

8 Commentary and Discussion on Mechanisms of Imitation

8.1 Reflections on Mirror Systems

J. N. P. Rawlins on Rizzolatti and on Decety and Chaminade

8.1.1 Imitation

Imitation requires a mapping between one's own behavior and the behavior of some other or others. This mapping may be necessary but not sufficient for a behavior to count as imitative, since definitions of imitation vary. Views differ on how precise the mapping and how novel the resulting behavior must be in order to count as true imitation. For example, Rizzolatti uses Thorndike's definition, "learning to do an act from seeing it done." Decety uses a more restrictive definition, which specifies that imitation must be intentional and effortful; moreover, the behavior that results must be novel and must share the goal of the observed behavior and use similar means to attain it. My purpose here is not to argue for one definition or another. Rather, I wish to consider the neural circuitry activated when monkeys or humans engage in action or observe the same action in others—the conditions under which this circuitry is effectively activated and the psychological functions it might serve—and to propose some new experimental approaches to these issues.

8.1.2 Why Should Monkeys Have Mirror Neurons?

The discovery of mirror neurons in the brains of monkeys demonstrates the existence of a link between vision and action. Mirror neurons are as powerfully activated when a monkey observes some action being performed by another individual as they are when the observer itself engages in the same action. But what are monkey mirror neurons for?

There are many examples of the behavior of one individual conforming to the behavior of another. Flocks of birds move together; people in conversation tend to adopt similar postures. These kinds of behavioral

conformity can be observed informally, with no need for sophisticated experiment, and are widely attributed to response facilitation or priming. There is no need to assume that they must depend upon understanding the behavior of others, or result from some intentional or effortful component. However, they do require a mechanism by which a sensory input—typically a visual one—can appropriately drive the motor system. Could monkey mirror neurons provide such a mechanism?

Mirror neurons in monkey area F5 are activated by object-related hand or mouth movements; precision grip is also represented there. Such a system could have a role during development if infant monkeys learn to forage by observing what other foraging monkeys pick up and eat, and then doing the same themselves. I do not know if there is evidence for this kind of learning. However, as Rizzolatti indicates, for such a system to increase foraging efficiency, it would have to represent the right kind of objects to pick up and ingest. It would be counterproductive to copy hand and mouth movements without relating them to the kinds of objects that the observed actors actually manipulate.

In fact, in monkeys there is no evidence of such copying of mere movements, unrelated to goals (see Voelkl & Huber, 2000, on goal-related imitation of movements by marmosets). Moreover, there is reason to believe that monkey mirror neurons may respond only to goal-directed actions. For example, Rizzolatti describes mirror neurons that require the observer monkey to have seen the target object of the observed movement, even if the object is then hidden behind a screen. Under these conditions, a movement that does elicit a mirror response may be identical to a movement that does not, if the observer monkey has not seen the target object placed behind the screen. To activate such mirror neurons, the observer monkey must detect not just a movement but also a specific goal; so simple response priming does not appear to be the function of these neurons.

It may come as a relief to dancing instructors that the human mirror system can be entrained by nongoal-directed, intransitive movements as well as by goal-directed movements. Human beings are enthusiastic imitators, in multiple senses of the term. But as Decety convincingly argues, mirror systems offer neural mechanisms that could underpin a range of distinctively human processes, beyond imitation. If area F5, where mirror neurons were first observed, is really the monkey homologue of the human Broca's area, which plays a crucial role in our linguistic ability, then perhaps a system that originally conferred quite different evolutionary benefits

now subserves our effortless acquisition of language.¹ Still, monkeys do not acquire language, so what advantages would a mirror system have for a monkey?

8.1.3 Are Monkeys Beginner Mind Readers? Is Mind Reading an Exclusively Human Preserve?

Rizzolatti attributes to the monkey mirror system a role in enabling monkeys to begin to understand the actions of others. Seeing the actions of another activates one's own mirroring motor system and thus allows the other's behavior to be matched to one's own repertoire and its consequences predicted, just as the consequences of one's own motor activity might be predicted by a forward model. However, there is at present no experimental evidence for such proto-"mind reading" in monkeys either.

The following experimental approach may yield fresh evidence on this issue. Single-unit recording experiments have found *place cells* in the rat hippocampus, which are active when the rat is located in particular places. Place cells in the monkey hippocampus have also been identified, but these are active when the monkey simply looks at a particular place; it need not be physically located at that place. Consider an experiment with two monkeys. Monkey A has a place cell that fires when the monkey looks at place x. We now arrange that monkey A can see monkey B, and that monkey B but not monkey A can see place x. Under these conditions, if place cells for place x are activated in monkey A's hippocampus just when monkey B is looking at place x and monkey A is watching monkey B, that would suggest that monkey A has inferred the place that monkey B is looking at. Such a result would seem to be evidence of elementary mind reading by monkeys. If such a result were obtained, its relationship to the monkey mirror system would be worth investigating.

Decety suggests that mind reading is a uniquely human ability. I know of no convincing counterevidence for this claim, but I would nonetheless like to speculate about the possibility that similar abilities could have evolved in other social animals. One particularly interesting group to consider might be social cetaceans, which use echolocation. Whereas primates generally have to be looking in the right direction to see what others in their social group are attending to, there is surely no such constraint in

1. In acquiring language, we effortlessly copy verbalizations we hear. If imitation must be effortful, as Decety requires, how would this effortless copying be better described?

whales. If one individual is using active sonar to interrogate some object of interest, the return echo will presumably be available to all the other individuals within range. Each individual is in a sense immersed in the current preoccupations of all the other animals in the group. This offers a wonderful potential handle for investigating the evolution and range of mind-reading abilities.

8.1.4 The Human Mirror System: Is Imitation Its Key Function?

Under what conditions are human mirror systems activated? Does this activity function primarily to drive imitation? Given the constraints on single-unit recording in the human brain, much of what we know is derived from functional imaging of regional blood flow changes. Although evidence from other methodologies is also available, none allows us to characterize the range of drivers to which particular mirror neurons may respond. Our conclusions are therefore derived from general changes in the activity of brain regions. We cannot at present know whether there are neurons within those regions whose activities are quite different from the group as a whole. Despite this limitation, some striking findings have been obtained, such as the finding Decety reports of a hemispheric difference in parietal activation between imitating and being imitated.

