Review

Where do mirror neurons come from?

Cecilia Heyes

All Souls College & Department of Experimental Psychology, University of Oxford, All Souls College, High Street, Oxford OX1 4AL, UK

ABSTRACT

Debates about the evolution of the 'mirror neuron system' imply that it is an adaptation for action understanding. Alternatively, mirror neurons may be a byproduct of associative learning. Here I argue that the adaptation and associative hypotheses both offer plausible accounts of the origin of mirror neurons, but the associative hypothesis has three advantages. First, it provides a straightforward, testable explanation for the differences between monkeys and humans that have led some researchers to question the existence of a mirror neuron system. Second, it is consistent with emerging evidence that mirror neurons contribute to a range of social cognitive functions, but do not play a dominant, specialised role in action understanding. Finally, the associative hypothesis is supported by recent data showing that, even in adulthood, the mirror neuron system can be transformed by sensorimotor learning. The associative account implies that mirror neurons come from sensorimotor experience, and that much of this experience is obtained through interaction with others. Therefore, if the associative account is correct, the mirror neuron system is a product, as well as a process, of social interaction.

© 2009 Elsevier Ltd. All rights reserved.

Contents

The remarkable thing about a mirror neuron is that it fires not only when a monkey is performing an action, such as picking up a raisin using a precision grip, but also when the monkey passively observes a similar action performed by another agent. Neurons with this capacity to match observed and executed actions have been found in the premotor and parietal cortex of the monkey (Fogassi et al., 2005; Gallese et al., 1996), and there is evidence of a similar system in the human brain (Gazzola and Keysers, 2009; Iacoboni et al., 1999). Mirror neurons seem to bridge the gap between one agent and another; to represent 'my action' and 'your action' in the same way. Therefore, it has been suggested that mirror neurons are the key to explaining many aspects of social

E-mail address: Cecilia.Heyes@all-souls.ox.ac.uk.
URL: http://www.ceciliaheyes.co.uk/.

Table 1
Adaptation.

A characteristic of an organism, C, is an adaptation for a particular function, F, if C evolved because it helped organisms to do F (Sober, 2008; Williams, 1966). Therefore, the hypothesis that mirror neurons (C) are an adaptation for action understanding (F) concerns the origins, rather than the current utility, of mirror neurons. It asserts that a certain process – genetic evolution – produced mirror neurons, and that they were favoured by natural selection because they supported action understanding. Not all biological characteristics are adaptations. Many are 'byproducts'—characteristics that result from adaptations, but were not 'foreseen', or specifically favoured, by natural selection. For example, it is likely that the high calcium content of bones is an adaptation for skeletal strength, whereas the whiteness of bones is a byproduct of selection for calcium content. Adaptations are specialised for a particular function; they solve a specific problem in an efficient, reliable and precise way. Byproducts can have a range of beneficial effects, but they are not specialised for a particular function. We can ask two questions about any characteristic: where does it come from? What does it do? In the case of adaptations, both questions can be answered in the same way, by identifying the characteristic's function. In the case of byproducts, two answers are needed; the origins of a byproduct are not explained by its current effects.

Table 2
Associative learning.

Associative learning is a form of learning that results from exposure to a relationship between two events. 'Conditioning procedures' arrange different types of relationship between events. Research examining the effects of conditioning procedures on animal behaviour has shown that associative learning depends on 'contiguity' – the closer the two events occur in time, the stronger the association – and 'contingency' – there needs to be a correlation or predictive relationship between them. For example, the probability of the second event occurring must be higher when the first event has occurred than when the first event has not occurred. Both psychological and neural models of associative learning (Schultz and Dickinson, 2000) suggest that the change in behaviour results primarily from the strengthening of existing connections between event representations. In the neural case, this consists in the facilitation of synaptic transmission.

The associative hypothesis considered in this article suggests that associative learning is responsible for changes in the behaviour of neurons rather than whole animals. In discussions of neuronal behaviour, associative learning is often called 'Hebbian learning', but the two terms have different connotations. Hebb's (1949) famous dictum, 'neurons that fire together wire together', implies that contiguity is sufficient for the strengthening of connections. To say that mirror neurons are products of associative learning implies that contingency is also necessary for their formation, and, more broadly, that the extensive literature on associative learning in animals can be used to predict and explain the conditions in which neurons acquire mirror properties. Therefore, the 'Hebbian account' of mirror neurons (Keysers and Perrett, 2004), which was formulated after the associative hypothesis, does not draw on the same body of theory and evidence, or make the same predictions, as the associative model (Cook et al., in press).

