whether the target and input signals match and the degree of any mismatch or error (e.g., the room is still five degrees below the desired temperature);
- (5) the output of the control system, (such as the level of heat output) which is determined by comparison between target and input signals (e.g., heat output is turned up if the room temperature is measured to be below the thermostat setting);
- (6) a feedback loop, by which output has effects on the succeeding input signal (e.g., actual room temperature rises).



**Figure 1. First layer: basic adaptive feedback control**

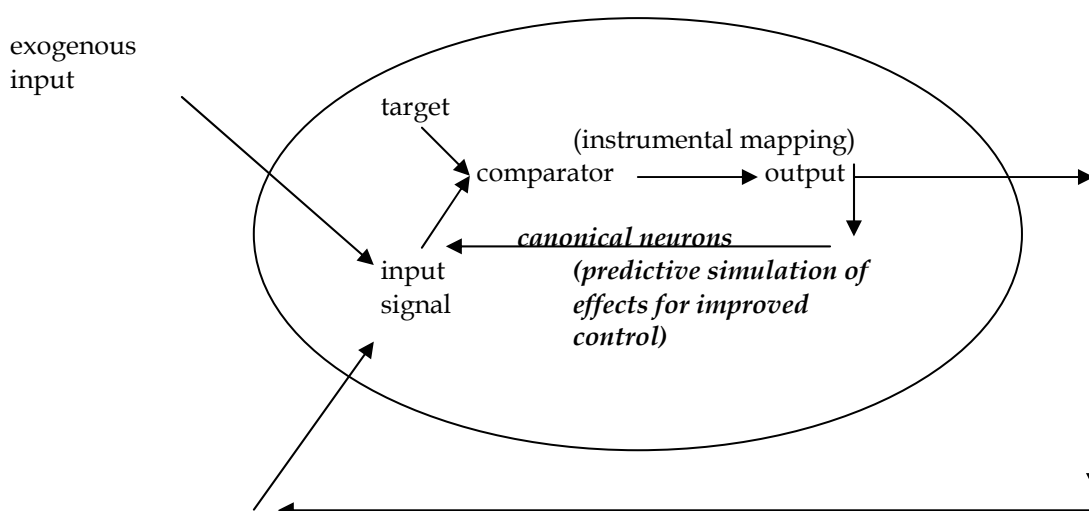
Consider the function that takes the system from target, in the context of actual input, to instrumental output.<sup>64</sup> I'll call this the *instrumental mapping*: in effect, it maps goal to means, or specifies the means that will be used to approach the target, in given circumstances. The feedback loop at this layer is relatively slow, since it operates in real

<sup>64</sup> An 'inverse model', in engineering terminology.

time (e.g., the room takes a while to warm up after the heat has been turned up). In organisms, this is input to the system that results from the organism's own activity<sup>65</sup>, by contrast with input to the system resulting from exogenous events<sup>66</sup>. Feedback from the organism's own activity includes, for example, visual and proprioceptive inputs resulting from movements of one's own hands, or movement through space, or manipulation of objects. Inputs from events originating in the external environment, both inanimate and animate, would include visual inputs resulting from movements by other creatures in a social group.

This kind of feedback control system is *adaptive* because it adjusts itself to changing environmental conditions and compensates for exogenous disturbances: in the presence of different exogenous events, different output is needed to achieve the target. The control process is cyclical and dynamic; it does not have discrete steps or a nonarbitrary start or finish. Input is as much an effect as a cause of output. Information about inputs is not segregated from information about outputs; the dynamic relations among inputs and outputs are critical for control. This feature will be preserved as further layers are added; to the extent that perception and action arise out of a system with this basic feature, they share a fundamental information space (see Hurley 1998, 2001).

Layer 2: low-level simulation. Predictive simulation of effects added to basic adaptive feedback control. An inner loop is now added, which maps a copy of the output signal directly onto the 'expected' input signal, or means to results.<sup>67</sup> Over time an association is established between copied output and subsequent input, so that in effect a copy of the motor output signals comes to evoke the associated input signal. It can then operate as a simulation of feedback, to predict the consequences of output on input. (See Figure 2; new aspects are italicized.)



<sup>65</sup> 'Reafference'.

<sup>66</sup> 'Exafference'.

<sup>67</sup> In organisms this mapping is often understood in terms of *efference copy* (or *corollary discharge*); in engineering it is referred to as a *forward model*.

external feedback loop

**Figure 2. Second layer: predictive simulation of effects for improved control**

This process provides a general purpose improvement in the functioning of the instrumental motor control system, because the system no longer needs to wait for output to produce actual feedback effects. Rather, the system can bypass relatively slow processes by learning and then anticipating the likely effects of output on input. For example, a temperature control system with this feature doesn't need to wait for the room actually to heat up to a certain level when the heat output is increased by a certain amount; it can predict that this will happen and can adjust its output accordingly. A *simulation* of the expected effects of the system's output speeds up the control process and smoothes the appropriate behavioral trajectory. In the case of a significant mismatch between real and simulated input, a local switch can default back to actual feedback control while the predictive simulation is further fine-tuned to improve its subsequent predictions.<sup>68</sup> This simulation is low-level, in the sense that it can perform its speeding and smoothing control functions without the system needing to monitor continuously or to access globally whether it is using actual or simulated feedback.

Recall that the order of the layers is heuristic and does not necessarily represent the order of evolution, development, or learning. For example, in the learning of particular tasks, layer 2's forward models may be acquired from feedback, enabling motor prediction, before layer 1's instrumental mappings are acquired, enabling motor control.<sup>69</sup> One does not necessarily have to be pursuing a goal in order to learn to predict the sensory consequences of movement, even if it is natural to conceive of such prediction of feedback in an instrumental context.

Notice, however, that a system that includes predicted feedback from the organism's own movements in addition to actual feedback has the resources to track the distinction between information about events in the world and information about goal-directed activity originating in the organism, that is, its behavior. When the train I am on pulls out of the station, I register movement relative to the train on the next platform, but this does not necessarily give me information about whether my train or the train on the next platform has begun to move. Comparison of predicted feedback from action with actual input gives an organism the resources to resolve the parallel subpersonal ambiguity, and hence provides information about the distinction between activity by the self and activity by the world.<sup>70</sup> This information could provide part of the basis for the personal level distinction between action and perception, which on this view would emerge from a

---

<sup>68</sup> See Grush forthcoming on how Kalman filters combine real and simulated input continuously, and recall that Grush's sense of 'emulation' is very different from the sense used in the summary above; see Wolpert & Kawato 1998; Haruno et al 2001; Wolpert et al 2003; on the mosaic model and selection among different forward models.

<sup>69</sup> Flanagan et al 2003; here I am indebted to comments from Marco Iacoboni.

<sup>70</sup> For discussion of this familiar point, and references, see Hurley 1998, 140-141 and *passim*. 'Self' here does not entail 'person', but is neutral between persons and other animals.

shared information space and processing resources. Note that information for the action/perception distinction is prior to and more general than information for the self/other distinction (see layers 3 and 4 below). In this sense there are more and less fundamental layers of information about self.

At this point it could be predicted that cells that mediate the association between copies of motor signals and actual input signals might come to have both motor and sensory fields. Suppose an animal typically acts in a certain way on the perceived affordances of a certain kind of object: eating a certain kind of food in a certain way, for example. There will be associations between copies of the motor signals for the eating movements and a multimodal class of inputs associated with such objects and the eating of them. Any cells that mediate this association might thus have both sensory and motor fields that between them capture the affordances of the objects in question. *Canonical neurons* are candidates for such predicted sensorimotor affordance neurons.<sup>71</sup>

Layer 3: mirroring for priming, emulation, and imitation. Now consider how the system described so far would apply to movements that produce visual feedback for their agent, as when a creature watches his own hand movements. (The contrast here is with movements such as facial expressions: while they produce proprioceptive feedback, the creature cannot normally see his own facial expressions.) As the creature watches his own hand movements, an association is formed between copies of motor signals for such movement and visual feedback from such movements. Here it could be predicted that cells that mediate this association might have matching sensory and motor fields. If the first creature watches another creature perform hand movements of the same kind and he receives similar visual inputs, these will also activate his sensorimotor matching neurons with their motor fields. The sensory fields of such matching neurons cannot tell the difference between his own actions of this kind and similar actions by others; they will fire when he does something or observes someone else do the same thing. *Mirror neurons* are of course candidates for such predicted matching sensorimotor neurons, and provide the neural underpinning for the kind of primitive blended intersubjective information space described by Gallese in terms of a shared manifold and by Gordon in terms of constitutive mirroring. Note the intimate relationship between the sharing of circuits for action and perception and for self and other: the blended intersubjective information space is a specification of and presupposes the generic blended sensorimotor information space.

Assume now that the sensorimotor matching association is bi-directional. Then, as well as copies of motor signals predictively simulating input signals, as described so far, input signals can also evoke motor signals. That is, assume that the predictive simulation can run in reverse.<sup>72</sup> The result of such *reverse simulation* would be *mirroring*: motor copying at some level or levels. (See Figure 3.) If a particular shared circuit controls details of movement<sup>73</sup>, a predicted result would be mirroring of similar movement: motor priming.

---

<sup>71</sup> Rizzolatti 2004; Iacoboni 2004; Gallese 2004.

<sup>72</sup> See and cf. Gallese & Goldman 1998; Blakemore & Decety 2001.

<sup>73</sup> As in Rizzolatti's *low-level resonance*; 2004.

If the circuit instead controls the result of movement<sup>74</sup>, rather than the detailed movements that are the means to these results, a predicted result would be mirroring of goals: emulation. If shared circuits for both motor means and results are themselves flexibly associated so that they can function instrumentally together, they would enable full-fledged imitation in which means as well as ends are copied (as revealed by the two-action methodology for identifying imitation). Such full-fledged imitation would be predicted to be rarer than either response priming or emulation separately, since it would require circuits for both means and ends, appropriately linked together. And indeed it is rarer.