Functional imaging studies all require subtraction methods. The experimental condition is contrasted with a comparison condition intended to be identical in all but the crucial variable of interest. Just the same kind of formal relationship between experimental and control conditions exists in classical conditioning designs for the study of associative learning. When Pavlov's dogs learned that the sound of a bell was followed by the delivery of food, they started to salivate in response to the bell. An innate response to the delivery of food was now elicited by the *signal* of food. But how can we be certain that this new response really depends on having detected and learned the association that we so carefully arranged between the sound of the bell and the delivery of food? In Pavlovian conditioning we arrange a correlation between a signal and an outcome, and assume that the conditioned response to the signal that develops does so because that relationship has been learned. If we had arranged that the signal and the outcome were uncorrelated, and the same response to the signal had developed, then this could not be because the animal had learned the correlation. We would therefore assume that some mechanism other than Pavlovian conditioning was responsible for the change in behavior. Control procedures for associative learning experiments are designed to test the possibility that an apparently conditioned response in fact arises via some other, non-

associative route. For example, does the bell come to elicit drooling even if the bell did not reliably predict food?

There are two ways of settling that question. One is to arrange a negatively correlated comparison condition. In this case, food is never presented when the bell has just sounded. Animals exposed to this kind of contingency typically differ in their responses from those exposed to Pavlov's positively correlated condition. This control is not usually the preferred comparison, though, because it has become clear that animals learn the negative correlation—they expect that food will *not* be delivered when the bell sounds. Any difference in behavior between the positively correlated and negatively correlated groups might therefore stem from this inhibitory learning, rather than from the excitatory learning in the positively correlated group. As a result, the more generally accepted comparison condition is a “truly random” control condition, in which the bell and the food are each presented from time to time, but food is no more likely to be delivered after the bell sounds than when the bell has not sounded.²

The functional imaging paradigms used to study imitation and mirror systems have typically used comparison conditions that entail observation alone, or observation of an unrelated action. Imitation leads to more activation than either of these conditions. But just as in classical conditioning, unrelated comparison conditions may differ in important ways from negatively correlated ones. Moreover, in studying mirror systems, the negatively correlated control condition is, I suggest, more analytically informative than the unrelated comparison condition typically used.

If mirror systems are important for imitation as such, then they should not be activated when one must do the opposite of what the actor does—indeed, one might predict deactivation. If, on the other hand, they play a key role in understanding the actions of others, then they should be at least as powerfully activated in a paradigm in which an observer needs to do the opposite of an actor as in the standard imitation paradigm in which the observer needs to do the same as an actor. Imagine that we watch an actor assemble a complex puzzle that requires a strict sequence for assembly and disassembly. If we know that our job is going to be to assemble the puzzle ourselves, we clearly need to remember and subsequently reproduce the actor's sequence of actions. That will presumably powerfully activate the mirror system. But what if we know that our job will be to *disassemble*

2. This comparison also has a potential drawback because there is reason to believe that animals may actively learn the lack of a relationship between the two stimuli—a learned irrelevance paradigm.

the construction? We now need not only to observe and identify what the actor does, but also to plan for ourselves the reversed sequence of actions we will need to use in order to attain our goal of achieving the status quo ante. This would not be imitation in Decety's sense, because it has a different goal and a different sequence of actions. Nonetheless, it is undoubtedly very different from seeing someone carry out actions that are unrelated to one's own task. I propose that such an experiment would be a more rigorous way to determine whether the mirror system functions primarily to drive imitation or the understanding of actions.

8.2 Action Recognition, Imitation, and Language Are Different

Michael Arbib on Iacoboni

Marco Iacoboni's chapter shows how study of the monkey mirror system has inspired a body of excellent work on human imitation using functional magnetic resonance imaging (fMRI) and transcranial magnetic stimulation (TMS). Since space is limited, I will just comment on section 2.3, which focuses on the key issue of the relationship between the action recognition system and the emergence of language. This will raise broader issues concerning the status of linguistic theory and the brain mechanisms that underlie the evolution of language.

Let us begin with the specific. In chapter 2 Iacoboni summarizes a previous discussion by myself and Rizzolatti as follows:

How can such a formalized structure [of a parsing tree for grammar] emerge from a relatively primitive action recognition system? A type of answer ... (Rizzolatti and Arbib, 1998) [henceforth R&A] ... is that "gestures may be a primitive form of grammar." The problem with both question and answer is that they accept a view of language as a phenomenon that can be essentially reduced to formal constructs such as grammar. (Iacoboni, vol. 1, ch. 2, p. 93)

However, the apparently quoted phrase, "gestures may be a primitive form of grammar" occurs nowhere in R&A! Instead, these authors:

examine whether or not a "prelinguistic grammar" can be assigned to the control and observation of actions. If this is so, the notion that evolution could yield a language system "atop" of the action system becomes much more plausible. (Rizzolatti & Arbib, 1998, p. 191)

I can see why some readers may have mistaken talk of a "prelinguistic grammar" for the claim that "gestures may be a primitive form of grammar," but (1) R&A's approach is semantic rather than syntactic, and (2) emphasizes differences as well as commonalities. R&A stated:

We view the activity of "canonical" F5 neurons as part of the code for an imperative case structure, for example,

Command: grasp-A(raisin)

as an instance of grasp-A(object), where grasp-A is a specific kind of grasp, to be applied to the raisin. *Note that this case structure is an "action description," not a linguistic representation...* [Again,] we might say that the firing of "mirror" F5 neurons is part of the code for a declarative case structure, for example,

Declaration: grasp-A(Luigi, raisin)

which is a special case of grasp-A(agent, object), where grasp-A is a specific kind of grasp, applied to the raisin (the object) by Luigi (the agent). *Again, this is an "action description," not a linguistic representation.* (Rizzolatti & Arbib, 1998, p. 192; italics added.)

In the case analysis by Fillmore (1966), the sentence "John hit Mary with his hand" is viewed as the "surface structure" for a case structure "hit (John, Mary, John's hand)," which is an instance of the case frame "hit (agent, recipient, instrument)," which makes explicit the roles of "John," "Mary," and "John's hand." However, being able to grasp a raisin is different from being able to say "I am grasping a raisin," and R&A are clear that the neural mechanisms that underlie the doing and the saying are different. However, the case structure lets us see a commonality in the underlying representations, thus helping us understand how a mirror system for grasping might provide an evolutionary core for the development of brain mechanisms that support language.

The key point of language is that it can provide sentence structures that can describe very different actions (and much, much more). The "case structure" of an animal's action is a human description of a limited capability; the "case structure" of a sentence exemplifies a human linguistic "frame" that the human can use to describe freely novel situations that have never arisen before.