Associative learning is a task-general process of learning, which has been documented in a wide range of vertebrate and invertebrate species, including not only rats and pigeons (model species), but also humans, monkeys, cats, ferrets, bony fish, ants, bees and sea slugs (MacPhail, 1996). This taxonomic distribution implies that the mechanisms of associative learning are phylogenetically ancient.

cognition, including the ability to understand the actions of others (Rizzolatti et al., 2001), to 'read minds' (Gallese and Goldman, 1998), to imitate (Iacoboni et al., 1999; Rizzolatti et al., 2001) and to communicate using gestures and speech (Rizzolatti and Arbib, 1998; Rizzolatti and Craighero, 2004). It has even been proposed that mirror neurons are 'the driving force behind "the great leap forward" in human evolution' (Ramachandran, 2000).

Where do mirror neurons come from? How can we explain their capacity to match observed with executed actions? One possibility is that mirror neurons are an adaptation—a characteristic that evolved to fulfil a particular function (see Table 1). The 'adaptation hypothesis' implicit in many discussions of mirror neurons (e.g. Rizzolatti and Arbib, 1998; Rizzolatti and Craighero, 2004) suggests that they are an adaptation for action understanding; that mirror neurons were favoured by natural selection because they helped individuals to understand what others were doing. According to this view, monkeys and humans are born with mirror neurons, or experience plays a relatively minor role in their development. Sensory experience (observing an action; Ferrari et al., 2005) or motor experience (performing an action; Calvo-Merino et al., 2006) may trigger or facilitate the development of mirror neurons, but their capacity to match observed with executed actions is genetically inherited.

An alternative possibility, which has received relatively little attention, is that mirror neurons are a product of associative learning—the same kind of learning that produces Pavlovian conditioning (see Table 2). This 'Associative Sequence Learning' or 'associative hypothesis' (Heyes, 2001, 2005; Heyes and Ray, 2000) suggests that each mirror neuron is forged through sensorimotor experience—correlated experience of observing and executing the same action (Fig. 1). This kind of learning creates the matching properties of mirror neurons in the course of individual development. The motor neurons that become mirror neurons, and the mechanisms that mediate associative learning, are products of evolution, but motor neurons and associative learning did not evolve for the 'purpose' of producing mirror neurons.

The adaptation and associative hypotheses do not represent a nature-nurture dichotomy – each acknowledges genetic and experiential contributions to mirror neuron development – but they have very different implications. The adaptation hypothesis asserts, and the associative hypothesis denies, that to explain the defining feature of mirror neurons – their capacity to match observed with executed actions – we can refer to genetic evolution. Furthermore, while both hypotheses allow that sensory experience and motor experience may facilitate mirror neuron development, the associative account does, and the adaptation account does not, regard sensorimotor experience as crucial.

In this review, I argue that the adaptation and associative hypotheses both offer plausible accounts of the origin of mirror neurons, explaining their basic and more subtle properties. However, the associative hypothesis currently has three advantages. First, it provides a straightforward, empirically testable explanation for the differences between monkeys and humans that have led some researchers to question the existence of a 'mirror neuron system'. Second, it is consistent with evidence indicating that mirror neurons contribute to a range of social cognitive functions, but do not play a dominant, specialised role in action understanding. Finally, the associative hypothesis is supported by recent data showing that, even in adulthood, the mirror neuron system can be reconfigured by sensorimotor learning.