Instrumental mappings and reverse simulation mirroring (henceforth: *mirroring*, for short) both go from input signals to output. The distinction between them is functional (the neural paths that perform these functions might overlap). An instrumental mapping functions instrumentally (!), to bring about a goal by matching input to a target within a comparator system. Mirroring does not in itself have this instrumental function.<sup>75</sup> The priming of my own action by observing someone else's similar action is rather a by-product of the presence of the predictive simulation, which functions at layer 2 to improve the functioning of the control system. However, this priming may in due course be exapted for other functions, such as those associated with imitation and simulation for action understanding. The neural mechanism by which such reverse functionality might be acquired is a matter of speculation.<sup>76</sup>

Circuits with this mirroring aspect could function in a variety of useful ways to capture information about the instrumental structure of action. If seeing someone move in a way that produces a result generates motor activation associated with a similar *goal* for the observer (though not necessarily associated with the same *movements* in the observer), that could provide information for action understanding, about the goal-directness of the observed action. If seeing someone move in a way that produces a result also generates motor activation associated with the same *movements* in the observer, that could provide information for imitative learning, that is, about how to achieve the goal by means of the observed movements. Such circuits could be linked to enable understanding that observed movements are the means to a sub-goal, which is in turn the means to a further goal, and so on; or to enable imitative learning of how to achieve a goal by achieving a more proximal goal, and so on. Again, the distinction between ends and means is not absolute; movements can be the means to other movements, which are in turn the means to something else. Such linked circuits could generate behavioral building blocks or modules that could be strung together in program level imitation, of sequences (Byrne, 2004) or of hierarchical structures (Whiten, 2004). They could allow an infant to form three-way associations between observed behavior by its parents (who have survived to reproduce, so may have adaptive behaviors not all of which are heritable), observed circumstances in which its parents perform such behavior, and its own similar behavior. Such associations could drive contextual imitation: act like that, when the environment is like this (Byrne, 2004).

---

<sup>74</sup> As in Rizzolatti's *high-level resonance*, for example, in the monkeys in whom mirror neurons were discovered.

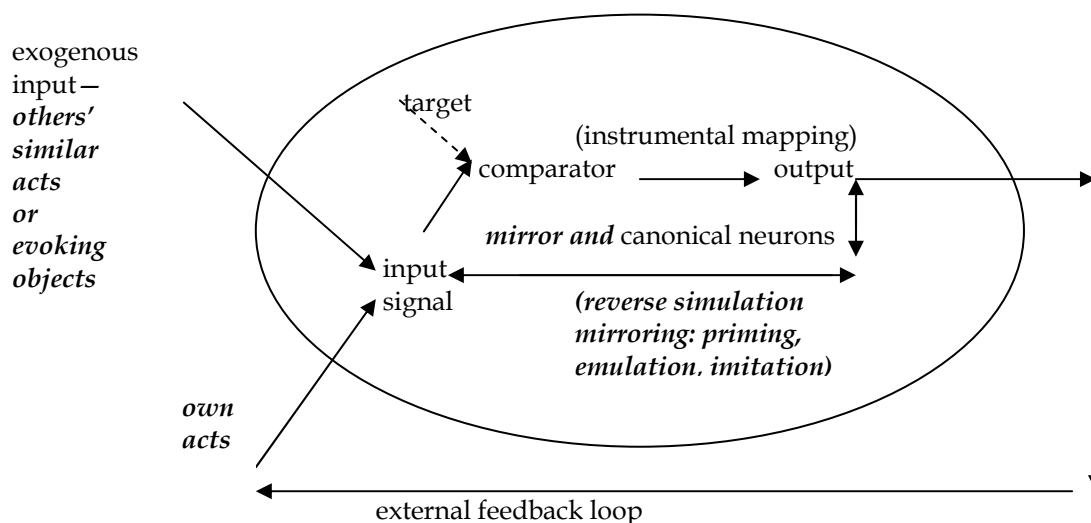
<sup>75</sup> See and cf. Peterson & Trapold, 1982.

<sup>76</sup> See and cf. Heyes, 2004 on cells that fire together and wire together.

Note that the sensorimotor affordance associations described in the second layer (mediated by canonical neurons?) could also be bilateral. If so, observation of an object that affords some type of action would be predicted to prime the type of action afforded.<sup>77</sup>

---

<sup>77</sup> See and cf. Lhermitte's utilization syndrome patients; Lhermitte 1983, 1986; Lhermitte et al 1986.



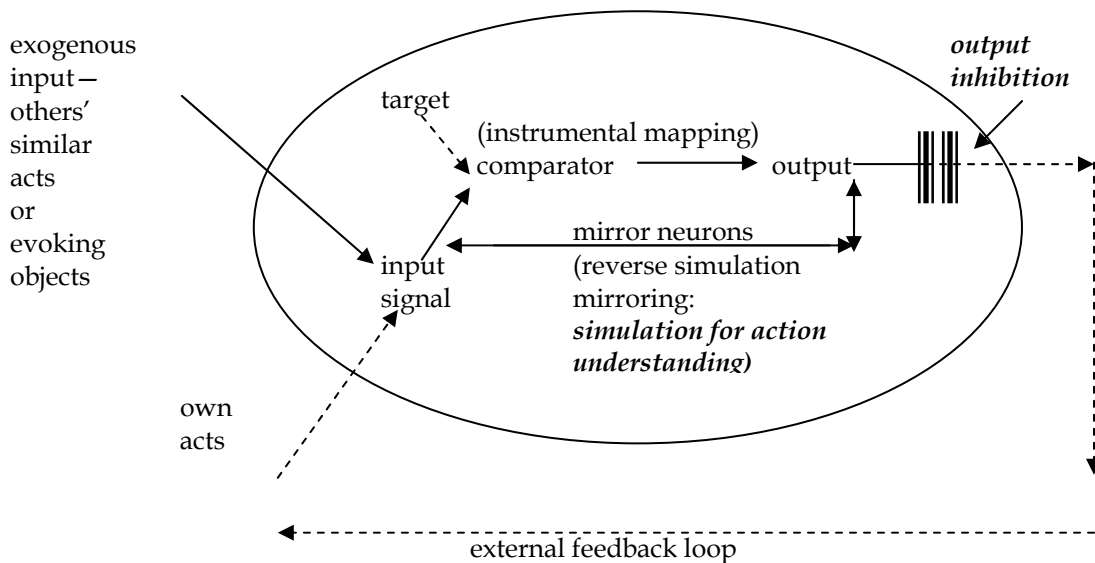
**Figure 3. Third layer: reverse simulation mirroring for priming, emulation, imitation**

So far, the reverse simulation and mirroring account does not explain copying of movements that cannot be seen by their agents. How can a correspondence be established between one's own acts and similar acts by others, when there is no feedback from the organism's own actions in the same modality as its observations of others' acts? For example, a creature receives visual input when observing another's facial expressions, but normally only receives proprioceptive, not visual feedback from its own facial expressions. How can these be compared and matched across modalities? How then can an association be established between my seeing another's facial expression and my making a similar expression myself? One answer is that some such supramodal correspondences are innate (Meltzoff, 2004). Another is that they are acquired in a variety of ways, through experience with mirrors, or with being imitated (Heyes, 2004).

The shared circuits model is compatible with these suggestions; it has no commitments about whether such opaque correspondences are innate, acquired, or both. It also naturally accommodates another suggestion: that *stimulus enhancement* can establish associations between one's own and others' similar acts in such cases. Suppose a social creature repeatedly visually observes others' actions of a certain type, and its attention is thereby drawn to the characteristic objects of such actions. Such stimulus enhancement repeatedly evokes in the observer an innate or otherwise acquired response to those objects. As a result, an association is formed between visual observations of others' actions and one's own similar action. This is not initially imitation or any kind of copying; the object independently evokes others' and one's own acts. But while the link is initially indirect, nevertheless an association between own and others' acts may be established. Cells that mediate this association may acquire mirror properties such that subsequently merely observing another's act comes to prime similar action by the observer. In this way mere stimulus enhancement may develop into copying, and an indirect stimulus

enhancement link into a direct sensorimotor matching link. This suggestion about how opaque correspondences could be established is similar to one Heyes (2004) makes about the mediating role of words, but it applies to stimulus enhancement in general.<sup>78</sup>

Layer 4: simulation for action understanding with output inhibition. Next consider the possibility that a creature might observe another's act, which primes a similar act in the usual way, yet its own action is inhibited so that the observed behavior is not actually copied. In effect, the output of the mirroring function is taken "off-line" prior to motor output. Since observing the other's act is still associated with motor priming even when copying is inhibited, such observation could be interpreted as providing the observer with simulated information about the intentional character of the observed act. Simulation for action understanding is off-line copying. But now what is being simulated are not effects, but causes. Instead of predicting the feedback that results from action, we now have simulation of the motor signals that would generate feedback similar to what is observed. Applied to emulation circuits that control the result of movement, the simulation would provide information about the goal to which the other's movement is directed, or the instrumental structure of the action. The ability to pick up the information that another's movement is directed toward a certain goal can be regarded as enabling an early stage in understanding other agents and hence other minds. (See Figure 4.)



<sup>78</sup> Heyes' ASL model (2004) claims that visual and motor representations are linked according to the same Hebbian principles whether or not the movement can be seen by its agent. The only difference is that in the case where it can be seen, self observation will lead to the formation of links between movements that are the same from a third party perspective. What I'm here regarding as stimulus enhancement could be regarded as acquired equivalence learning. The ASL model cites words as examples of the kind of stimuli that could act as the 'third term' in acquired equivalence learning, but acknowledges that, as in most experiments on acquired equivalence in animals, the third term is often a non-linguistic stimulus. Thanks here to Cecilia Heyes.