I agree with Iacoboni that one must "factor in the changes that the evolutionary process might have produced," but am concerned when he says that even though "the imitative abilities of monkeys are limited, ... one can also conceivably argue that the action recognition system made monkeys 'imitation-ready'" (vol. 1, ch. 2, pp. 81–82). In Arbib (2002), I introduced the term "language-ready" to indicate that early *Homo sapiens* might well have had brains like ours and yet not have had language—the brain was ready to learn language but the culture had not yet produced language to learn. However, all the evidence says that monkeys cannot learn to imitate

in any major sense, no matter what opportunities are provided by the social and physical environment, while chimpanzees exhibit imitation, but in a form limited with respect to the human's. I thus find Iacoboni's use of "imitation-ready" to be unfortunate, and suggest that what I would call an imitation-ready brain (i.e., one that could master imitation given the right environment) did not evolve until well into the hominid line (or at least millions of years after the common ancestor of monkeys and great apes, if "simple" imitation is all that is of interest). Indeed, I hypothesize (Arbib, 2002, in press) that recognition of manual actions, imitation, and the ability to acquire and use language rest on a nested, evolutionary progression of brain mechanisms. I take seriously our ability to produce myriad novel sentences, seeing the openness of language as both grounding for and emerging from the ability to translate between cognitive structures and verbal structures within the framework of social communication.

In section 2.3.3 Iacoboni presents a transcription of "everyday speech" (figure 2.7) and stresses that

not only do violations of grammar ... occur ..., but ... this segment of speech is full of phenomena that are ... not ... studied by traditional linguists.

A salient feature of typical conversations that is ignored by traditional linguists is *turn-taking* [which depends on] ... the hearer's tracking ... [of] transiently appearing *opportunities for taking a turn—which are not exclusively grammatical* but rhythmic and pitch contour-intonational as well. [vol. 1, ch. 2, pp. 93–94; my italics.]

Iacoboni is right to emphasize the importance of turn-taking, but I think that he is wrong to be so dismissive of the work of traditional linguists (who do include phonologists and not just grammarians, after all). In his figure 2.7, speech is not only labeled with "?" for "stuff" but also with syntactic markers for what the speaker and the hearer would recognize as the "real" message. Just as Newton made progress by first treating the sun, Earth, and moon as points, and then others developed the study of the tidal effects of the moon, so it makes sense to start with a model of the "perfect" speaker and then seek to understand why limitations of memory, delays in production, and changing rhetorical goals (competition of different thoughts for expression, new thoughts as one speaks, etc.) produce imperfect utterances. "Ums" and "ahs" are imperfections of language production, not its essence. I do take the point that the understanding of such elements will come from motoric concepts (such as signaling of delays in reaching a communicative goal) but still suggest that the motoric analysis should be complemented by a more idealized production model that is closer to traditional linguistics. However, when Iacoboni (vol. 1, ch. 2,

pp. 94–95) urges us to "consider carefully the incontrovertibly motor elements that are at work in conversation [whose processing] ... requires a fast functional architecture that is not dissimilar to the one needed in motor control," my concern is twofold. (1) I do not know of any studies in monkeys that relate neural processes to turn-taking. (2) Most species have a fast functional architecture for motor control, but only humans have language, so that the last statement ignores the very differences whose bridging across evolutionary time one needs to explain.

8.3 Evidence for Memetic Drive?

Susan Blackmore on Iacoboni

I was thrilled when I learned of Iacoboni's discovery that when a chimpanzee's brain is morphed onto a human brain, the areas of greatest expansion are those that are used in imitation. "Yes!" I thought. "This is exactly what I predicted on the basis of memetic theory. Whoopee—memetics is right!" But then I had to pause, because this is how to make the worst mistake in the book. Construct a wacky theory, derive a prediction from that theory, discover the prediction is correct and then (illegitimately) conclude that the theory must be true. So I would like to describe the prediction and consider whether these findings do have any implications for memetics.

Dawkins's (1976/1989) original idea in coining the term "meme" was to point out that when people imitate each other, they not only copy information, they must select what to copy, and their copies are not perfect. This is all that is required to apply the principles of universal Darwinism and, by definition, the information people copy is a replicator. Dawkins called that new replicator the meme.

One implication of the theory of memetics is that the capacity for imitation must inevitably let loose a new evolutionary process and, as Dawkins originally put it, "Once this new evolution begins, it will in no necessary sense be subservient to the old" (Dawkins, 1976/1989, pp. 193–194). He criticized his colleagues because "In the last analysis they wish always to go back to 'biological advantage'" to answer questions about human behavior (Dawkins, 1976/1989, p. 193). But if memes are replicators, then we must consider memetic advantage too. Humans must be the product of two replicators, not just one, and this should be obvious in the way they have evolved.

In exploring the implications of memetic advantage, I hypothesized that the interests of the memes might force the genes to take a direction

different from that which they would have taken otherwise; they would be forced to follow the direction taken by memetic evolution. This is the coevolutionary process I called *memetic drive* (Blackmore, 1999, 2001).

Put simply, the hypothesis is this. Once human ancestors could imitate, memes appeared and began competing to be copied, their success depending on the type of meme and the preferences and abilities of the people doing the copying. Given that at least some of the memes would provide survival benefits, this means an advantage to genes for the ability to copy those memes. If better imitation requires a bigger brain, then this process alone will tend to increase brain size and improve the ability to imitate. As this ability increases, more memes will appear and their evolution will take off in various directions, perhaps including the creation of rituals, clothes, body decoration, or music, including behaviors that are of more advantage to those memes themselves than to the genes of the people copying them. If being able to display the latest memes provides status (which is a reasonable assumption), then it will pay everyone to copy the best imitators, and to mate with them. Either way, this creates an advantage for genes for the ability to copy the latest memes. In this way genes would be expected to track the direction taken by purely memetic evolution and thus we humans have ended up with brains that are not only much larger, but are specially designed to be good at music, ritual, art and, of course, language.

This hypothesis allows for some (admittedly rather general) predictions. In particular, if brain size has been meme driven, then within groups of similar species brain size should correlate with the ability to imitate. Of course there are few species capable of imitation, but this prediction holds for humming birds (Jarvis et al., 2000). Other aspects of the big brain hypothesis have been confirmed using simulations and mathematical modeling (Bull et al., 2000; Higgs, 2000). More specifically, I predicted that brain scans of people either initiating or imitating actions should reveal that "imitation is the harder part—and also that the evolutionarily newer parts of the brain should be especially implicated in carrying it out" (Blackmore, 2000b, p. 73). This implies that the parts of the brain that differ most between chimpanzees and humans should be those involved in imitation (assuming that present-day chimpanzees are closer to our common ancestor than humans are). Finally, if memetic drive is responsible for the evolution of language, then we should expect the language areas in the human brain to be derived from areas originally used for imitation.

