

## 6 Imitation by Association

Cecilia Heyes

### 6.1 Introduction

In a small class, a fun-loving student is reading aloud the essay she finished at 4 a.m. Glancing up at the professor, she sees his eyebrows rise sharply from silent concentration to a look of dubious surprise. Another student immediately copies the gesture, showing ingratiatingly that he too has spotted the error. The reader knits her brows in a query.

Imitative movements have a great deal in common with nonimitative movements. Both the ingratiating and the fun-loving student were reacting to the same observed body movement, and both reactions depended on a host of processes that mediated the detection and encoding of the distal stimulus and the selection and control of the motor response. Furthermore, in this example, performance of the imitative and nonimitative movements was based on the attribution of mental states. Both students understood the professor to be expressing doubt and disapproval, and both intended their reactions to communicate how much (or how little) they understood.

The obvious difference between imitative and nonimitative movements is that in the former case but not the latter there is a relationship of similarity, a match, between the observed movement, the stimulus, and the reactive movement, the response. This relational property provides a minimal definition of imitation; it distinguishes imitation from other contingent reactions to observed body movements.

If a fundamental feature of imitation is contingent behavioral similarity, research on imitation has the task of explaining the causes and consequences of this similarity. Investigating consequences, contributors to these volumes consider whether the practice of imitation (rather than that of producing dissimilar or nonimitative reactions to body movements) is a requirement for cumulative cultural evolution (Tomasello & Carpenter,

vol. 2, ch. 5; Donald, vol. 2, ch. 14; and Gil-White, vol. 2, ch. 16), whether it promotes mutual understanding (Gallese, vol. 1, ch. 3; Pepperberg, vol. 1, ch. 10; Meltzoff, vol. 2, ch. 1; Goldman, vol. 2, ch. 2; and Kinsbourne, vol. 2, ch. 7) and what its effects are on economic behavior (Sugden, vol. 2, ch. 15) and interpersonal aggression (Eldridge, vol. 2, ch. 11; Huesmann, vol. 2, ch. 12; and J. Prinz, vol. 2, ch. 13). Addressing causes, researchers are investigating which neural systems are involved in the production of imitative reactions (Rizzolatti, vol. 1, ch. 1; Iacoboni, vol. 1, ch. 2; and Decety & Chaminade, vol. 1, ch. 4), whether common representational structures are involved in detecting a model's movement and producing an imitative movement (Prinz, vol. 1, ch. 5 and Dijksterhuis, vol. 2, ch. 9), and how features of a model's performance are analyzed and selected for imitative production (Byrne, vol. 1, ch. 9; Whiten et al., vol. 1, ch. 11; Anisfeld, vol. 2, ch. 4; and Harris & Want, vol. 2, ch. 6).

This chapter is about the causes rather than the consequences of imitation. It addresses the "correspondence problem" (Alissandrakis et al., 2002a; Nehaniv and Dautenhahn, 2002a), a question posed only by the phenomenon of imitation. How is sensory input from another individual's movements translated into matching motor output? The chapter begins by outlining the *associative sequence learning* (ASL) model, which offers a simple account of the kind of information-processing mechanisms responsible for this translation, and their origins in experience (Heyes & Ray, 2000; Heyes, 2001a; Hoppitt & Laland, 2002). Subsequent sections compare the ASL model with other hypotheses about the causes of imitation and identify evidence that does or would help to resolve key issues.

## 6.2 Associative Sequence Learning

The ASL model offers an account of the information-processing mechanisms involved in imitation of both familiar and novel movements, of movements that were and were not previously part of the observer's repertoire. Cases of both kinds are described here as imitation, but only the latter constitute imitation learning.

### 6.2.1 Imitation

The ingratiating student provides an example of the imitation of a familiar movement. He had, presumably, raised his eyebrows many times before he saw the professor doing so; he did not learn in this episode *how* to raise his eyebrows. Nonetheless, there was something inside him providing the

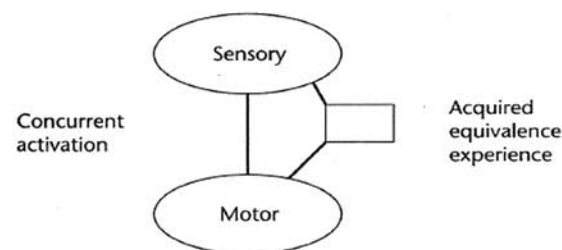


Figure 6.1

Direct and indirect vertical associations.

potential to react to the sight of the professor's movement with a matching movement, and the source of this potential is far from obvious. Note that the professor's and the ingratiating student's movements looked the same from a third-party perspective (e.g., to the fun-loving student), but not to the imitator himself. Watching the model, he saw arcs rising above dark blobs in the upper part of an oval face, but raising his own eyebrows yielded little, if any, distinctive visual input; the movement was felt rather than seen. Any solution to the correspondence problem must be able to explain cases like this, in which the imitated movement is "perceptually opaque," as well as those, such as imitation of finger movements, in which it is "perceptually transparent" (Heyes & Ray, 2000).