**Figure 4. Fourth layer: simulation for action understanding  
with output inhibition**

Although it uses the same circuit in reverse, simulation for action understanding can function at a higher level than the predictive simulation for speeding and smoothing control described in layer 2.<sup>79</sup> Recall that the basic functions of predictive simulation in a control system do not require the system to monitor continuously whether it is relying on actual or predicted feedback, though it should be able to switch between them as needed. In other words, as long as the predictive simulation works well and there is no significant mismatch in retrospect, the system does not need to know that it is using the predictive simulation to improve its own functioning. The distinction between actual and simulated feedback does not have global significance for the system.

By contrast, for simulation flexibly to subserve, as needed, understanding as opposed to copying an action, the system has to monitor the distinction between states in which the system's output is inhibited and states in which it is not; this distinction *is* of global significance. While we have seen that copying can have many beneficial functions, unselective overt copying of an action can in some circumstances have disastrous results for the observer. So the capacity selectively to inhibit overt copying is also adaptive, and should be expected to develop. Moreover, the system needs to keep track of whether the action it is 'entertaining' is merely simulated, for purposes of understanding another's behavior, or under execution, since these in general have quite different consequences and make quite different demands on subsequent behavior. Information about whether a movement is another's or one's own thus comes to overlay the primitive blended intersubjective manifold. Information about the distinction between self and other in this way emerges. Keep in mind, however, that in this hypothesis the level of description is subpersonal; while this subpersonal information is enabling, it is a further question how it is used at the personal level.

In particular, the shared intersubjective space is here prior to the self/other distinction, so that subpersonal information about instrumental actions in effect arrives in the first-person plural, in a form that does not distinguish or infer between self and others. Subpersonal processing of information about other agents is a matter of simulated recentering of the first-personal information processing that enables intentional action, rather than of inference from first-person information to third-person information. At the level of subpersonal information, the problem of 'knowledge' of other minds is reconfigured: it is neither one of starting from information about the self and constructing a bridge across a gulf to information about other persons, nor one of starting from information about other persons and from the resources it provides somehow generating information about the self. The shared circuits hypothesis gives concrete if subpersonal form to the interdependence and parity of information about self and other minds.

---

<sup>79</sup> By 'control' here, I refer to the overall function of the control system, not merely to that of the instrumental mapping component.

Again, it is a further question how these subpersonal relations are reflected at the personal level. Do they give any support to a parallel priority of the first person plural at the personal level? How should 'priority' indeed be understood in this question: as a question about development, or about the structure of mature understanding of other persons, and what is the relation between these? Can personal level understanding and knowledge of other minds be noninferentially based on or enabled by reliable subpersonal information? Is there any reason, conceptual or empirical, to believe that the problem of knowledge of other minds is similarly reconfigured at the personal level, so that it is neither one of starting from the first person perspective and constructing a bridge across a gulf to the third person, nor one of starting from the third person perspective and from the resources it provides somehow creating the first person perspective? Careful further thought is needed here. We should not simply help ourselves to an isomorphic projection from the subpersonal to the personal levels, but nor should we assume that the structure of subpersonal information processing has no implications for the personal level.

One way of responding to these issues is suggested by the affinities between the shared circuits hypothesis and Gordon's version of simulation theory.<sup>80</sup> In Gordon's felicitous phrase, constitutive mirroring *multiplies the first person*, through a process of making sense of observed behavior and the self's matching response together, under a common scheme of reasons, a process that assigns incoherent mental states to different persons (2004). While the shared circuits model offers a subpersonal description in which first-person plural information is prior to first-person singular and third-person singular information, Gordon's account of the multiplication of the first person under a scheme of reasons is more ambitious in linking subpersonal constitutive mirroring to personal level understanding of other minds.<sup>81</sup>

Layer 5: counterfactual input simulation for deliberation and planning. Finally, the system can be taken off-line on the input side as well as the output side. Counterfactual inputs of possible acts and affordances can be simulated and the resulting motor

---

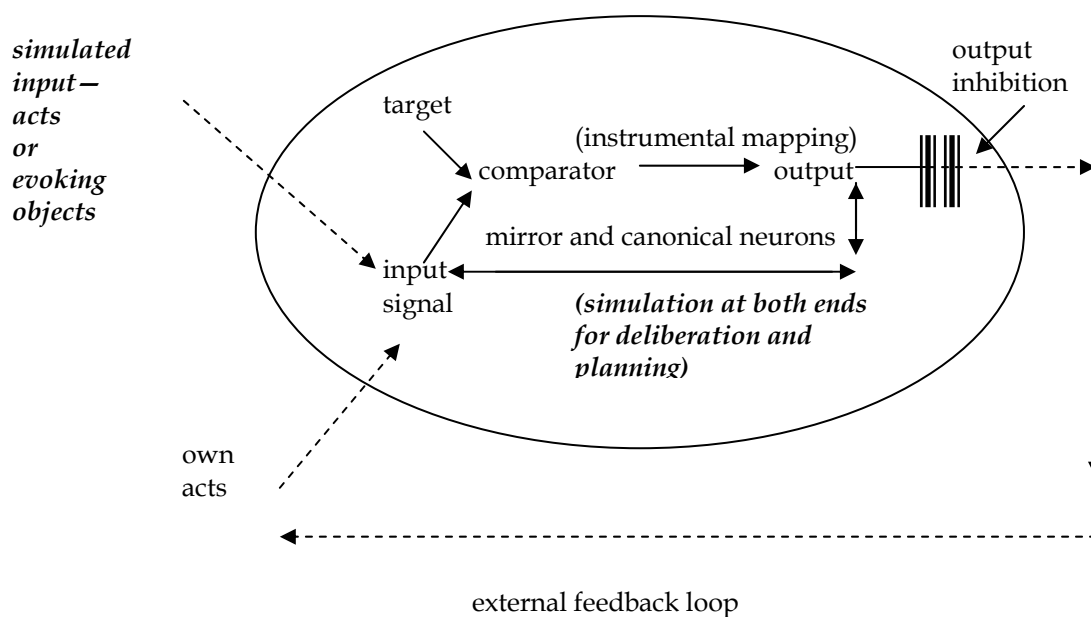
<sup>80</sup> See especially Gordon 1995a, pp. 56, 58, 68; see also 2002, 2004.

<sup>81</sup> Gordon appeals to *ascent routines* to explain how simulation can underwrite mind-reading without depending on inference from the first to the third person, as other versions of simulation theory do (see Gordon 1995a, 2004; compare Gallese and Goldman 1998). When I use an ascent routine, I answer a meta-question about my own or another's mental states by looking at the world; ascent routines are as well suited in principle to answering questions about another's mental states as about one's own. For example, to answer a question about whether I believe *p*, I consider whether *p* is true; to answer a question about whether another believes *p*, I perform an egocentric shift and imaginatively recenter myself to the other's perspective, and then again consider whether *p* is true. Similarly, for questions about what I or another perceive or intend: I look out at the world and the reasons it provides, though in the case of others having first transformed myself imaginatively. Note that on this view, to answer questions about what I or others believe, perceive, or intend, someone must first have the ability to perceive and act in the world. There is here another parallel, between Gordon's conception of ascent routines and the first aspect of the shared circuits model I have noted in the text: the way a shared intersubjective space is distilled out of and simulatively employs the shared perception/action information space.

activations entertained and compared without commitment to action and its costs<sup>82</sup>; circuits for means and ends can be linked and recombined flexibly. Simulation at both ends could provide subpersonal information to enable imagination, deliberation and planning, and counterfactual and more elaborate instrumental reasoning at the personal level. For these functions, it would also be essential for the system to monitor whether it is simulating or not: the informational basis for the distinctions between the imagined/supposed/apparent and the real thus emerges, close on the heels of the distinction between self and other. (See Figure 5.) These distinctions are necessary (if not sufficient) for some aspects of the normativity and possibility of mistake that are characteristic of the personal level, so understanding how they emerge at the subpersonal level may help in turn to understand how the personal level emerges from the subpersonal.

---

<sup>82</sup> See Millikan's squirrel (forthcoming).



**Figure 5. Fifth layer: counterfactual input simulation for deliberation and planning**

Again, keep in mind that the order of the layers presented here is heuristic and does not necessarily represent the order of evolution, development, or learning; those are questions for further investigation. Layer 4's simulation and inhibition of output may accompany or follow rather than precede layer 5's simulation of input. That is, the shared circuits hypothesis does not specify the phylogenetic or developmental priority between subpersonal information about self vs. other and subpersonal information about the imagined vs. the real. Rather, it provides generic, adaptable tools for framing more specific hypotheses.

Concluding remarks. The shared circuits hypothesis provides a unified subpersonal architecture for control, imitation, and simulation at a middle level of description: a functional level above that of neural implementation but below that of the normatively constrained and/or conscious personal level. We can now step back from the details to see the broad outline: The feedback effects of certain motor outputs, such as visual inputs that result from certain movements, are predicted via simulation; such prediction can benefit instrumental control. Mirroring reverses this predictive simulation, so that observation of movements of a certain kind by another induces in the observer motor output that would typically cause such movements. While copying can be beneficial, the capacity to inhibit actual copying is also adaptive. When overt copying is inhibited, the causes of the observed behavior are nevertheless simulated, enabling action understanding. The progression is from simulation of effects through mirroring to simulation of causes.