This is what Iacoboni and his colleagues have demonstrated, thus confirming these predictions. In chapter 2 (vol. 1, p. 91) Iacoboni concludes

that "from a relatively simple neural mechanism of matching the observation and execution of an action (mirror neurons), more complex functional properties were built and more complex behaviors were supported." The question now is, Why? The standard evolutionary view must be that it happened in the interests of the genes. Like Dawkins's colleagues, most people will presumably "wish always to go back to 'biological advantage'." But the wider alternative remains; that when it comes to human evolution there may be more than one replicator competing for survival.

There is nothing mysterious about memetics. Memes are not mystical entities floating about in a few theorists' minds. They are nothing more nor less than whatever it is that people copy when they imitate. So if you admit that people (imperfectly and selectively) copy each other, and you define a replicator as information that is copied with variation and selection, then you have to conclude that memes exist. All the doubt must be about whether memetics can ever prove itself useful as a science, and whether memes really have played the crucial role in human evolution that memetic theory suggests. Iacoboni's findings fit perfectly with the predictions made, but then, as he discusses in his chapter, there are many possible explanations for them. Memetics has made a start, but it has a great deal further to go if it is to prove its worth in understanding human evolution.

8.4 The Role of Mirror Neurons in Imitation

Susan Jones on Gallese

The mirror neuron was second only to imitation itself as the hardest working concept at the conference at Royaumont. Some participants spoke of the mirror neuron as though, all by itself, it could be a mechanism for imitation. More specifically, some seemed to have the idea (too briefly and therefore crudely expressed here) that mirror neurons might be capable of directly converting observed behavior into executed behavior. If this were true, then the problem of finding a plausible common mechanism for different imitative acts would be as good as solved. But in my view, how imitation actually happens is not much clearer since the discovery of mirror neurons than it was before.

The burden of these comments is that the role of mirror neurons in producing imitative behavior is not likely to be as a means of converting visual input about behavior into motor output of a copy of that behavior. Instead, I would put my money on Gallese's vision of mirror neurons as sources of the experience of common experience with other people, animals, or robots like ourselves.

Gallese has first-hand experience of mirror neurons, and has probably thought about their possible functions more than most. He proposes that mirror neurons are part of a mechanism for the automatic, subpersonal, nonpropositional recognition and understanding of the actions of others. This recognition and understanding of just the nature of others' behaviors might then feed into an understanding of the intentions and states behind those actions—in other words, it might produce empathy—and other forms of intersubjectivity. Then, to the degree that these experiences and the knowledge they yield are prerequisites for the ability to imitate, mirror neurons would be part of a mechanism for imitation.

Gallese does not suggest that mirror neurons function directly or even primarily to produce imitation, except perhaps in newborn human infants. This seems like appropriate caution, given what Gallese and Rizzolatti tell us in this volume about their single-cell recordings from mirror neurons in the monkey premotor cortex—and also, given what they do *not* tell us. They do not tell us about any data linking mirror neurons to imitation. This does not mean that mirror neurons are not involved in imitation—only that we don't yet know whether they are and if so, how.

We are told instead about data from monkeys on the responsiveness of mirror neurons to both the sight and the production of the act of reaching for and grasping an object. Given that the data are about reaches and grasps, the question becomes how what we know about the production of these actions can inform, first, speculation about the role of mirror neurons in producing the grasping action; and second, speculation about how imitative behavior is produced.

The first thing to note is that these reaching and grasping actions in the monkeys (and in the experimenters) are not imitative; and the responsiveness of mirror neurons to these actions does not give rise to imitation. Specifically, Gallese does *not* tell us that when the mirror neuron fires in response to either the sight or the sound of the grasping action, the monkey moves its own hand. Mirror neurons, then, are not transducers. They respond to both sensory input and motor events; they do not respond *to* sensory input *with* motor events. They do not automatically convert visual or auditory input into a motor response (at least, not in the adult; more about this qualification in a moment).

The idea that mirror neurons link observed actions directly to stored "motor plans" also seems wrong. For example, there has been a lot of work done on targeted reaching and its development (see, e.g., K. Newell & Molenaar, 1998; Rosenbaum et al., 1996). The data indicate that even such an apparently simple behavior is actually mind-numbingly complex in ex-

ecution. That is because of the number of elements that have to be controlled when you (or Gallese's monkeys) reach toward an object. These include a host of different muscles (twenty-three on each side of your body can be involved in reaching) and different joints, acting from different starting points under highly variable conditions in the nervous system, in the muscles, in the forces produced by gravity interacting with posture, and so on. Because of the number of degrees of freedom in the conditions for reaching, it is thought that each reach is pretty much unique. The particular combination, sequence, and intensities of muscle movements in any one reach cannot be anticipated, even at the beginning of that particular reach. Instead, it seems most likely that reaches and other motor behaviors are dynamic and assembled online. That means that in every instance where someone reaches out and grasps an object, the action is assembled in that moment, is tailored to the contextual conditions of that moment, and is continuously monitored and adjusted even after it is launched. It is therefore very unlikely that the control of the huge number of different sequences that produce actual reaches is permanently represented in the central nervous system as "motor plans."

This is what reaching for objects looks like from the beginning (e.g., Thelen, 2001). Human infants do not reach at birth. In fact, they put in a lot of hard work before they succeed in reaching to grasp objects at about 5 months of age. Moreover, the developmental course of reaching is so different in different infants that it provides a strong argument against the idea of innate neural systems for reaching (Thelen, 2001; Thelen et al., 1993). Certainly, newborn infants do nothing to suggest that they come equipped with mirror neurons preprogrammed to play a role in reaching.

In short, the data so far suggest that any role played by mirror neurons in object-directed reaches and grasps is indirect and acquired in development. Mirror neurons are not transformers of visual input into motor output. And mirror neurons do not initiate the implementation of stored motor plans.

Our topic is not targeted reaching or grasping, but the much broader metabehavioral category of imitation. What, then, are the implications for the role of mirror neurons in enabling us to imitate any of a host of motor behaviors that we see, or hear, or feel, or imagine others performing? Gallese seems to be proposing two quite distinct roles—one for imitation in newborn infants and another for imitation beyond the newborn period.

Gallese suggests that perhaps imitation in newborn humans is the product of a transducerlike mirror neuron. Again, there are no data. We know nothing about the development of mirror neurons in monkeys, let alone in humans. This point is worth emphasizing. We do not know when or how

monkeys develop the mirror neurons responsive to object grasping. And we certainly do not know that human newborns have mirror neurons that respond during tongue protrusions, which is the most commonly matched behavior in studies of newborn imitation.