The ASL model suggests that both when the movement is transparent and when it is opaque, imitation is made possible by a vertical association between a mental picture of the movement performed by the model, and a mental image of what it feels like to perform the same movement, i.e., a movement that is perceived as the same from a third-party perspective. In other words, imitation is made possible by an excitatory link between a sensory representation of the observed movement (encoding properties of the movement detected via the distal senses) and a motor representation (encoding somatosensory properties and motor commands) (see figure 6.1). When such a link exists, excitation of the sensory (typically visual) representation, by observation or recollection of the represented movement, will lead to excitation of the motor representation.<sup>1</sup> If the link or association is

1. Vertical associations are likely to be bidirectional, allowing the sensory representation to activate the motor representation, as in imitation, but also enabling the motor representation to activate the sensory representation, as in some cases of internal mental practice or motor imagery.



sufficiently strong, excitation of the motor representation will result in the activation of muscles involved in execution of the represented movement (e.g., Strafella & Paus, 2000; Stürmer et al., 2000), but it may or may not lead to overt performance of the movement. Typically, at least in healthy human adults, activation of the motor representation can be inhibited so that performance of imitative movements, like that of nonimitative movements, can be brought under intentional control.

Where do these vertical associations come from? The ASL model suggests that whereas a few vertical associations may be innate, the majority are formed through experience that provokes concurrent activation of sensory and motor representations of the same movement. This experience may consist of concurrent observation and execution of the same movement, leading to a “direct” vertical association, or it may involve exposure to a common stimulus in conjunction with, on some occasions, observation of the movement, and on other occasions with its execution.<sup>2</sup> For example, a child may hear the sound of a word such as “frown,” sometimes when she is frowning and at other times when she sees another person frowning. As a consequence of this “acquired equivalence” experience (Hall, 1996), sensory and motor representations of frowning will each become linked to a representation of the sound of the word. This “indirect” vertical association enables activation of the sensory representation to be propagated to the motor representation via the word representation, and to the extent that it allows the sound of the word concurrently to activate sensory and motor representations of frowning, to the formation of a direct vertical association between them (Rumiati & Tessari, 2002; Tessari & Rumiati, 2002).

The ASL model assumes that nonmatching vertical associations, links between a sensory representation of one movement and a motor representation of another, providing the basis for systematically non- or counter-imitative behavior, can be formed in the same way as matching vertical associations, links between sensory and motor representations of the same movement. It suggests that if a system contains more matching than nonmatching vertical associations, this discrepancy is due primarily to the environment in which the system has developed.

The human information-processing system typically develops in an environment that favors the formation of matching vertical associations in

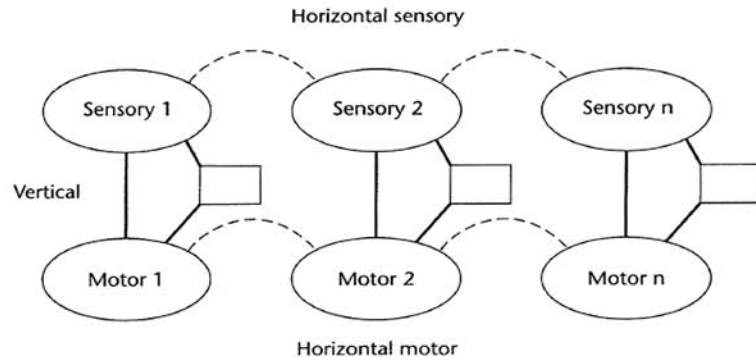
2. Although concurrent activation or “contiguity” is emphasized here for clarity of exposition, the ASL model assumes, in line with contemporary theories of associative learning (see Hall, 1994 for a review), that the formation of vertical associations depends on contingency in addition to contiguity.

a number of ways.<sup>3</sup> First, gross human anatomy is such that many movements of one’s own distal appendages can be viewed in much the same way as those of another person. When I watch many of my own hand and finger movements, the appropriate motor representations are activated concurrently with sensory representations, arising from visual feedback, which are similar to the visual percepts that arise when I observe someone else performing the same movements. Second, the typical environment of human development contains optical mirrors and other reflecting surfaces—instruments that allow one’s own facial and whole-body movements to be viewed from a third-party perspective—but not video playback devices that provide visual feedback from one movement during execution of another. Third, during early development, humans are surrounded by other humans who imitate them. Learning over a cot, we coo when a baby is cooing, grimace when a baby is grimacing (Field et al., 1985; Papousek & Papousek, 1989). We do not react to cooing with grimacing or grimacing with cooing in a way that would promote the formation of nonmatching vertical associations. Finally, there is language. Generally speaking, the range of movements constituting the referents of each action word look more alike, from a third-party perspective, than those of other action words. Only if natural languages contained words like “frile,” referring to you frowning and to me smiling, would the use of language promote the formation of nonmatching vertical associations.