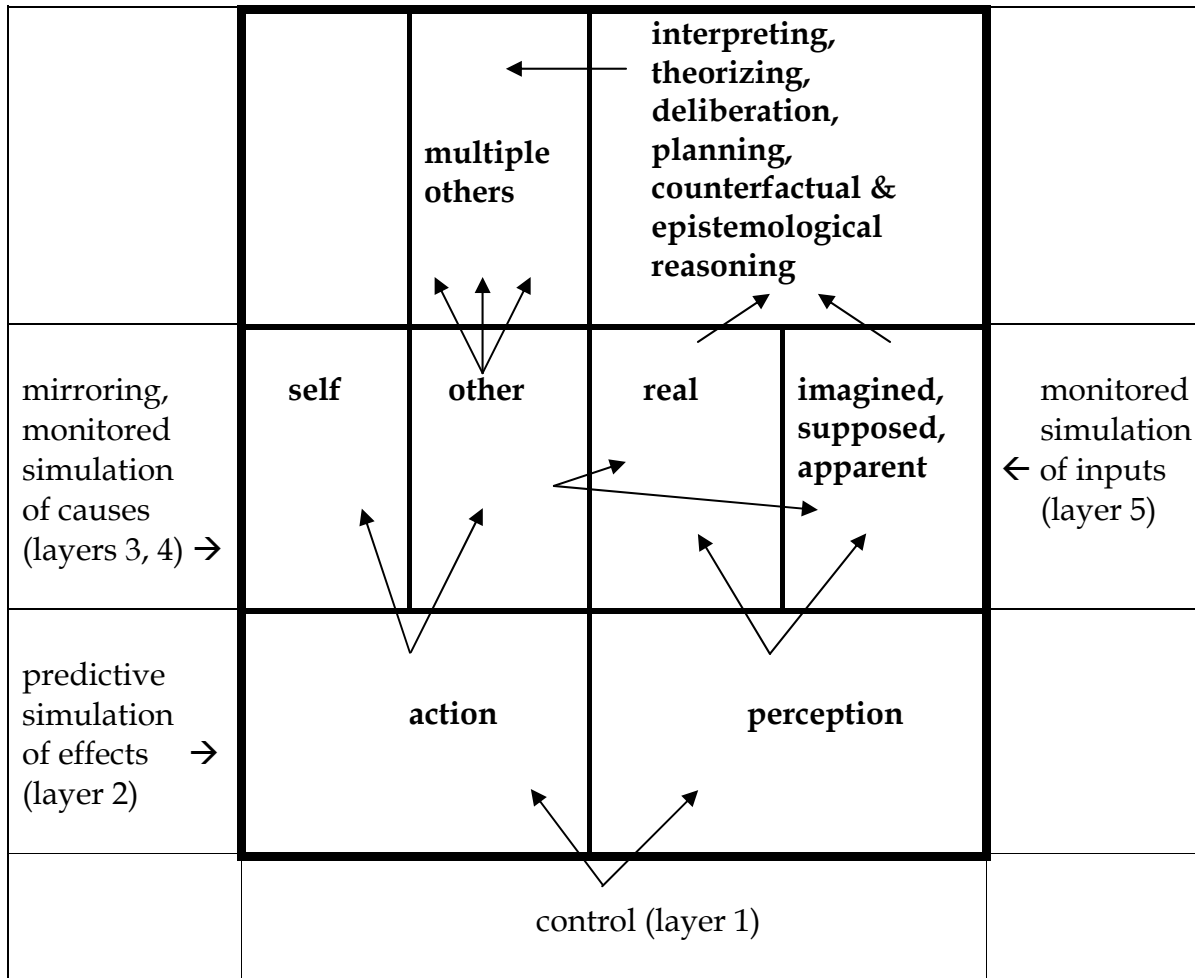
Hand in hand with this building up of mechanisms goes the building up of a subpersonal informational structure that stands in an *enabling* relation to various aspects of the

personal level (see and cf. McDowell 1994). Such enabling subpersonal information is *not* necessarily the conscious basis or, or the reason for, or sufficient for actions or judgements by the person (or animal). Nevertheless, what people can do depends on and is enabled by subpersonal informational resources. Figure 6 displays the way subpersonal informational resources described in the shared circuits model (outside the bold box) could be drawn on to enable various personal-level distinctions and capacities (inside the bold box).

At the lower levels, simulative prediction of effects within an instrumental control system provides information that enables actions to be distinguished from perceptions. In this sense, the action-perception distinction is informationally fundamental; but this is not yet information for a full-fledged self-world distinction, since the actions of the self and those of others are not yet distinguished.

At the higher levels, consider first the action side. The mirroring of actions provides information about the instrumental structure of action, but does so in the subpersonal first-person plural: one's own intentional acts and others' are gathered together (for simplicity, this is not depicted separately in Figure 6). Then, monitored inhibition of copying and simulation for action understanding impose a basic self-other distinction on intersubjective information about intentional action. The self-other distinction is thus at its informational base a distinction between self and others as intentional agents, i.e. the agents of instrumentally structured actions. Capacities for interpretation and theorizing provided by counterfactual simulation (see next paragraph) may build on this fundamental simulative self-other distinction to fine-tune differentiations and identifications of multiple other intentional agents. From the action side of the basic distinction thus emerges information for perception of the actions of others.

Focus next on the perception side of the basic distinction: monitored simulation of inputs to control systems (in addition to predicted effects) provides counterfactual information and information for distinctions between actual and possible states of the world: between what is real and what is merely imagined or supposed or apparent. These resources in turn provide information for theorizing of various kinds, and may enable the higher-level interpretation of multiple other agents, as well as counterfactual and epistemological reasoning, deliberation and planning.



**Figure 6. Structure of subpersonal informational resources**

In considering how the shared circuits subpersonal functional architecture might map onto the neural and personal levels, care is needed to avoid over-simple interlevel projections and isomorphism assumptions. Here, the model may play a useful heuristic role even if it proves to be wrong in details. For example, looking down to the level of neural implementation, we can ask: where might the postulated comparators be located?<sup>83</sup> Are shared circuits for the results of action found in prefrontal areas while those for detailed movements are in parietal areas? Does the model bear any relationship to the distinction between ventral and dorsal processing streams? Does it cast any light on the presence and function of mirror neurons in Broca's area and their relation to linguistic capacities? I have sketched the dynamics of the shared circuits model in cybernetic terms, but if neural implementations can be found, their interactive behavior through time could be represented as the evolution of a phase space in the manner of dynamical systems theory, and its attractor structure investigated.

<sup>83</sup> PF? STS? Compare Rizzolatti 2004 and Iacoboni 2004.

Looking up to the personal level, we can ask: What behavioral and cognitive predictions does the model provide? Intentional agents achieve their goals by means that can be given successively finer specifications, related by an asymmetrical 'do  $x$  by doing  $y$ ' relation: for example, I turn on the light by flipping the switch by moving my fingers. If we envisage a series of spectra with control of the ultimate result or goal of action at one extreme, and control of detailed fine movements that are the means to the result at the other extreme, then the shared circuits model could apply at successive linked points along such spectra. Thus the means outputted to the target of one circuit could be the target of the next circuit. A network of such linked circuits would support hierarchical control while permitting the flexible decomposition and recombination of goals and means. What relationship might such recombinant flexibility have to the recombinant flexibility characteristic of language? What does the model suggest about the functional relationships among three distinctive human capacities, for imitation, mind-reading, and language?<sup>84</sup> The shared circuits model has a cybernetic rather than a conceptual structure, yet it may provide informational resources for sophisticated cognitive abilities usually understood to display conceptual structure, such as the abilities to deceive and to cooperate, to interpret other minds, to use language, to deliberate and plan, to reason instrumentally and counterfactually. The model thus has a bearing on issues about whether the subpersonal resources on which conceptual abilities depend must themselves have conceptual structure. What implications might it have for the issue of whether simulation approaches to mind-reading require an inference from the first to the third person? What constraints does the model suggest on the relationships among various personal level distinctions: between action and perception, between self and other, between the real and the imagined? Comparator structures play a prominent role in both the shared circuits model and in various recent approaches to consciousness; does this suggest any particular way of linking theories of consciousness and social cognition, or of updating social theories of consciousness?<sup>85</sup> Can the model play any role in distinguishing conscious and unconscious mental states and processes? How might the model be extended to include the perception of facial expression and the emotional mirroring postulated by various researchers?<sup>86</sup> Might the layers of the model usefully be mapped onto evolutionary or developmental stages, in theorizing the imitative and mind-reading abilities of other animals or children?

I conclude by summarizing the shared circuits hypothesis. Theories about the control, imitative, and simulative functions of the mirror system, and evidence from imitation studies for ideomotor and common coding theories, suggest that perception and action share a fundamental information space that is preserved as higher cognitive capacities and distinctions are built on it. The distinction between results and the means to those results, essential to goal-directed, perceptually-guided intentional action as well as to imitative learning, emerges as a flexible articulation of this shared processing. But perception remains fundamentally enactive, in a way that challenges orthodox views of perception

---

<sup>84</sup> Cf. Rizzolatti & Arbib, 1998; Iacoboni, 2004; Meltzoff, 2004.

<sup>85</sup> See and cf. Hesslow 2002; Frith et al 2000; Gray forthcoming; Jeannerod 1997; Milner and Goodale 1995, 64.

<sup>86</sup> I have work in progress on this question. See Rizzolatti 2004, Gallese 2004; Iacoboni 2004; Decety and Chaminade 2004; see also Hesslow 2002, which I am grateful to John Cummins for drawing to my attention.

and action as separately constituted and hence of perception as motivationally inert. The intersubjectivity characteristic of human beings, their distinctive capacity to understand and empathize with one another, emerges as a specialization of enactive perception: I perceive your action enactively, in a way that immediately engages my own potential similar action, thus enabling me to understand, or to imitate, your action. Shared processing of the actions of other and self is a special aspect of the shared processing of perception and action. This subpersonal informational structure may have implications for the epistemology of other minds. Within this informational structure, it is not so much that intersubjectivity bridges a self/other gap as that the self/other distinction is imposed on the fundamental information space that self and other share. Simulation theories of mind-reading can be right about shared processing for self and other with respect to this fundamental intersubjectivity, even if more advanced aspects of mind-reading require theorizing, in ways enabled by language.