The appeal of the idea that the newborns have mirror neurons is obvious. It has been argued for decades that imitation in the neonatal period is achieved via active intermodal mapping (Meltzoff & Moore, 1997, 1983a)—that is, the matching of mental representations in different modalities. But how this mapping might be accomplished has never been plausibly spelled out. Gallese suggests that the match is achieved by mirror neurons. In previous accounts of imitation in newborns, the proposed match was between two kinds of sensory information—visual and proprioceptive representations of tongue protrusion, for example. Now the proposal seems to be that mirror neurons both respond to visual input of an observed action and also initiate a motor behavior that replicates that same action. This idea resembles the classic reflex loop much more than the mirror neuron as observed in Gallese's own experiments.

The data on behavioral matching in newborns do not fit the hypothesis all that well. First, newborn infants in imitation experiments do not reproduce the adult model's behavior in a one-to-one fashion, as might be expected of a "visual in—motor out" device. Typically, they produce fewer tongue behaviors than they see, and these are clustered irregularly throughout the trial. If mirror neurons are mediating between the visual input and the motor output, what is the source of these long delays?

Until the discovery of mirror neurons, much was made of the fact that behavioral matching by newborns in many experiments *did* occur after a delay, as the delay was thought to indicate something more interesting than an automatic response that might be a fixed action pattern. For example, in Meltzoff and Moore's 1977 study (which Meltzoff re-presented at Royaumont³), the infants had pacifiers in their mouths as they watched an adult do tongue protrusions at a rate of 4 in 20 seconds, then had 2.5 minutes each to produce tongue protrusions in the absence of a model. Their graph indicates that the infants produced tongue protrusions at a rate of about 0.5 in 20 seconds (40+ tongue protrusions/(12 infants × 2.5 minutes))/3 segments per minute. The same graph indicates that the infants produced less than 1/10 of a mouth opening for every 4 mouth openings produced by the adult (Meltzoff & Moore, 1977). In later experiments (e.g.,

3. The conference with the same name as this volume, held at Royaumont Abbey, France, in May 2002. ED.

Meltzoff & Moore, 1983a), a model typically would produce 4 tongue protrusions in a 20-second period, then assume a still face for another 20 seconds—and it was while the model was *not* tongue protruding that the infant's tongue protrusions were most numerous.

A second potential problem for the idea that mirror neurons are the intervening mechanism between the adult model and the imitating newborn is the fact that the one behavior that infants reliably produce in imitation experiments—tongue protrusions (Anisfeld, 1996)—is also produced with equal enthusiasm when the stimulus is not an adult model doing tongue protrusions, but is instead some other arousing stimulus such as flashing colored lights (Jones, 1996) or music (Jones, 2001). Clearly, these tongue protrusions match nothing and are not produced by mirror neurons. It is possible that infants in very similar circumstances produce tongue protrusions in response to different arousing stimuli via different mechanisms, but it is certainly not parsimonious. An alternative is that tongue protrusions by newborns in response to all of these stimuli, the sight of adult tongue protrusions included, are by-products of arousal, and that newborn imitation itself might be a chimera.

One further consideration. I have said why, to me, the evidence says that the role of mirror neurons in producing behavior is neither simple nor direct. The mechanism for producing behaviors such as directed reaches and object grasps—with which mirror neurons are empirically (as opposed to speculatively) related—is almost certainly not a bunch of mirror neurons acting like a big reflex loop with a "stop-go" function added at some point in development. This is also a very unlikely description of a mechanism for imitation. Thus, if behavioral matching by newborns *did* turn out to be the product of a simple, direct, reflexlike, behavior-in/behavior-out mechanism with mirror neurons in the middle, it would mean that behavioral matching by newborns is not mechanistically related to imitation beyond the newborn period. In other words, if mirror neurons *are* the mechanism underlying behavioral matching by newborns, then behavioral matching by newborns goes nowhere developmentally and is consequently less interesting than we thought.

I like Gallese's core idea, which is that mirror neurons are the seat of our experiences of identification and empathy. This idea captures the one thing about mirror neurons that makes their functions potentially different from the functions of two separate but closely interacting populations of cells. The one thing about mirror neurons that is different is that these cells can fire for a specific instance of a broader category of actions—say, for a particular object grasp—but not "know" whether the action was mine or

yours. What would such cells be good for if not to blur the lines between me and you and let us each know the other to be like ourself?

I am afraid, however, that this new idea is in danger of being lost in the rush to make the mirror neuron fit existing holes in our theories. We are such good synthesizers, and we get such aesthetic pleasure out of making the pieces fit, that we are understandably tempted to go far beyond the data. This has happened to a great extent with the mirror neuron. However, we should probably resist the temptation to assume that nature works as elegantly as it would if we designed it ourselves. Often it doesn't, and we are in the business of finding out what's true, not what would make a good and intellectually satisfying story.

8.5 Overlapping Brain States while Viewing and Doing

Marcel Kinsbourne on Decety and Chaminade

Decety and Chaminade introduce their broad goal as one of making "a tentative step toward a better understanding of intersubjectivity" (vol. 1, ch. 4, p. 119). With respect to the "intimate psychological relation between self and other (p. 139)," they arrive at the view that there is a unique human capacity to identify with others and thus share subjectivities. I focus here on the theoretical implications of some of the intriguing discoveries that have arisen from Decety's elegantly conceived research program.

8.5.1 Two-Way Traffic

According to Decety and Chaminade, "the discovery of mirror neurons ... has encouraged the search for a comparable mechanism in humans" (vol. 1, ch. 4, p. 124). Mirror neurons fire both when a particular action is perceived and when the animal itself executes this action. By "comparable mechanism," they refer to brain areas that activate on functional imaging both when an action is perceived and when it is performed. They do not claim that there are mirror neurons in these areas, and there is currently no way of knowing this. The statement is rhetorical, drawing an analogy between a cell population and a single cell type.

Distinct neural systems subserve knowledge of persons and of objects (Mitchell et al., 2002). Decety organizes his functional neuroimaging research in line with the presumed two-way traffic in the case of persons. He differentiates between knowing something about oneself and about others. He sets up comparable eliciting conditions, which differ in only one critical respect, along the lines of perceiving versus performing, self versus other. The conditions will generate a nearly identical activation profile in the

cortical manifold. Those regions of activation that do not overlap both conditions are presumed to code for the critical difference between the conditions. In one study he can, for example, attribute the self versus other distinction to a corresponding left versus right hemisphere engagement. The strategy works seamlessly. How then should we understand the analogy with mirror neurons? Does it refer to a real functional parallel or is it mere metaphor? Do mirror neurons tie percept and action together or are they merely indifferent to the distinction between them?