### 6.2.2 Imitation Learning

The ASL model assumes that a novel behavior consists of familiar elements or “primitives” arranged in a novel sequence and that two kinds of processes are initiated when a novel behavior is observed (see figure 6.2). First, sensory representations of the sequence components are activated and “horizontal” links are formed between them. The model says little about these horizontal links (see Byrne, vol. 1, ch. 9, for discussion) because it assumes that they are not specific to imitation; that they are formed through the same processes of sequence learning regardless of whether the novel behavior is imitated, counterimitated, or merely stored to allow future recognition. However, to the extent that each sensory representation

3. I am grateful to Giacomo Rizzolatti for suggesting, in a question during my talk, that the challenge for the ASL model is not to show that humans have experience that would lead to matching vertical associations, but to explain why they have more experience of this kind than of the kind that would promote the formation of nonmatching vertical associations.



**Figure 6.2**

The associative sequence learning (ASL) model of imitation.

in the sequence is part of a matching vertical association, formation of the horizontal links between sensory representations will allow the second, imitation-specific process to occur. That is, successive activation of each sensory representation, by observation or recollection of the model's behavior, will provoke activation of matching motor representations in the same order, providing the potential for overt performance of a rough copy of the observed novel movement. Furthermore, repetitive activation of this sequence of motor representations allows them to become horizontally linked. This horizontal linkage of motor representations constitutes motor learning, produces a new motor primitive, and improves the potential fluidity of imitative movement. Thus, according to the ASL model, imitation learning occurs when matching vertical associations allow sensory input from another's behavior, rather than feedback from one's own, to provide the input for motor learning (Heyes, 2003).

### 6.3 Evolution and Development

The ASL model implies that the capacity to imitate and to engage in imitation learning does not depend on an innate module, a special-purpose cognitive mechanism shaped by natural selection (cf. Meltzoff & Moore, 1997; Decety & Chaminade, vol. 1, ch. 4). Rather, it suggests that the capacity to imitate depends on experience, that the effect of this experience is to reconfigure relatively general-purpose cognitive processes, and that this effect is mediated by associative learning; arguably the most general of all psychological processes. Thus it proposes that vertical associations

are formed by the same processes that mediate Pavlovian and instrumental conditioning and that a repertoire of these vertical associations links together two sets of horizontal processes—one involved in learning sequences of distal stimuli and the other in motor learning. Vertical associations reconfigure these into a common system in the sense that they allow sensory input from the movements of others to drive motor learning.

#### 6.3.1 Animals

Evidence that chimpanzees can imitate (e.g., Whiten, 1998 and Whiten et al., vol. 1, ch. 11) is consistent both with the ASL model and with a more nativist conception of imitation. It is plausible that chimpanzees have the opportunity to form some vertical associations and that an innate module for imitation is ape- or primate-specific, rather than present in humans alone. Studies reporting that "enculturated" chimpanzees, those that have had extensive contact with humans, are better imitators than their wild-born and/or mother-reared conspecifics (Tomasello, 1996) suggest that experience plays a substantial role either in enhancing imitative performance or in generating imitative ability. A role of the latter kind is less compatible with a nativist perspective than with one that emphasizes ontogenetic processes, but these data certainly do not provide specific support for the ASL model. This model would imply that the active ingredients of human contact, the experiences that make a specific contribution to the development of imitative ability in chimpanzees, involve being imitated, being rewarded for imitation, exposure to mirrors, and hearing words or other signifiers—experiences providing the opportunity to form vertical associations. However, until we know the results of research comparing the imitative abilities of chimpanzees raised with different kinds of human contact, it will remain equally plausible that the active ingredient is, for example, engagement in joint attention behavior (Carpenter et al., 1995).

Because avian species are such distant relatives of hominids, research on imitation in birds has a more direct bearing on the question of whether human imitation is based on an innate module or on more species- and domain-general processes of learning. Recent work has provided evidence of imitation in birds that is at least as strong as the current evidence of imitation in chimpanzees. It suggests that pigeons, quail, starlings, and budgerigars can imitate simple movements such as thrusting the head upward or downward to remove a stopper from a food box (F. Campbell et al., 1999; Heyes & Saggerson, 2002), and operating a treadle with a foot or beak (e.g., Akins & Zentall, 1996). Furthermore, the most advanced research in this field indicates that like imitation in humans, motor imitation by

birds can be deferred, and that it can be modulated by the perceived consequences of the model's behavior (Akins & Zentall 1998; Dorrance & Zentall, 2001). These data do not "disprove" the existence of an ape- or human-specific innate module for imitation, but they certainly imply that such a device is unnecessary to produce imitative behavior, and given the small brain and limited intellect of the average bird, make it plausible that in the right rearing environments, associative learning is sufficient to yield a capacity for imitation.<sup>4</sup>