I have emphasized three noteworthy aspects of the shared circuits hypothesis. First, it distills a shared intersubjective space (enabling imitation, intersubjective empathy, and action understanding) from a shared information space for action and perception (understood in terms of control processes). Second, it illustrates how the distinctions between perception and action, between self and other, and between the real and the imagined can be imposed on these shared information spaces.<sup>87</sup> These informational distinctions play important roles in enabling the mental life characteristic of persons. In particular, the shared intersubjective space is viewed as prior to the self/other distinction, and information about intentional agents arrives in the first person plural, in a form that does not register the self/other distinction. At the subpersonal level, the problem of 'knowledge' of other minds is reconfigured: it is neither one of starting from information about the self and constructing an inferential bridge across a gulf to information about other persons, nor one of starting from information about other persons and from the resources it provides somehow generating information about the self. The shared circuits hypothesis gives concrete form to the interdependence and parity of information about self and other minds. Finally, the shared circuits hypothesis avoids the classical sandwich and illustrates an alternative *horizontally modular* architecture.<sup>88</sup> On this view, perception and action are dynamically co-constituted and cognitively significant resources, such as the self/other and imagined/real distinctions and information for action understanding and planning, emerge from the information space that perception and action share.

---

<sup>87</sup> In a way that converges with the broadly Kantian view that the latter two distinctions are intimately connected; cf. Strawson 1959, 1966.

<sup>88</sup> Hurley 1998, 2001; see also Brooks 1999.

## References

- Adolphs, R. (2002). Recognizing emotion from facial expressions: Psychological and neurological mechanisms. *Behavioral and Cognitive Neuroscience Reviews* 1(1), 21-61.
- Akins, C., & Zentall, T. (1996). Imitative learning in male Japanese quail (*Conturnix japonica*) using the two-action method. *Journal of Comparative Psychology*, 110, 316-320.
- Akins, C., & Zentall, T. (1998). Imitation in Japanese quail: The role of reinforcement of demonstrator responding. *Psychonomic Bulletin & Review*, 5, 694-697.
- Akins, C., Klein, E., & Zentall, T. (2002). Imitative learning in Japanese quail (*Conturnix japonica*) using the bidirectional control procedure. *Animal Learning & Behavior*, 30, 275-281.
- Anisfeld, M. (1979). Interpreting "imitative" responses in early infancy. *Science*, 205, 214-215.
- Anisfeld, M. (1984). *Language development from birth to three*. Hillsdale, NJ: Erlbaum.
- Anisfeld, M. (1991). Neonatal imitation. *Developmental Review*, 11, 60-97.
- Anisfeld, M. (1996). Only tongue protrusion modeling is matched by neonates. *Developmental Review*, 16, 149-161.
- Anisfeld, M. (2004). No compelling evidence to dispute Piaget's timetable of the development of representational imitation in infancy. In S. Hurley & N. Chater (Eds.), *Perspectives on Imitation: From Neuroscience to Social Science* (Vol. 2, pp. 000). Cambridge, MA: MIT Press.
- 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, 111-122.
- Arbib, M. (2002). The mirror system, imitation, and the evolution of language. In C. Nehaniv & K. Dautenhahn (Eds.), *Imitation in animals and artifacts* (pp. 229 - 280). Cambridge, MA: MIT Press.
- Arbib, M. (forthcoming). From monkey-like action recognition to human language: An evolutionary framework for neurolinguistics, behavioral and brain sciences. *Behavioral and Brain Sciences*.
- Arbib, M., Billard, A., Iacoboni, M., & Oztop, E. (2000). Synthetic brain imaging: Grasping, mirror neurons and imitation. *Neural Networks*, 13, 975-997.
- Arbib, M., & Rizzolatti, G. (1997). Neural expectations: a possible evolutionary path from manual skills to language. *Communication and Cognition*, 29, 393-424. Also in P. van

Looche (ed.), *The nature of concepts. evolution, structure, and representation* (pp. 128-154). London: Routledge.

Baldwin, D. (1995). Understanding the link between joint attention and language. In C. Moore & P. Dunham (Eds.), *Joint Attention: Its origin and role in development* (pp. 131-158). Hillsdale, NJ: Erlbaum.

Baldwin, J. (1896), "A new factor in evolution", *American Naturalist* 30 (1896), 441-51, 536-53.

Bargh, J. (1997). The automaticity of everyday life. In R. Wyer (Ed.), *Advances in Social Cognition* (Vol. 10, pp. 1-61) Mahwah, NJ: Erlbaum.

Bargh, J. (1999). The most powerful manipulative messages are hiding in plain sight, *The Chronical of Higher Education* (January 29), B6.

Bargh, J. (forthcoming), Bypassing the Will: Towards Demystifying the Nonconscious Control of Social Behavior, in R. Hassin, J. Uleman, and J. Bargh, eds., *The New Unconscious*. New York: Oxford University Press.

Bargh, J., & Chartrand, T. (1999). The unbearable automaticity of being. *American Psychologist*, July, 462-479.

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, 230-244.

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, 1014-1027.

Bavelas, J., Black, A., Lemery, C., & Mullett, J. (1987). Motor mimicry as primitive empathy. In N. Eisenberg & J. Strayer (Eds.), *Empathy and its development* (pp. 317-338). Cambridge, UK: Cambridge University Press.

Bekkering, H., & Wohlschläger, A. (2002). Action perception and imitation: a tutorial. In W. Prinz & B. Hommel (Eds.), *Attention & performance XIX. Common mechanisms in perception and action* (pp. 294-314). Oxford: Oxford University Press.

Blackmore, S. (1999), *The Meme Machine*. Oxford: Oxford University Press.

Blackmore, S. (2000). The meme's eye view. In R. Aunger (Ed.), *Darwinizing culture: The status of memetics as a science* (pp. 25-42). Oxford & New York: Oxford University Press.

Blackmore, S. (2001). Evolution and memes: The human brain as a selective imitation device. *Cybernetics and Systems*, 32, 225-255.

- Blakemore, S., & Decety, J. (2001). From the perception of action to the understanding of intention. *Nature Reviews Neuroscience*, 2, 561-567.
- Boyd, R. and Richerson, P. (1982), 'Cultural Transmission and the Evolution of Cooperative Behavior', *Human Ecology* 10, 325-351
- Boyd, R. and Richerson, P. (1985), *Culture and the Evolutionary Process*. Chicago: University of Chicago Press.
- Boyd, R., & Richerson, P. (2000). Memes: Universal acid or a better mousetrap? In R. Aunger (Ed.), *Darwinizing culture: The status of memetics as a science* (pp. 143-162). Oxford & New York: Oxford University Press.
- Brass, M. (1999). *Imitation and Ideomotor Compatibility*. Unpublished dissertation, University of Munich, Germany.
- Brass, M., Bekkering, H., & Prinz, W. (2001). Movement observation affects movement execution in a simple response task. *Acta Psychologica*, 106, 3-22.
- Brooks, R. (1999). *Cambrian Intelligence*. Cambridge, MA: MIT Press.
- Byrne, R. (1995), *The Thinking Ape: Evolutionary Origins of Intelligence*. Oxford: Oxford University Press.
- Byrne, R. (1998). Imitation: the contributions of priming and program-level copying. In S. Braten, (Ed.), *Intersubjective communication and emotion in early ontogeny* (pp. 228-244). Cambridge, UK: Cambridge University Press.
- Byrne, R. (1999). Imitation without intentionality. Using string parsing to copy the organization of behaviour. *Animal Cognition*, 2, 63-72.
- Byrne, R. (2002a). Imitation of complex novel actions: What does the evidence from animals mean? *Advances in the Study of Behavior*, 31, 77-105.
- Byrne, R. (2002b). Seeing actions as hierarchically organized structures: Great ape manual skills. In A. N. Meltzoff & W. Prinz (Eds.), *The imitative mind. Development, evolution and brain bases* (pp.122-140). Cambridge, UK: Cambridge University Press.
- Byrne, R. (2004). Detecting, understanding, and explaining animal imitation. In S. Hurley & N. Chater (Eds.), *Perspectives on Imitation: From Neuroscience to Social Science* (Vol. 1, pp. 000). Cambridge, MA: MIT Press.
- Byrne, R., & Russon, A. (1998). Learning by imitation: A hierarchical approach. *Behavioral and Brain Sciences*, 21, 667-721.