8.5.2 The Spell of Mirror Neurons

Why have mirror neurons so captivated the imagination of neuroscientists? Perhaps it is because they run counter to a seemingly irresistible trend to atomization in neural brain models. The contest between the lumpers and the splitters in cognitive neuroscience is unequal. Focal brain lesion effects reveal which functions are dissociable, but functions that do not dissociate very well might be shown to do so in the future, when the patient with the strategically located lesion turns up. So case-by-case, instances of differentiation of functions increasingly preponderate over communality of functions. In contrast, demonstrating what is indissoluble, or fieldlike, or "resonant" (Shepard, 1984) in the brain's functioning generally has to await technology that is only minimally within our current reach. So "assembly" models of brain function, a collage somehow pasted together ("integrated"), predominate. Mirror neurons offer a tangible integration of perceived and performed action and better still, they do so by means of an experimentally accessible specialized single cell type. The opportunity is hard to resist. Based on this microscopic edifice, theories of the brain mechanism of "theory of mind," the nature of autism, cultural advance, etc., are copiously proposed.

8.5.3 Shared Representations

Did mirror neurons expressly evolve to unite percepts with actions or intentions for adaptive advantage? The fact that they respond during both perceiving and performing does not prove that they are instrumental in coordinating these two domains. Perhaps they evolved, not to represent both perception and production, but to represent neither in particular. They might simply exist within a sphere upon which both perception and performance draw, namely, representing the action in question, leaving it to other circuitry to represent the perception-production distinction.

The notion that a single representation serves alternative mental states, rather than that separate representations are secondarily shared or united,

is strongly supported by the work of Decety and others that shows common coding of perception and action. Consider one of his designs in which three conditions are implemented: representing one's own action, another's action in imagination, and another's action directly observed. Instead of there being three distinct processes that activate distinct areas or even "modules," the findings suggest that the three conditions utilize the same core neural representation of the action in question.

If self- and nonself-related representations have the same neural activity in common, rather than a specifically evolved connection, this might make the neural intertwining seem incidental rather than functionally critical. Shared representations, and also mirror neurons, would only incidentally tie percepts and actions together, being indifferent to which is being experienced at the time. If this is so, it does not invalidate the generally held belief that the communality of perception and action subserves the adaptively useful function of "simulation." Simulation is variously nominated to be an instrument of perceptual recognition, mind reading, emotional affiliation, and empathy. Any or all of that could still be the case. The core neural organization in question would merely not have specifically evolved for any of the stated purposes. Instead, it would represent a preadaptation on which human natural selection capitalized when the above-listed complex cognitive processes evolved.

8.5.4 Selectivity in Imitation

What does it take to elicit an imitative response? In the case of a neuron, as far as we know, it requires a percept that is biologically and therefore motivationally relevant. The grasping hand elicits firing from the mirror neuron if it is grasping an object, but not if it is grasping thin air. In the case of a human and specifically an infant, the imitated motion derives from a person rather than a surrogate, such as a toy, robot, or other machinery. Nor are all human motions that are within sight automatically imitated overtly. The individual seems overtly to imitate only movements that are high on a hierarchy of relevance, firmly fixated in focal attention, and performed by other humans. Covert enactive encoding has broader boundary conditions, implicating specific movements as well as more abstractly conceived actions.

By what means is "resonance" in motor representations that correspond to perceived actions typically held covert, so as not to result in overt movement? If the perception and the production were represented in distinct neural machinery, then an inhibitory barrier between them could hold overt imitation in check. If the resonance arises from an early stage of

unified representation, then an inhibitory barrier at the output side would be more likely. The fact that unwanted imitation occasionally breaks through in pathological cases, both neuropathology and psychopathology, somewhat favors downstream inhibition.

8.5.5 Interpretation of Functional Neuroimaging

Decety and Chaminade present fMRI evidence for differential local activations in relation to both perceptual-motor and socioemotional variables. Dramatic and convincing as this evidence is, it is of course subject to the well-known uncertainties that are common to the field of human functional neuroimaging and have not yet been resolved. Are the most activated areas those that are most critical to the process being studied? Are some inhibitory? As for areas that did not notably activate, we know that areas may participate in a function without producing activation maxima in PET or fMRI scans.

At present, the interpretation of functional images remains tentative, and calls for support from converging findings, particularly in lesion studies. A striking convergence between neuroimaging and lesion neuropsychology is offered by the finding of Decety et al. (2002) that a self-generated action and its observed counterpart exhibit mirror-image activations in opposite hemispheres. Although other areas were activated under both conditions, the left inferior parietal lobule was activated in the self-generated condition, and the opposite, the right inferior parietal lobule, in the other-generated condition. This dissociation is consistent with a proposed dissociation of function between the hemispheres, which attributes pursuing one's own action plan ("continue ongoing behavior") to the left hemisphere, and suspending it to monitor the (animate or inanimate) environment ("interrupt ongoing behavior") to the right hemisphere (Kinsbourne, 1989).

For lesion neuropsychology, the findings of Decety's research program are generally heuristic rather than confirmatory. His studies, and similar ones that he and Chaminade reference, suggest a host of functional localizations that have never been uncovered by traditional neuropsychological studies, although patients with focal brain lesions have been assiduously investigated for a century and a half. This exemplifies an edge for functional imaging over classical neuropsychology, but by the same token it also raises questions.

What focal lesion selectively renders a patient unable to distinguish between actions he performed spontaneously and ones by means of which he imitated another? This specific deficit has not been described in focal brain

lesions. Perhaps that is because it has not been looked for. But then again, confusion between the self and another person would surely be expected to lead to some quite dramatic complaints and observations. Findings from classical lesion neuropsychology and from functional neuroimaging have not been systematically reconciled. Even consensual guidelines as to what lesion effects can be predicted from activation peaks in imaging, and vice versa, do not exist.

When functional brain imaging pinpoints an area that is particularly activated during a given activity, converging supportive evidence is required before firm conclusions can be drawn about the localization of the relevant function. When a lesion destroys this area, is performance impaired, as would be predicted based on the neuroimaging results? For converging validation, one can supplement lesion data with specifically planned and precise momentary inactivation of cortical territories by transcranial magnetic stimulation. Understanding what the activation maxima of functional images convey about how the brain works is a work in progress. Decety and colleagues constructively and imaginatively address the relationship between perception and action. Do they cast light on the brain's basis of intersubjectivity (Trevvarthen & Aitken, 2001), a phenomenon that is regarded as a building block of socioemotional development?