The ASL model suggests that a capacity for associative learning is the basic psychological requirement for imitation, and this capacity is known to be present in a broad range of taxa. So why is there relatively little evidence of imitation in animals? Why apes and birds, and not monkeys and rats? One possible answer is that the ASL model is wrong; ape and avian imitation could be mediated by different, specialized processes that are products of convergent evolution. This view would be supported if as much high-quality research effort was devoted to imitation in monkeys and rats as in apes and birds, and if this effort revealed, not merely ambiguous findings (e.g., C. Mitchell et al., 1999; Visalberghi & Frigaszy, 1996), but that, given comparable opportunity to form vertical associations, they are less able to imitate. The ASL model predicts that given the right experience, many animals will be able to imitate. However, it also suggests that

4. It may seem strange to emphasize birds' capacity for motor imitation, their ability to copy body movements, given that they have long been famous for their vocal imitation. There is no doubt that some birds have a remarkable capacity to imitate conspecific vocalizations and/or arbitrary human sounds (see, e.g., Pepperberg, vol. 1, ch. 10), and that this is important from a number of theoretical perspectives. However, in relation to the specific question of whether human imitation is based on general processes or on a dedicated innate module, avian vocal imitation is less informative than avian motor imitation. This is because vocalizations are relatively perceptually transparent; they give rise to similar sensory inputs when observed and executed; and therefore imitation of vocalizations could be achieved by an error-correction process that adjusts motor output until sensory feedback matches sensory input from the model (Mowrer, 1960). An error-correction process of this kind is insufficient to explain imitation of opaque movements (e.g., facial expressions) because when an opaque movement is imitated, the sensory feedback to the observer does not match the sensory input received from the model. Human imitative competence includes opaque movements, and therefore it is easier to argue that human imitation and avian vocal imitation are based on distinct psychological mechanisms than to show that human and avian motor imitation have independent roots (Thorndike, 1898).

humans acquire many of their vertical associations through cultural artifacts and practices, and therefore the model could be compromised by too much, as well as too little, imitation in animals.

### 6.3.2 Infants

An obvious implication of the ASL hypothesis is that the range and identity of the movements that a child, or indeed an adult, can imitate will depend on their experience, and specifically on the range of movements that they have either concurrently seen and done, or for which they have associates in common, such as verbal labels. To suppose that a few matching vertical associations are innate, rather than experience dependent, would not radically change the model, but it would not be viable if, as suggested by a substantial body of research from several laboratories, newborn human infants can imitate a broad range of face and hand gestures (see Meltzoff & Moore, 1997 for a review).

Having reanalyzed all published experimental data on imitation in neonates, Anisfeld (1991, 1996; Anisfeld, vol. 2, ch. 4; and Anisfeld et al., 2001) has concluded that the evidence is compelling for only one movement—tongue protrusion. For other candidates, such as mouth opening and lip protrusion, he argues that the data are inconclusive, either because there is a preponderance of negative findings or because the reported effects could be artifacts produced by imitation of tongue protrusion. Supporting and extending this view, recent studies have reported further evidence that neonatal imitation is confined to tongue protrusion (Couturier-Fagan, 1996), the failure to find imitation of tongue protrusion (Ullstadius, 1998), and results consistent with the idea that the tongue protrusion effect in early infancy is due either to an "innate releasing mechanism" or vertical association (Heimann & Ullstadius, 1999) or is not sufficiently specific to constitute imitation (Jones, vol. 1, ch. 12). The latter studies show that the frequency of tongue protrusion in very young infants increases not only when they have observed tongue protrusion but also when they are exposed under comparable conditions to flashing lights (Jones, 1996) or rousing music (Jones, 2002).

Many infancy researchers continue to believe that newborns can imitate a range of movements and that this provides decisive support for the idea that human imitation is based on an innate module. However, the "believers" seldom refer to the work of Anisfeld, Jones, and other skeptics, and have not yet addressed the specific questions their work has raised (e.g., Nadel & Butterworth, 1999; Heyes, 2000). Until there is open exchange on these issues, it will be difficult to draw any firm conclusions



about neonatal imitation, and thereby to take an important step in resolving the question of whether imitation is mediated by a dedicated, innate module or by relatively domain- and taxon-general processes of learning.

## 6.4 Intentionality and Culture

It has been proposed that imitation is intrinsically “goal directed” (e.g., Gattis et al., 2002) or “intentional,” and that it is this property that enables imitation to contribute to cumulative cultural evolution (Tomasello et al., 1993a). Depending on what is meant by goal and intention, the ASL model is compatible with both of these proposals.

### 6.4.1 Outcomes

The term “goal” is sometimes used to refer to an observable outcome or effect of a movement; to a state of the actor’s body, or of an object, or a spatial configuration between the actor’s body and an object, which obtains at the end of a whole movement or at the end of a component of a movement sequence. For example, a fist may be the outcome of a hand-closing movement, and gripping a glass may be the outcome of reaching toward a glass. If a goal is understood to be an outcome, then the ASL model is perfectly consistent with the suggestion that goals are an intrinsic part of the imitation process.