- Byrne, R., and Whiten, A., eds. (1988). *Machiavellian Intelligence: Social Expertise and the Evolution of Intellect in Monkeys, Apes and Humans*. Oxford: Oxford University Press.
- Call, J., & Carpenter, M. (2002). Three sources of information in social learning. In K. Dautenhahn & C. Nehaniv (Eds.), *Imitation in animals and artifacts* (pp. 211-228). Cambridge, MA: MIT Press.
- 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.
- Call, J. and Tomasello, M. (1994). The Social Learning of Tool Use by Orangutans (*Pongo pygmaeus*). *Human Evolution* 9(4):297-313.
- Call, J. and Tomasello, M. (1999). A nonverbal theory of mind test: the performance of children and apes. *Child Development* 70: 381-395.
- Carpenter, M., Akhtar, N., & Tomasello, M. (1998). Fourteen- through 18-month-old infants differentially imitate intentional and accidental actions. *Infant Behavior & Development*, 21, 315-330.
- Carpenter, M., Call, J., & Tomasello, M. (in press). Understanding 'prior intentions' enables 2-year-olds to imitatively learn a complex task. *Child Development*.
- Carpenter, M., Tomasello, M., & Savage-Rumbaugh, S. (1995). Joint attention and imitative learning in children, chimpanzees, and enculturated chimpanzees. *Social Development*, 4, 217-237.
- Carruthers, P., & Smith, P. (1996). *Theories of Theories of Mind*. Cambridge, UK: Cambridge University Press.
- Chartrand, T., & Bargh, J. (1999). The chameleon effect: The perception-behavior link and social interaction. *Journal of Personality and Social Psychology*, 76, 893-910.
- Chartrand, T., Maddux, W., & Lakin, J. (in press). Beyond the perception-behavior link: The ubiquitous utility and motivational moderators of nonconscious mimicry. In R. Hassin, J. Uleman, & J. Bargh (Eds.), *The new unconscious*. New York: Oxford University Press.
- Christiansen, M. (1994). *Infinite languages, finite minds: Connectionism, learning and linguistic structure*. Unpublished PhD dissertation. University of Edinburgh.
- Christiansen, M. (2004). On the relation between language and (mimetic) culture. In S. Hurley & N. Chater (Eds.), *Perspectives on Imitation: From Neuroscience to Social Science* (Vol. 2, pp. 000). Cambridge, MA: MIT Press.

- 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.
- Dautenhahn, Kerstin, and Nehaniv, C. (2002), *Imitation in Animals and Artifacts*. Cambridge: MIT Press.
- Davies, M., & Stone, T. (1995a). *Folk Psychology*. Oxford, Blackwell.
- Davies, M., & Stone, T. (1995b). *Mental Simulation*. Oxford: Blackwell.
- Dawkins, R. (1976/1989). *The Selfish Gene*. Oxford: Oxford University Press.
- Dawkins, R. (1982). *The Extended Phenotype*. Oxford: Oxford University Press.
- Deacon, T. (1997), *The Symbolic Species: The Coevolution of Language and the Human Brain*. London: Penguin Books; New York: Norton.
- Decety, J. (2002). Neurophysiological evidence for simulation of action. In J. Dokic, & J. Proust (Eds.), *Simulation and knowledge of action* (pp. 53-72). Philadelphia: Benjamins Publishing Company.
- Decety, J., & Chaminade, T. (2003). Neural correlates of feeling sympathy. *Neuropsychologia*, *41*(2), 127-138.
- Decety, J., & Chaminade, T. (2004). The neurophysiology of imitation and intersubjectivity. In S. Hurley & N. Chater (Eds.), *Perspectives on Imitation: From Neuroscience to Social Science* (pp. 000). Cambridge, MA: MIT Press.
- Decety, J., & Grèzes, J. (1999). Neural mechanisms subserving the perception of human actions. *Trends in Cognitive Sciences*, *3*, 172-178.
- 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*, 1763-1777.
- Dennett, D. (1995). *Darwin's dangerous idea: Evolution and the meanings of life*. New York: Simon & Schuster. Also London: Penguin.
- Dijksterhuis, A. (2004). Why we are social animals: The high road to imitation as social glue. In S. Hurley & N. Chater (Eds.), *Perspectives on Imitation: From Neuroscience to Social Science* (Vol. 2, pp. 000). Cambridge, MA: MIT Press.
- 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. Also in: M. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 30, pp. 1-40). New York: Academic Press.

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, 865-877.

Eidelberg, L. (1929), 'Experimenteller beitrage zum Mechanismus der Imitationsbewegung', *Jahresbucher fur Psychiatrie und Neurologie* 45, 170-73.

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.

Fadiga, L., Fogassi, L., Pavesi, G., & Rizzolatti, G. (1995). Motor facilitation during action observation: A magnetic stimulation study. *Journal of Neurophysiology*, 73, 2608-2611.

Flanagan, J., Vetter, P., Johanson, R., and Wolpert, D. (2003), Prediction precedes control in motor learning. *Current Biology* 13:146-150.

Frith C., Blakemore S. & Wolpert D. (2000), "Abnormalities in the awareness and control of action", *Philosophical Transactions of the Royal Society B (Biological Sciences)* 355, 1771-1788.

Galef, B. (1988). Imitation in animals: History, definition and interpretation of data from the psychological laboratory. In T. Zentall & B. Galef (Eds.), *Social learning: Psychological and biological perspectives* (pp. 3-28). Hillsdale, NJ: Erlbaum.

Galef, B. (1998). Recent progress in the study of imitation and social learning in animals. In M. Sabourin, F. Craik & M. Roberts (Eds.), *Advances in psychological science, Vol. 2: Biological and cognitive aspects* (pp.275-279). Hove, UK: Psychological Press.

Galef, B. (2004), "Breathing new life into the study of animal imitation: what and when do chimpanzees imitate?" In S. Hurley & N. Chater (Eds.), *Perspectives on Imitation: From Neuroscience to Social Science* (). Cambridge, MA: MIT Press.

Gallese, V. (2000). The inner sense of action: Agency and motor representations. *Journal of Consciousness Studies*, 7(10), 23-40.

Gallese, V. (2001), The 'shared manifold' hypothesis: from mirror neurons to empathy. *Journal of Consciousness Studies*, 8, 33-50.

Gallese, V. (2003b). The manifold nature of interpersonal relations: The quest for a common mechanism. *Philosophical Transactions of the Royal Society of London*, 358, 517-528.

Gallese, V. (2004). "Being like me": Self-other identity, mirror neurons and empathy. In S. Hurley & N. Chater (Eds.), *Perspectives on Imitation: From Neuroscience to Social Science* (Vol. 1, pp. 000). Cambridge, MA: MIT Press.

Gallese, V., & Goldman, A. (1998). Mirror neurons and the simulation theory of mind-reading. *Trends in Cognitive Sciences*, 2, 493-501.

Gergely, G., Bekkering, H., & Király, I. (2002). Rational imitation in preverbal infants. *Nature*, 415, 755.

Gerrans, P. (forthcoming). *The measure of madness: Philosophy and cognitive neuropsychiatry*. Cambridge, MA: MIT.

Gibson, J. (1986). *The ecological approach to visual perception*. Hillsdale, N.J.: Lawrence Erlbaum Associates.

Gil-White, F. (2004). Common misunderstandings of memes (and genes). The promise and the limits of the genetic analogy to cultural transmission processes. In S. Hurley & N. Chater (Eds.), *Perspectives on Imitation: From Neuroscience to Social Science* (Vol. 2, pp. 000). Cambridge, MA: MIT Press.

Goldman, A. (1989). Interpretation psychologized. *Mind and Language*, 4, 161-85.

Goldman, A. (1992a). In defense of the simulation theory. *Mind and Language*, 7, 104-19.

Goldman, A. (2004). Imitation, mind-reading, and simulation. In S. Hurley & N. Chater (Eds.), *Perspectives on Imitation: From Neuroscience to Social Science* (Vol. 2, pp. 000). Cambridge, MA: MIT Press.

Gordon, R. (1986). Folk psychology as simulation. *Mind and Language*, 1, 159-71.

Gordon, R. (1995a). Simulation without introspection or inference from me to you. In M. Davies & T. Stone (Eds.), *Mental Simulation* (pp. 53-67). Oxford: Blackwell.

Gordon, R. (1995b). Sympathy, simulation and the impartial spectator. *Ethics*, 105, 727-742.

Gordon, R. (1996). 'Radical' simulationism. In P. Carruthers & P. Smith (Eds.), *Theories of theories of mind* (pp. 11-21). Cambridge, UK: Cambridge University Press.

Gordon, R. (2002). Simulation and reason explanation: the radical view. Special Issue of *Philosophical Topics*, 29, Nos. 1 & 2.

Gordon, R. (2004). Intentional agents like myself. In S. Hurley & N. Chater (Eds.), *Perspectives on Imitation: From Neuroscience to Social Science* (Vol. 2, pp. 000). Cambridge, MA: MIT Press.

Gray, J. (forthcoming), *Constructing Consciousness*. Oxford: Oxford University Press.

- Graziano, M., Taylor, C., Moore, T., and Cooke, D. (2002). The cortical control of movement revisited. *Neuron* 36:349-362.
- Greenwald, A. (1970). Sensory feedback mechanisms in performance control: With special reference to the ideomotor mechanism. *Psychological Review*, 77, 73-99.
- Greenwald, A. (1972). On doing two things at once: Time sharing as a function of ideomotor compatibility. *Journal of Experimental Psychology*, 94, 52-57.
- Grush, R. (1995) *Emulation and Cognition*. Doctoral Dissertation, Department of Philosophy, University of California at San Diego.
- Grush, R. (forthcoming) The emulation theory of representation: motor control, imagery, and perception. *Behavioral and Brain Sciences*.
- Hare, B., Call, J., Agnetta, B. and Tomasello, M. (2000). Chimpanzees know what conspecifics do and do not see. *Animal Behaviour* 59, 771-785.
- Hare, B., Call, J. and Tomasello, M. (2001). Do chimpanzees know what conspecifics know and do not know? *Animal Behaviour* 61, 139-151.
- Hari, R., Forss, N., Avikainen, S. Kirveskari, E., Salenius, S., & Rizzolatti, G. (1998). Activation of human primary motor cortex during action observation: A neuromagnetic study. *Proceedings National Academy of Science, USA*, 95, 15061-15065.
- Harris, P., & Want, S. (2004). On learning what not to do: The emergence of selective imitation in young children's tool use. In S. Hurley & N. Chater (Eds.), *Perspectives on Imitation: From Neuroscience to Social Science* (Vol. 2, pp. 000). Cambridge, MA: MIT Press.
- Haruno, M., Wolpert, D., & Kawato, M. (2001). Mosaic model for sensorimotor learning and control. *Neural Computation*, 13, 2201-2220.
- 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(5):1123-1128.
- Henrich, J., & Boyd R. (1998). The evolution of conformist transmission and the emergence of between-group differences. *Evolution and Human Behavior*, 19, 215-241.
- 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, 165-196.
- Herman, L. (2002). Vocal, social, and self-imitation by bottlenosed dolphins. In K. Dautenhahn & C. Nehaniv (Eds.), *Imitation in animals and artifacts* (pp. 63-106). Cambridge, MA: MIT Press.