Their stated objective was to further understanding of intersubjectivity. "Intersubjectivity is the process by which mental activity—including conscious awareness, motives and intentions, cognitions and emotions—is transferred between minds" (Trevvarthen, 1999, pp. 415–417). Taking simulation theory for granted, they have conclusively identified at least part of the neural substrate that is activated both by self-action and the action of others. But more simply construed, they show how both acting and perceiving draw on some common and on some disparate neural circuitry. In my opinion, they have not yet promoted understanding of intersubjectivity, but they have in place a research design that can be adapted to that purpose. To verify that among the activated regions are some that interpret the perceived action in terms of motives, intentions, and so on, they might use a different, still more exacting experimental design in which in one condition the subject observes without making inferences ("interobjectivity" only) and in the other the subject observes the same activity, but makes intersubjective inferences. Subtracting the activation pattern of the first from that of the second might possibly reveal components of the inferential brain mechanism.

8.6 Action, Ideation, and Perception

Michael Arbib on W. Prinz

In chapter 5 Wolfgang Prinz rejects sensorimotor approaches in which actions come into being as a consequence of stimulation, in favor of ideomotor approaches, in which everything starts with intention, and actions come into being as the means for realizing those intentions. I would rather stress a cycle of perception-ideation-action that denies primacy to either approach (Neisser, 1976; Arbib, 1989). Since it is a cycle, it does not matter which term comes first. Our action may be in response to an unexpected noise as much as to an idea, while the action may lead to new sensory stimuli whose interpretation, and ensuing course of action, depends on our current internal state (of which the intention may be a small part).

Prinz claims that in the sensorimotor framework, perception and action are subserved by separate and incommensurate representational structures. He points out that on the perceptual side, representations stand for patterns of stimulation in sense organs and their derivatives, while on the action side, representations stand for motor commands and patterns of excitations in muscles. In general, perception does not register sensory stimulation, but instead interprets these as signals for "things in the world," and these "things" can include actions ("sensing" a movement and interpreting it in terms of its observed or inferred goal). This is as true for ideomotor theory as for its alternatives.

Indeed, Jordan and Rumelhart (1992) have shown how learning may create forward models that bring motor output and sensory input into congruence, and Arbib and Rizzolatti (1997) have outlined the relevance of such ideas in exploring the functionality of the mirror system. Moreover, I think Prinz himself supports the action-ideation-perception cycle when he later notes that "regular connections between motor acts and perceivable bodily and environmental effects ... become functional in two different ways. One is to *expect certain effects, given certain acts*; that is, to predict an ongoing action's perceivable consequences. The other way is to *select a certain act, given an intention to achieve certain effects*" (vol. 1, ch. 5, p. 143), and his first endnote discusses the distinction between forward models and inverse models in motor control.

Prinz presents the Lotze-James theory as requiring for an action to be voluntary that "two conditions be met: (1) There must be an idea, or representation, of what is being willed ... and (2) conflicting ideas must be absent or be removed" (p. 142). Unfortunately, this is circular, since it says

no more than that an action is voluntary if it is willed! Will aside, it is also consistent even with involuntary behaviors within sensorimotor theory, for which I take prey capture by frogs and predator avoidance to be a paradigm case (Arbib, 1987). For example, a frog confronted with several flies builds a representation in its tectum that encodes for each fly the "idea" that it should be snapped at (condition 1); then a winner-take-all circuit may determine which fly is indeed snapped at (condition 2). I find it reasonable to believe that a complete theory of human action will include many cases of pairing of action and response that do not require invoking the idea of the triggering action (as when we swerve suddenly to avoid a collision, without consciously recognizing the actions of the other driver).

As Prinz notes, the defining feature of an imitative act is some form of similarity between the act perceived in the other and the act performed by oneself. Space does not permit me to comment on the elegant experiments that Prinz summarizes, save to suggest that they seem more consistent with an action-ideation-perception cycle than purely ideomotor approaches. Instead I want to stress the subtleties that arise in the imitation of new actions. Imagine (for those old enough to remember manual gearshifts) that you have learned to change among the forward gears of a car. You developed a generic skill (a parameterized set of actions) of using compliant motion to get the gearshift to a desired end position. But having mastered that, if you try to imitate someone changing into reverse gear, it is highly likely that the first few times there will be a grinding sound and the car may stall. The problem is that you "recognized" the action as if it were just the "forward-gear action" to move to a new position. It requires a new act of attention, and perhaps explicit instruction, to learn that an additional movement is involved, such as pressing down the head of the gearshift in a direction orthogonal to the overall trajectory to augment correctly the class of actions already in your repertoire.

Thus, what I want to add to the ideomotor story and the mirror neuron story is the fact that when we observe an action or try to understand or imitate an action, it is only in some cases that we can already have the complete idea of the action. In general we need not only the representation of what the action is like but also the representation of how it differs from the thing it is most like. Of course, if there is no significant difference, then we are back in ideomotor territory. But if we are in a situation where finding that difference and factoring it into our behavior is required, then in some sense the perception of the failings of an old idea is driving the creation of a new idea that will then result in a skilled action.

8.7 The Application of Ideomotor Theory to Imitation Merlin Donald on W. Prinz

Wolfgang Prinz has presented several elegant demonstrations of strong coupling between percepts and actions that support an ideomotor approach to imitation. Although I have no problem with his experiments, I am not completely convinced of the applicability of his theoretical model to imitation. Or perhaps it would be more accurate to say that the model cannot be as simple, or as universal as he proposes, for several reasons.

First, it is worth reminding readers that the building blocks of the vertebrate motor system do not follow an ideomotor principle. Simple segmental reflexes, and even some of the more complex suprasegmental reflexes, are sensorimotor in nature, and ubiquitous, even in humans. They are present in the nervous systems of all vertebrates and form the evolutionary foundation for all voluntary action. For very good adaptive reasons, they are generally quite resistant to the kinds of perceptual influences Prinz describes. An example is the so-called "H-reflex," which balances movement patterns that engage antagonistic muscle groups. It adheres to what Prinz calls a sensorimotor principle, in that it is the output of a highly quantifiable, reliable, and linear set of responses to a specific stimulus. The same is true of most basic protective reflexes, such as sneezing and vomiting, even when they involve significant suprasegmental coordination. For many species, this is the only type of action available. Of course, higher vertebrates and human beings have additional kinds of movement control, but reflexes are nevertheless built into their motor systems. In some instances, these reflexes can be overcome by corticospinal influences, but in most cases they cannot. In every case, voluntary action systems evolved on the backbone of reflexes, adding certain modifications but not replacing them. While some classes of action are undoubtedly ideomotor in their governance, many are not, and thus the ideomotor principle is far from universal.