Intransitive movements, such as hand and facial gestures, have been used to illustrate the ASL model (e.g., Heyes & Ray, 2000) because when imitation is understood to consist of contingent, similar reactions to body movements, intransitive movements provide the clearest examples. With transitive (i.e., object-related) movements, such as dropping a ball in a cup, it is not always obvious that the model’s body movement was part of the stimulus configuration to which the observer was reacting; he or she may have been reacting to the object’s movement alone. However, the ASL model applies to imitation of both transitive and intransitive movements—to contingent, similar reactions to stimulus configurations that include body movements plus movements of an object, and to body movements alone—and assumes that in both cases a large proportion of what is encoded in sensory and motor representations is information about outcomes. Specifically, and in accordance with what has long been known about motor control, it assumes that the motor commands encoded in motor representations specify outcomes, not “muscle twitches” (D. Campbell, 1954).

### 6.4.2 Intentions

A goal can be more specifically characterized, not just as any outcome, but as an outcome toward which a movement is directed by an error-correction process, and in this case the outcome may be described as an intention and the movement as intentional. Thus, dropping a ball in a cup would be a goal if it followed adjustments to the actor’s hand position made by a process sensitive to the distance between hand and cup, but this outcome would not be a goal if the ball just happened to land in the cup when the actor sneezed and lost his grip. The ASL model suggests that this kind of goal directedness is a property of many but not all cases of imitation, and that it has limited potential to explain the key feature of an imitative action—its similarity to that of the model.

Error-correction processes can contribute to the formation of matching vertical associations. For example, an error-correction process comparing visual feedback from an observer’s own finger movements with a visual representation of a model’s finger movement would terminate when the observer was performing the same movement as the model, i.e., when a sensory (visual) representation and motor representation of the same action were concurrently activated. However, this is only one of many routes to the formation of matching vertical associations. For example, mirror exposure, synchronous movement, and being imitated all allow concurrent activation of matching sensory and motor representations without the involvement of error-correction processes, and error correction is available only for perceptually transparent movements.

An error-correction process comparing visual feedback from opaque movements, such as facial gestures, with a visual representation of a modeled movement would not be able to reduce the discrepancy substantially, and when it had done the best it could, the executed movement would hardly resemble the one observed, i.e., the sensory representation would be activated concurrently with a nonmatching motor representation. Perhaps, as proposed by Meltzoff and Moore (1997), there are error-correction processes that compare movements, not in terms of their visual features, but in relation to some properties that even opaque movements have in common when they are observed and executed. This is an interesting possibility, but it is difficult to test without clear hypotheses about the “common currency”, about the nature of the nonsensory or “amodal” properties used for comparison and the processes through which they are derived from sensory input.

In addition to enabling formation of some matching vertical associations, and thereby contributing to the potential to imitate, it is likely that

error-correction processes are typically involved in translating this potential into imitative performance. Thus, when a motor representation is activated by a sensory representation via a vertical association, the observer may intend to perform the represented act. In other words, performance of the act may be regulated by an error-correction process that compares somatosensory information encoded in the motor representation with somatosensory feedback from movement, as well as visual information encoded in the sensory representation and visual feedback (Wolpert & Kawato, 1998). However, this is not a distinctive feature of imitation. Performance of many nonimitative and counterimitative movements is also goal directed or intentional in this sense, and like those other categories of movement, imitation is not always intentional. The occurrence of non-intentional imitation is indicated in healthy adult humans by the chameleon effect (Chartrand & Bargh, 1999; Dijksterhuis, vol. 2, ch. 9), and the phenomenon of perceptual induction (Knuf et al., 2001; W. Prinz, 2002), and in patients with frontal lesions by their utilization behavior (Lhermitte, 1983).

#### 6.4.3 Higher-Order Intentions

In the third and most specific sense of goal to be considered here, imitation is intrinsically goal directed if it is mediated by higher-order intentions; if imitation invariably involves the observer making inferences about the outcome that the model intended to achieve through an action. In this sense the observer's goal is an outcome that is represented by the observer as being the same as the outcome intended by the model. Thus, imitative performance after watching a ball dropped in a cup would be guided by the intention, not merely to drop the ball in the cup, but to drop the ball in the cup *as intended by the model*.

It is not yet certain at what age it becomes possible for imitation to be guided by higher-order intentions (Heyes, 2001a; Huang et al., 2002; Meltzoff, 1995), but introspection alone leaves little doubt that once we reach adulthood, many of our imitative actions are guided in this way. In an aerobics class, or when being trained to operate a machine, one selectively imitates intended actions—the pirouette and the key press, not the jarring of the table or the sneeze. However, the ASL model does not distinguish imitative actions guided by higher-order intentions from other imitative behavior for two reasons. First, invoking higher-order intentions does not help to solve the correspondence problem, the explanatory challenge uniquely posed by imitation. My intention to imitate an intended action

may on some occasions help to explain *why* I did the same thing as a model, but it does not explain *how* I was able to do it.