- Hesslow, Germund (2002), Conscious thought as simulation of behaviour and perception. *Trends in Cognitive Sciences* 6(6):242-247.
- Heyes, C. (1993). Imitation, culture and cognition. *Animal Behaviour*, 46, 999-1010.
- Heyes, C. (1996). Genuine imitation? In C. Heyes & B. Galef Jr. (Eds.), *Social learning in animals: The roots of culture* (pp. 371-389). San Diego: Academic Press.
- Heyes, C. (2001). Causes and consequences of imitation. *Trends in Cognitive Sciences*, 5, 253-261.
- Heyes, C. (2004). Imitation by association. In S. Hurley & N. Chater (Eds.), *Perspectives on Imitation: From Neuroscience to Social Science* (Vol. 1, pp. 000). Cambridge, MA: MIT Press.
- Heyes, C. (1998). Theory of mind in nonhuman primates. *Behavioral and Brain Sciences* 21:101-114.
- Heyes, C., & Dickinson, A. (1993). The intentionality of animal action. In M. Davies & G. Humphreys, eds. *Consciousness*. Oxford: Blackwell.
- Heyes, C. and Galef, C. B., eds. (1996), *Social Learning in Animals: The Roots of Culture*. San Diego: Academic Press.
- Howard, J. (1988). Co-operation in the prisoner's dilemma. *Theory and Decision* 24: 203-213.
- Hunt, G., & Gray, R. (2003). Diversification and cumulative evolution in New Caledonian crow tool manufacture. *Proceedings of the Royal Society London B*. In press.
- Hurley, S. (1989). *Natural Reasons*. New York: Oxford University Press.
- Hurley, S. (1998), *Consciousness in Action*. Cambridge: Harvard University Press.
- Hurley, S. (2001). Perception and action: Alternative views. *Synthese*, 291, 3-40.
- Hurley, S. (2003a), Animal action in the space of reasons. *Mind and Language* 18(3): 231-256.
- Hurley, S. (2003b), The limits of individualism are not the limits of rationality. *Behavioral and Brain Sciences* 26:
- Hurley, S. (2004). The shared circuits hypothesis: A unified functional architecture for control, imitation, and simulation. In S. Hurley & N. Chater (Eds.), *Perspectives on Imitation: From Neuroscience to Social Science* (Vol. 1, pp. 000). Cambridge, MA: MIT Press.
- Hurley, S. (forthcoming). Rationality, cooperation, and mind-reading. In N. Gold (Ed.), *Teamwork*. Palgrave.

- Hurley, S., & Chater, N. (2004a). *Perspectives on Imitation: From Neuroscience to Social Science*. Cambridge, MA: MIT Press.
- Hurley, S., and Chater, N. (2004b), The importance of imitation. In S. Hurley and N. Chater (Eds) *Perspectives on Imitation: From Neuroscience to Social Science*. 2 volumes. Cambridge, MIT Press.
- Hurley, S. and Noë, A. (2003). Neural Plasticity and Consciousness. *Biology and Philosophy* 18:131-168.
- Iacoboni, M. (2004). Understanding others: imitation, language, empathy. In S. Hurley & N. Chater (Eds.), *Perspectives on Imitation: From Neuroscience to Social Science* (Vol. 1, pp. 000). Cambridge, MA: MIT Press.
- Iacoboni, M., Woods, R., Brass, M., Bekkering, H., Mazziotta, J., & Rizzolatti, G. (1999). Cortical mechanisms of human imitation. *Science*, 286, 2526-2528.
- James, W. (1890). *Principles of psychology*. New York: Holt.
- Jeannerod, M. (1997). *The Cognitive Neuroscience of Action*. Oxford: Blackwell.
- Jeannerod, M. (2001). Neural simulation of action: A unifying mechanism for motor cognition. *Neuroimage*, 14, S103-S109.
- Kinsbourne, M. (2004). Imitation as entrainment: Brain mechanisms and social consequences. In S. Hurley & N. Chater (Eds.), *Perspectives on Imitation: From Neuroscience to Social Science* (Vol. 2, pp. 000). Cambridge, MA: MIT Press.
- Lhermitte, F. (1983). 'Utilization behaviour' and its relation to lesions of the frontal lobes. *Brain*, 106, 237-255.
- Lhermitte, F. (1986). Human autonomy and the frontal lobes, Part II. *Annals of Neurology* 19(4), 335-343.
- Lhermitte, F., Pillon, B., & Serdaru, M. (1986). Human autonomy and the frontal lobes, Part I. *Annals of Neurology* 19(4), 326-34.
- McDowell, J. (1994), 'The content of perceptual experience'. *Philosophical Quarterly* 44 (175): 190-205.
- Meltzoff, A. (1988), 'Infant imitation after a 1-week delay: Long-term memory for novel acts and multiple stimuli'. *Developmental Psychology* 24, 470-76.
- Meltzoff, A. (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 D. Cicchetti & M. Beeghly (Eds.), *The self in transition: Infancy to childhood* (pp. 139-164). Chicago: University of Chicago Press.

Meltzoff, A. (1995). Understanding of the intentions of others: Re-enactment of intended acts by 18-month-old children. *Developmental Psychology*, 31, 838-850.

Meltzoff, A. (1996). The human infant as imitative generalist: A 20-year progress report on infant imitation with implications for comparative psychology. In B. Galef & C. Heyes (Eds.), *Social Learning in animals: The roots of culture* (pp. 347-370). New York: Academic Press.

Meltzoff, A. (1999). Origins of theory of mind, cognition, and communication. *Journal of Communication Disorders*, 32, 251-269.

Meltzoff, A. (2002a). Elements of a developmental theory of imitation. In: A. Meltzoff & W. Prinz (Eds.), *The imitative mind: development, evolution and brain bases* (pp. 19-41). Cambridge, UK: Cambridge University Press.

Meltzoff, A. (2002b). Imitation as a mechanism of social cognition: Origins of empathy, theory of mind, and the representation of action. In U. Goswami (Ed.), *Handbook of childhood cognitive development* (pp. 6-25). Oxford: Blackwell.

Meltzoff, A. (2004). Imitation and other minds: The "like me" hypothesis. In S. Hurley & N. Chater (Eds.), *Perspectives on Imitation: From Neuroscience to Social Science* (Vol. 2, pp. 000). Cambridge, MA: MIT Press.

Meltzoff, A., & Decety, J. (2003). What imitation tells us about social cognition: A rapprochement between developmental psychology and cognitive neuroscience. *Philosophical Transactions of the Royal Society B*, 358, 491-500.

Meltzoff, A. & Moore, M. (1977). Imitation of facial and manual gestures by human neonates. *Science*, 198, 75-78.

Meltzoff, A., & Moore, M. (1979). Note responding to Anisfeld, Masters, and Jacobson and Kagan's comments on Meltzoff and Moore (1977). *Science*, 205, 217-219.

Meltzoff, A., & Moore, M. (1983a). Imitation of facial and manual gestures by human neonates. *Science*, 198, 75-78.

Meltzoff, A. & Moore, M. (1983b) Newborn infants imitate adult facial gestures. *Child Development*, 54, 702-709.

Meltzoff, A., & Moore, M. (1983c). The origins of imitation in infancy: Paradigm, phenomena, and theories. In L. Lipsitt & C. Rovee-Collier (Eds.), *Advances in infancy research* (Vol. 2 pp. 265-301). Norwood, NJ: Ablex.

Meltzoff, A., & Moore, M. (1989). Imitation in newborn infants: Exploring the range of gestures imitated and the underlying mechanisms. *Developmental Psychology*, 25, 954-962.

Meltzoff, A., & Moore, M. (1992). Early imitation within a functional framework: The importance of person identity, movement, and development. *Infant Behavior and Development*, 15, 479-505.

Meltzoff, A., & Moore, M. (1994). Imitation, memory, and the representation of persons. *Infant Behavior and Development*, 17, 83-99.

Meltzoff, A., & Moore, M. (1995). Infants' understanding of people and things: From body imitation to folk psychology. In J. Bermúdez, A. Marcel & N. Eilan (Eds.) *Body and the self* (pp. 43-69). Cambridge, MA: MIT Press.

Meltzoff, A., & Moore, M. (1997). Explaining facial imitation: A theoretical model. *Early Development and Parenting*, 6, 179-192.

Meltzoff, A., & Moore, M. (1998). Infant intersubjectivity: broadening the dialogue to include imitation, identity and intention. In S. Braten (Ed.), *Intersubjective communication and emotion in early ontogeny* (pp. 47-62). Cambridge, UK: Cambridge University Press.

Meltzoff, A., & Moore, M. (1999a). Persons and representations: why infant imitation is important for theories of human development. In J. Nadel & G. Butterworth (Eds.), *Imitation in infancy. Cambridge Studies in Cognitive and Perceptual Development* (pp. 9-35). Cambridge, UK: Cambridge University Press.

Meltzoff, A., & Moore, M. (1999b). Resolving the debate about early imitation. In A. Slater & D. Muir (Eds.), *The Blackwell reader in development psychology* (pp. 151-155). Malden, MA: Blackwell.

Meltzoff, A., & Moore, M. (2000). Resolving the debate about early imitation. In D. Muir (Ed.), *Infant development: The essential readings* (pp. 167-181). Malden, MA: Blackwell.