Second, the ideomotor principle does not apply to all classes of voluntary movement. Prinz seems to be claiming that ideomotor theory provides a universal principle that governs voluntary action, including imitation. In humans, it appears that way. But humans are special, and demonstrations in human subjects are not necessarily representative of a universal principle. An ideomotor theory of imitation will have serious difficulty explaining why imitation is so difficult for many species when they obviously have very good control of voluntary movement in some domains. Many

primates have very precise control of voluntary movement in specific areas, such as visual-manual coordination, and yet lack precise imitative skills in those same areas. If their movements were governed entirely by ideomotor principles, then their imitative skills should reflect the precision of their actions and perceptions in various domains, but they do not. For example, apes can visually parse many subtle human gestures and respond to them appropriately, but they cannot reproduce those gestures, despite having sufficient motor control to do so.

The same criticism applies to the theory of “mirror” neurons. To a degree, mirror neurons behave as if they were components of ideomotor maps, and their existence seems to bolster the likelihood that ideomotor theory will prove useful in explaining some aspects of voluntary movement. However, the presence of mirror neurons does not guarantee the existence of imitative skills in a species. Monkeys have mirror neuron systems and learn to make excellent predictions about the consequences of their actions, but are nevertheless very poor at imitation. Imitation is different. Its explanation will not entail a simple extrapolation of a universal principle of movement control. In human evolution, the refinement of imitative skill has been linked to the emergence of mimetic gesture, role-playing, social transmission, and skilled rehearsal; in a word, to the intensification of social life, nonverbal communication, and group coordination.

8.8 How to Analyze Learning by Imitation

Bennett Galef on Heyes

For me, the most interesting feature of Heyes’s associative-sequence learning model of imitation is not that it predicts that the ability to imitate will be experience dependent or that imitative learning is simply the production of novel sequences of familiar acts, controversial though those notions may be. Rather, my attention is captured by the assertion that the ability to imitate rests entirely on processes that are not unique to imitation itself. It is here that the contrast with theories, such as Meltzoff’s active intermodal matching model (Meltzoff & Moore, 1999a), that postulate a dedicated, innate mechanism for imitation that transforms visual input into representations encoding modeled movements, is most pronounced.

The implications of Heyes’s approach are quite profound. If, as Heyes proposes, imitation results from the formation of horizontal links among visual representations and vertical links between sensory and motor representations, such general processes may be better studied in nonimitative

than in imitative situations, where they may be confounded. To the extent Heyes is correct, imitation becomes an epiphenomenon reflecting the activity of basic processes with functions other than support of imitative learning that can be studied without reference to imitation. For example, in Heyes’s view, as I understand it, studies of the effects of practice on learning motor sequences would inform our understanding of imitation, much as studies of rhyming skills and sensitivity to phonemes inform our understanding of reading (Bradley & Bryant, 1983; Bryant et al., 1990). Imitation, like reading, can be viewed as an emergent property of mechanisms evolved for other purposes (Gould & Vrba, 1982).

The second point that I would like to make is that identifying the general substrate of imitation, whether behavioral or neuronanatomical, may be a hopeless task. As Heyes, and others, have indicated repeatedly, the most convincing evidence of imitative behavior in nonhuman animals is found in chimpanzees and birds (Heyes, 2002; vol. 1, ch. 6). Despite decades of effort by numerous investigators (e.g., Visalberghi & Fragaszy, 2002), there is only the most limited evidence of imitative learning in monkeys (Voelkl & Huber, 2000) and none in rodents (C. Mitchell et al., 1999). Birds imitate (e.g., Akins & Zentall, 1998; Moore, 1992), chimpanzees imitate, and humans imitate. Probably dolphins (Herman, 2002) and orangutans (Russon & Galdikas, 1993; van Schaik & Knott, 2001) imitate as well. This unusual phylogenetic pattern suggests that the behavioral phenomena we conventionally label as imitative are products of convergent evolution rather than of descent from a common ancestor (Moore, 1992). If so, there is no reason to expect the same behavioral or physiological substrates to underlie imitation in all imitative species.

Compare imitation by quails, chimps, and humans, as described in the literature. A quail may show an increased probability of using one of two simple responses, say pecking at and stepping on a treadle, after seeing a model use one method rather than the other to obtain food (e.g., Akins & Zentall, 1998). A chimp, after extensive training may learn, marginally, to follow a do-as-I-do command (e.g., Tomasello et al., 1993b) or, when manipulating a complex object, will sometimes copy a demonstrator’s actions with greater or lesser fidelity (e.g., Whiten, vol. 1, ch. 11). A human adult can, without special training, closely imitate a near-infinite number of acts after seeing them but once.

Are such differences in performance quantitative or qualitative? We don’t know. Still, we describe chimps, quails, and humans as “imitators,” although quite different processes may support their imitative learning. Many models of imitation learning may be correct. Different models may

of birds now shown to imitate (but see Suddendorf & Whiten, 2001, and Whiten, 2000, on an apparent correlation between the ability to acquire a "concept" of imitation, and mirror self-recognition).

5. Language providing a correspondence "bridge" For example, the term "frown" is used to label both one's own expression and that of others. But (even setting aside the controversial neonatal case) human imitation emerges before language, and this route will not help budgerigars.

6. Seeing and moving one's own appendages, generalized to others' appendages that look similar Of all the hypothesized experiential routes to imitation, this one does not suffer the inherent problems noted for the others, and to that extent has the most apparent plausibility. However, it seems to almost dismiss the imitative correspondence problem as a problem. It is saying that, for example, I know through experience what to do to make my hands appear to give a "thumbs-up"; and when I see someone else give a thumbs-up, it looks recognizably like my own action, so I can do it. This sounds superficially so easy that we have then to ask why imitation is so elusive in the animal kingdom, and why even in certain human cases such as autism (J. Williams et al., 2001) and apraxia (Goldenberg & Hermsdörfer, 2002) it is problematic, despite associative learning processes evidently being in place. Moreover, as Heyes recognizes, this route will not serve for actions of the self one cannot see, and the budgerigars' copying of beak use is just one of many such cases.

As noted earlier, the ASL model can be seen as having two parts: the vertical correspondence component and the horizontal sequential learning component. However, one could argue that only the first of these counts as a model of imitation as such; for once the correspondence problem is solved, the second component can be seen as the learning of perceived sequences of events *per se*, something we already know certain birds and mammals can do, quite apart from the imitative context (Shettleworth, 1998). The concerns expressed here can therefore be summarized by questioning whether ASL really offers a model of imitation, or instead too often sidesteps, or has to assume solved, the thorny correspondence problem, a criticism that can in a similar way be directed to Byrne's string-parsing model in chapter 9 (i.e., is it really a model of imitation or does it rely on the correspondence problem being already solved?). None of this is to argue that imitation cannot *develop* significantly through experience (to the contrary; Whiten, 2000), but rather that the hypothesis that imitation can be learned *from scratch* through basic associative processes faces some quite fundamental logical difficulties.