Second, the available evidence suggests that in nature there is no special relationship between imitation and higher-order intentionality; imitation often occurs without the attribution of intentions to the model (e.g., Dorrance & Zentall, 2002; Chartrand & Bargh, 1999; Dijksterhuis, vol. 2, ch. 9; Knuf et al., 2001), and when the model's intentions *are* represented, it seems that, as in the case of the fun-loving and ingrating students, they are just as likely to lead to nonmatching action as to imitation.

Recent positron emission tomography (PET) studies have addressed these issues by comparing activation of the medial prefrontal cortex, an area implicated in the attribution of intentions to others (Shallice, 2001), during imitation and during the performance of carefully chosen control tasks (Decety & Chaminade, vol. 1, ch. 4). One of these studies (Chaminade et al., 2002) used video footage of a model's hand selecting and grasping a Lego block from an array and moving it to one of several locations on a board. On any given trial, the participants were allowed to observe the beginning of this sequence (the "means"), the end of the sequence (the "goal") or the whole sequence, before being required to select the same block and put it in the same place on their own board. The results showed that regional cerebral blood flow (rCBF) in the medial prefrontal cortex was greater during "means" trials than during "goal" trials. If we assume that activation of the medial prefrontal cortex is specific to higher-order intentionality, this effect implies that at least in the "means" condition, imitative performance was guided by the attribution of intentions to the model. However, the involvement of higher-order intentionality under these circumstances does not imply that imitation necessarily or even typically involves the attribution of intentions. At least two features of the experimental situation are likely to have promoted attribution of mental states. The participants were explicitly instructed to do the same thing as the model, and, on "means" trials, their view of the outcome of the model's movement was tantalizingly occluded.

In another PET study (Decety et al., 2002), rCBF in the medial prefrontal cortex was greater when the participants were imitating a model's hand moving objects around within an array (imitation) and when they were watching the hand of a person imitating their own spontaneous movements of the objects (being imitated), than when they simply received visual feedback from their own spontaneous movements of the objects



(control). This result implicates the medial prefrontal cortex, and therefore higher-order intentionality, in the processing of contingent body movements, but it does not demonstrate a special relationship between the attribution of intention and matching contingent body movements, i.e., imitation. To test for such a relationship, it would be necessary to compare activation of the medial prefrontal area when participants are required to respond to each modeled movement with a matching movement (imitation) or with a nonmatching movement (counterimitation).

If it turns out that imitation does not typically involve higher-order intentionality and that the attribution of intentions is no more likely to give rise to imitative than to counterimitative action, there may still be good reasons to focus research attention on behavior that is both imitative and guided by higher-order intentionality—on the intersection between these two sets. Even if, as the ASL model implies, higher-order intentionality does not play an especially significant or distinctive role in the causation of imitative behavior, it remains plausible that fostering the development of higher-order intentionality (Iacoboni, vol. 1, ch. 2; Gallese, vol. 1, ch. 3; Meltzoff, vol. 2, ch. 1; and Goldman, vol. 2, ch. 2), and promoting cumulative cultural evolution (Tomasello et al., 1993a) are its most important effects, and behavior at this intersection is crucial for testing these hypotheses. The ASL model is not incompatible with these proposals but it has a different focus. It offers a solution to the correspondence problem, and in the process emphasizes, not what imitation can do for culture, but what culture can do for imitation (Heyes, 2001b).

## 6.5 Mirror Neurons and Representation

The ASL model assumes that vertical associations are formed through concurrent activation, a Hebbian principle that can be expressed, in neurological rather than psychological or functional terms, as “neurons that fire together, wire together.” It is therefore unsurprising that the ASL model is compatible with the existence of mirror neurons (Rizzolatti, vol. 1, ch. 1), and, more generally, mirror tissue (e.g., Buccino et al., 2001). Broadly speaking, it suggests that mirror tissue is the neurological embodiment of vertical associations. Neurons that previously fired only during execution of an action become mirror neurons, which fire during observation as well as execution of an action through linkage with neurons that discharge only during observation of an action. A link of this kind is formed when the two neurons are activated at the same time, and this happens most commonly

when the action is simultaneously observed and executed (Iacoboni, vol. 1, ch. 2).<sup>5</sup>

In this section, the foregoing functional interpretation of mirror neurons is compared with two alternatives. The first alternative suggests that mirror tissue mediates symbolic or amodal representation of action, and the second encourages us to interpret the function of mirror neurons in the context of an ideomotor theory of action.<sup>6</sup>

### 6.5.1 Inference versus Association

A vertical association consists of a sensory and a motor representation linked so that one can excite or activate the other. It does not include an abstract, symbolic, or amodal representation of action, a representation of a kind that could be said in any traditional sense to support inferences or confer meaning (cf. Bandura, 1986; Meltzoff & Moore, 1997). However, some neurological studies seem to indicate that mirror neuron activity mediates symbolic representation.