Meltzoff, A.N., & Prinz, W., eds. (2001). *The imitative mind: Development, evolution, and brain bases*. Cambridge, MA: Cambridge University Press.

Milgram, S. (1963). Behavioral study of obedience. *Journal of Abnormal and Social Psychology*, 67, 371-378.

Millikan, R. (2004). Comments on Goldman by Ruth Garrett Millikan: Some reflections on the simulation theory - theory theory debate. In S. Hurley & N. Chater (Eds.), *Perspectives on Imitation: From Neuroscience to Social Science* (Vol. 2, pp. 000). Cambridge, MA: MIT Press.

Millikan, R. (forthcoming). Styles of rationality. In M. Nudds and S. Hurley, eds. *Rational Animals?* Oxford: Oxford University Press.

- Milner, A. D. & Goodale, M. (1995). *The Visual Brain in Action*. Oxford: Oxford University Press.
- Nadel, J., & Butterworth, G. (1999). *Imitation in infancy*. Cambridge, UK: Cambridge University Press.
- 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, 174-186.
- Nehaniv, C., & Dautenhahn, K. (2002a). The correspondence problem. In K. Dautenhahn & C. Nehaniv (Eds.), *Imitation in animals and artifacts*. (pp. 42-61). Boston, MA: MIT Press.
- Nudds, M. & Hurley, S., eds. (forthcoming). *Rational Animals?* Oxford: Oxford University Press.
- O'Regan, K., and Noë, A. (2001a), "A sensorimotor account of vision and visual consciousness", *Behavioral and Brain Sciences* 24(5), 883-917.
- O'Regan, K., and Noë, A. (2001b), "Acting out our sensory experience", *Behavioral and Brain Sciences* 24(5), 955-975.
- O'Regan, K., and Noë, A. (2001c), "What it is like to see: a sensorimotor theory of perceptual experience", *Synthese* 129, 1: 79-103:
- Oztop, E. & Arbib, M (2002). Schema design and implementation of the grasp-related mirror neuron system. *Biological Cybernetics* 87:116-140.
- 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.
- Pepperberg, I. (1999). *The Alex studies: Cognitive and communicative studies on grey parrots*. Cambridge, MA: Harvard University Press.
- Pepperberg, I. (2002). Allospecific referential speech acquisition in Grey parrots (*psittacus erithacus*): evidence for multiple levels of avian vocal imitation. In K. Dautenhahn & C. Nehaniv (Eds.), *Imitation in animals and artifacts* (pp. 109-131). Cambridge, MA: MIT Press.
- Pepperberg, I. (2004). Insights into vocal imitation in Grey parrots (*Psittacus erithacus*). In S. Hurley & N. Chater (Eds.), *Perspectives on Imitation: From Neuroscience to Social Science* (Vol. 1, pp. 000). Cambridge, MA: MIT Press.
- Peterson, G. and Trapold, M. (1982), "Expectancy mediation of concurrent conditional discriminations", *American Journal of Psychology* 95(4):571-80.

Povinelli, D. (1996). Chimpanzee theory of mind? In P. Carruthers and P. Smith (Eds), *Theories of Theory of Mind*. Cambridge, UK: Cambridge University Press, 293-329.

Povinelli, D. (forthcoming). We don't need a microscope to explore the chimpanzee's mind. In M. Nudds and S. Hurley (Eds), *Rational Animals?* Oxford: Oxford University Press.

Povinelli, D., & Giambrone, S. (1999). Inferring other minds: Failure of the argument by analogy. *Philosophical Topics*, 27, 167-201.

Preston, S. & de Waal, F. (in press). Empathy: Its ultimate and proximate bases. *Behavioral and Brain Sciences*.

Prinz, J. (2004). Imitation and moral development. In S. Hurley & N. Chater (Eds.), *Perspectives on Imitation: From Neuroscience to Social Science* (Vol. 2, pp. 000). Cambridge, MA: MIT Press.

Prinz, W. (1984). Modes of linkage between perception and action. In W. Prinz & A. F. Sanders (Eds.), *Cognition and motor processes* (pp. 185-93). Berlin & Heidelberg: Springer.

Prinz, W. (1987). Ideomotor action. In H. Heuer & A. Sanders (Eds.), *Perspectives on perception and action* (pp. 47-76). Hillsdale, NJ: Erlbaum.

Prinz, W. (1990). A common-coding approach to perception and action. In O. Neumann & W. Prinz (Eds.), *Relationships between perception and action: Current approaches* (pp. 167-201). Berlin & New York: Springer.

Prinz, W. (1997). Perception and action planning. *European Journal of Cognitive Psychology*, 9, 129-154.

Prinz, W. (2002). Experimental approaches to imitation. In A. Meltzoff & W. Prinz (Eds.), *The imitative mind: Development, evolution, and brain bases* (pp. 143-162). Cambridge, UK: Cambridge University Press.

Prinz, W. (2004a). An ideomotor approach to imitation. In S. Hurley & N. Chater (Eds.), *Perspectives on Imitation: From Neuroscience to Social Science* (Vol. 1, pp. 000). Cambridge, MA: MIT Press.

Proust, J. (forthcoming). Thinking of oneself as the same. *Consciousness and Cognition*.

Regan, D. (1980). *Utilitarianism and Co-operation*. Oxford: Clarendon Press.

Rizzolatti, G. (2004). The mirror neuron system and imitation. In S. Hurley & N. Chater (Eds.), *Perspectives on Imitation: From Neuroscience to Social Science* (Vol. 1, pp. 000). Cambridge, MA: MIT Press.

Rizzolatti, G., & Arbib, M. (1998). "Language within our grasp." *Trends in Neuroscience*, 21, 188-194.

Rizzolatti, G., & Arbib, M. (1999). From grasping to speech: imitation might provide a missing link: Reply. *Trends in Neuroscience*, 22, 152.

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.

Rizzolatti, G., Fadiga, L., Fogassi, L., & Gallese, V. (2002). From mirror neurons to imitation: facts and speculations. In A. Meltzoff & W. Prinz (Eds.), *The imitative mind: development, evolution, and brain bases* (pp. 247-266). Cambridge, UK: Cambridge University Press.

Stamenov, M. & Gallese, V., eds. (2002). *Mirror Neurons and the Evolution of Brain and Language*. Amsterdam/Philadelphia: John Benjamins.

Strawson, P. (1959). *Individuals*. London: Methuen.

Strawson, P. (1959). *Individuals*. London: Methuen.

Strawson, P. (1966), *The Bounds of Sense*. London: Methuen.

Thorndike, E. (1898). Animal intelligence: An experimental study of the associative process in animals. *Psychological Review and Monograph*, 2, 551-553.

Tomasello, M. (1996). Do apes ape? In C. Heyes & B. Galef (Eds.), *Social learning in animals: the roots of culture* (pp. 319-346). San Diego: Academic Press.

Tomasello, M. (1998). Emulation learning and cultural learning. *Behavioral and Brain Sciences*, 21, 703-704.

Tomasello, M. (1999), *The Cultural Origins of Human Cognition*. Cambridge: Harvard University Press.

Tomasello, M., & Call, J. (1997). *Primate Cognition*. Oxford: Oxford University Press.

Tomasello, M. & Call, J. (forthcoming). Do chimpanzees know what others see – or only what they are looking at? In M. Nudds and S. Hurley (Eds.), *Rational Animals?* Oxford: Oxford University Press.

Tomasello, M., & Carpenter, M. (2004). Intention-reading and imitative learning. In S. Hurley & N. Chater (Eds.), *Perspectives on Imitation: From Neuroscience to Social Science* (Vol. 2, pp. 000). Cambridge, MA: MIT Press.

- Tomasello, M., Kruger, A., & Ratner, H. (1993). Cultural learning. *Behavioral and Brain Sciences*, 16, 495-552.
- Voelkl, B., and Huber, L. (2000), True imitation in marmosets. *Animal Behaviour* 60:195-202.
- Want, S., & Harris, P. (2001). Learning from other people's mistakes: causal understanding in learning to use a tool. *Child Development*, 72, 431-443.
- Whiten, A. (1996). When does smart behaviour-reading become mind-reading? In P. Carruthers and P. Smith (Eds), *Theories of Theory of Mind*, 277-292.
- Whiten, A. and Byrne, R., eds. (1997), *Machiavellian intelligence II: extensions and evaluations*. Cambridge, UK: Cambridge University Press.
- 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.
- Whiten, A., Horner, V., & Marshall-Pescini, S. (2004). Selective imitation in child and chimpanzee: A window on the construal of others' actions. In S. Hurley & N. Chater (Eds.), *Perspectives on Imitation: From Neuroscience to Social Science* (Vol. 1, pp. 000). Cambridge, MA: MIT Press.
- Williams, J., Whiten, A., Suddendorf, T., & Perrett, D. (2001). Imitation, mirror neurons and autism. *Neuroscience and Biobehavioral Reviews*, 25, 287-295.
- Wohlschläger, A., Gattis, M., Bekkering, H. (in press) Action generation and action perception in imitation: An instantiation of the ideomotor principle. *The Philosophical Transactions of the Royal Society*.
- Wolpert, D. (1997) Computational approaches to motor control. *Trends in Cognitive Sciences*, 1, 209-216.
- Wolpert, D., & Kawato, M. (1998). Multiple paired forward and inverse models for motor control. *Neural Networks*, 11, 1317-1329.
- 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, B*, 358, 593-602.
- Wolpert, D., Ghahramani, Z., & Jordan, M. (1995). An internal model for sensorimotor integration. *Science*, 269, 1880-1882.
- Zentall, T. (2001). Imitation and other forms of social learning in animals: evidence, function, and mechanisms. *Cybernetics and Systems*, 32, 53-96.