One study of this kind shows that some mirror neurons in the ventral premotor cortex of the monkey fire when the animal reaches for and grasps an object, when it observes the experimenter executing the same sequence, *and* when the final part of the experimenter's action, the grasp, is hidden from view by a screen (Rizzolatti, vol. 1, ch. 1; Umiltà et al., 2001). Discharge occurs in the latter condition only if, prior to the trial, the monkey saw the object at its usual location. This striking finding can be naturally and vividly described by saying that the neurons in question “infer” the experimenter's movement when it is occluded, but this description should not obscure the possibility that associative, rather than strictly inferential, processes are responsible. In Pavlovian conditioning, a response that was

5. Some mirror neurons fire not only when a monkey observes and executes a particular action but also when it observes the object toward which the action is typically directed (vol. 1, ch. 1 by Rizzolatti and ch. 2 by Iacoboni). Viewed from the ASL perspective, this suggests that “Hebbian” connections can be formed between sensory (visual) representations of objects and motor representations, as well as between sensory (visual) representations of observed actions and motor representations. The term “vertical association” refers primarily to connections of the latter kind because by definition it is these that play a specific, functional role in imitation.

6. If mirror neurons are the neurological equivalent of matching vertical associations, one would expect the monkeys in which mirror neurons have been identified to be able to copy the specific actions for which their neurons have mirror properties. As far as I am aware, this hypothesis has not been tested.

once made only after the second of two stimuli comes to be elicited by the first. Applying this principle to mirror neurons, one would expect that after repeated observation of reaching toward an object, followed by grasping the object, the first stimulus, the sight of reaching, would be sufficient to make the neuron fire. However, if the first stimulus configuration was altered between training and testing by, for example, removal of the object from its usual location, generalization decrement would reduce the probability of conditioned responding.

Experiments implicating Broca's area (BA 44/45) in imitation (e.g., Iacoboni et al., 1999) may also seem to indicate that at the functional level, imitation is mediated by symbolic or amodal representations. This is because Broca's area is famous for its role in production of language, and we tend to think of language as a quintessentially symbolic system. However, three considerations make it clear that one cannot infer from Broca's activation that symbolic mediation is necessary for, or typical of, imitation. First, current evidence suggests that imitation is not always associated with activation in Broca's area (Grèzes & Decety, 2001). Second, Broca's area has nonlinguistic as well as linguistic functions (e.g., Iacoboni, vol. 1, ch. 2; Wise et al., 1998). Third, and most important, linguistic mediation does not necessarily imply symbolic mediation.

Even when Broca's area is involved in imitation because of its linguistic functions, when it reflects covert naming, it does not follow that a causal role in generating imitative behavior is being played by nonsensory, non-motor representation of the meaning of the action. In language users, naming an action enables it to enter a conceptual network of the sort that confers meaning and supports inferences. However, naming is itself an act—a motor event that is sometimes detectable through the senses—and it is fully possible that when naming plays a role in imitation, it does so via its sensorimotor properties, not through the access it provides to the conceptual system. This possibility is represented in the ASL model by indirect vertical associations, links between representations of the sight and the feel of an act that are formed through experience in which each has been paired, on separate occasions, with exposure to a common stimulus such as a name.

Recent behavioral experiments using the serial reaction time task (Nissen & Bullemer, 1987) suggest that symbolic mediation is unnecessary, not only for imitation, but also for imitation learning. In these experiments (Heyes & Foster, 2002), observers watched a model performing a complex sequence of finger movements in response to an asterisk moving between boxes arranged in a horizontal line on a computer screen. The observers

were subsequently required to perform the task themselves under three conditions: when the task was exactly the same as that performed by the model (basic transfer), when the screen stimuli were arranged vertically rather than horizontally (perceptual transfer), and when responses were made with the thumbs rather than the fingers (motor transfer). If the observers encoded what they saw symbolically, if they acquired abstract or amodally represented sequence knowledge by observation, they would be expected to perform well under each of these three conditions. In fact, however, when compared with controls who had not observed the model, the observers showed evidence of learning in the basic and perceptual transfer tests, but not in the motor transfer test. This implies that the observers' learning was effector specific, that it could not be transferred from fingers to thumbs, which is what one would expect if, via vertical associations, the sight of the model's finger movements excited motor representations of the same finger movements but did not give rise to symbolic processing.

### 6.5.2 Similarity versus Contiguity

The ideomotor framework (e.g., W. Prinz, 2002 and chapter 5) and the ASL model have a great deal in common. They are both compatible with William James's (1890) formulation of the ideomotor principle: "Every representation of a movement awakens in some degree the actual movement which is its object; and awakens it in a maximum degree whenever it is not kept from doing so by an antagonistic representation present simultaneously in the mind" (James, 1890, vol. II, p. 526). They both portray imitation as a product of more general psychological processes, rather than of a specialized module, and they are both consistent with recently discovered stimulus-response movement compatibility effects (e.g., Brass et al., 2001; Stürmer et al., 2000; W. Prinz, vol. 1, ch. 5). Finally, both models suggest that, in some sense, there is a common representational scheme underlying the perception and performance of action. However, whereas the ASL model claims that this commonality is acquired and based on contiguity, the ideomotor framework implies that it is intrinsic and based on similarity.

Thus, it would appear that the ASL model and the ideomotor framework differ with respect to two questions. First, is imitation mediated by one kind of representation or two? The ideomotor framework says that a single kind of representation of action has both sensory and motor properties, while the ASL model assumes that there are distinct sensory and motor representations, which become linked by vertical associations into a

common representational scheme. Second, is imitation governed by similarity or by contiguity? The ideomotor framework suggests that observation of a movement,  $x$ , will activate or “awaken” performance of actions to the extent that they are similar to  $x$ , while the ASL model suggests that sight of  $x$  promotes the performance of other actions to the extent that they have in the past been performed contiguously with the sight of  $x$ .

The first of these contrasts may be more apparent than real. The ideomotor framework may be expressing the view that even if integration of sensory and motor processing is acquired rather than intrinsic (e.g., even if mirror neurons are vertical associations), the level of integration typically achieved in mature humans is so profound that it is no longer useful to distinguish sensory and motor representations. If so, the contrast between the ideomotor framework and the ASL model merely reflects their different preoccupations; the former with mature, human competence and the latter with learning and development in human and nonhuman animals. However, if the “one representation or two” question is empirical, experiments examining the role of experience in the development of mirror neurons and tissue and, more broadly, in the capacity to imitate, may help to resolve this issue in addition to the “similarity versus contiguity” question.

For example, the ASL model predicts that neurons in monkey area F5 will develop mirror properties only for actions that have been contiguously seen and done, or for which a common stimulus has been paired on some occasions with observation of the action and on other occasions with its execution. Furthermore, it predicts that “countermirror” properties could be acquired in the same way; that if a monkey repeatedly saw a precision grip before executing a power grip, neurons would emerge that fire both during observation of the first action and execution of the second. Translating this prediction to the functional level, the ASL model suggests that if a person repeatedly saw a hand-opening movement while performing a hand-closing movement, the sight of the hand opening would no longer facilitate the same response, and that after sufficient training it would facilitate the opposite response (Stürmer et al., 2000).

Countermirror and counterimitation effects of this kind would not be expected if similarity, rather than contiguity, is the fundamental principle of imitation. However, even if these effects were observed, similarity could still play an important role in imitation. We could infer that similarity was irrelevant only if equally strong imitation and counterimitation (or mirror and countermirror) effects emerged from equivalent amounts of training, and even the literature on associative learning, the context of the ASL

model, provides reason to doubt that this would always be the case (Hall, 1994). Early associationists, such as Hume (1740/1984) and J. S. Mill (1843/1974), emphasized the importance of “resemblance,” and although experimental psychologists have been able to find little evidence that the formation of associations depends on similarity rather than temporal contiguity, the results of a few studies suggest that when contiguity is carefully controlled, similar stimuli are more readily associated (e.g., Rescorla & Furrow, 1977; Lolordo & Jacobs, 1983). If analogous results were obtained in research on imitation—where, by hypothesis, the conditioned and unconditioned stimuli derive from observation and execution of an action—the ASL model would need to be modified to incorporate the principle of similarity. This would include acknowledgment that matching vertical associations are formed more readily than nonmatching vertical associations, and would bring the ASL model and the ideomotor framework even closer together.

## 6.6 Conclusion

The ASL model is empiricist in that it emphasizes the role of experience in producing the capacity to imitate, but it is neither behaviorist nor anti-evolutionary. Associative learning mechanisms are inferred, rather than directly observable, causes of behavior, and the model assumes both that they are products of natural selection and that they operate according to an evolutionary algorithm of variation and selective retention (D. Campbell, 1974; Heyes, 2003).

It has been argued in this chapter that the ASL model is consistent with current behavioral and neuroscientific data on imitation in human and nonhuman animals, and that it is compatible both with the idea that imitation is intrinsically outcome directed and with many of the central tenets of the ideomotor framework. It is at odds with claims that imitative action is necessarily intentional, or that it invariably involves the attribution of mental states, but these may well turn out to be definitional issues. More interesting from an empirical perspective are the contrasts between the ASL model, the ideomotor framework, and theories suggesting that imitation is mediated by amodal or symbolic representations. The resolution of these issues will not only require experimental research of the kind discussed here, but also clearer and more complete specification of the models themselves, enabling each to generate differential, testable predictions. The ASL model may well turn out to be quite wrong—its fallibility may be its

greatest strength—but it will have fulfilled a function if it contributes to the development of a clear, detailed theory of imitation, with firm empirical support (Wimsatt, 1987).<sup>7</sup>

### Acknowledgments

I am grateful to the Economic and Social Research Council's Research Centre for Economic Learning and Social Evolution (ELSE) for their financial support, and to Marco Iacoboni for comments on an earlier draft.

7. See comments on this chapter by Galef, vol. 1, ch. 8.8, p. 218, and by Whiten, ch. 8.9, p. 220. ED.