1. Properties of mirror neurons in monkeys

Matching is the most basic property of mirror neurons—they fire when the monkey observes and performs similar actions. The adaptation hypothesis explains this matching in terms of the putative function of mirror neurons (see Table 1): the evolution of neurons that respond during observation and execution of similar actions has been favoured by natural selection because, in contrast with, for example, neurons that fire when one action is observed and different action is executed, they enable the monkey to understand observed actions by matching them to the monkey's motor plans for the same actions. On this view, the matching property of mirror neurons promotes action understanding because it puts the observer in the same causal state as the actor. Observation activates the same motor plan that is producing the
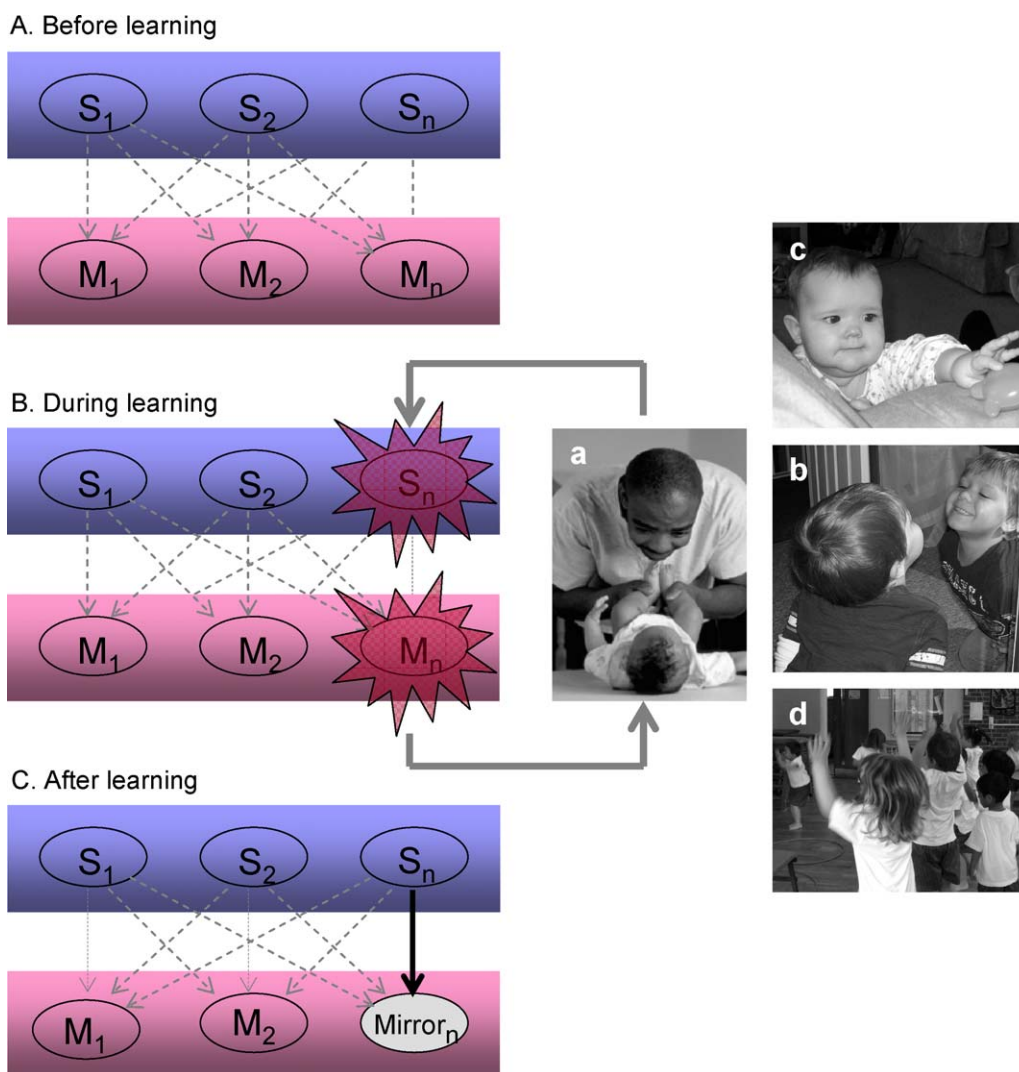
A. Before learning

B. During learning

C. After learning

Fig. 1. Mirror neurons from associative learning. The associative hypothesis (or 'Associative Sequence Learning' hypothesis; Heyes, 2001, 2005; Heyes and Ray, 2000) proposes that mirror neurons are acquired in the following way. Before learning (A), sensory neurons in the superior temporal sulcus, which are responsive to different high-level visual properties of an observed action (S_1, S_2, S_n; Oram and Perrett, 1994, 1996) are weakly and unsystematically connected (dashed arrows) to some motor neurons in the premotor (Rizzolatti et al., 1988) and parietal cortices (M_1, M_2, M_n; Gallese et al., 2002), which discharge during the execution of actions with different high-level properties. For example, S_1, which fires during observation of a precision grip, is weakly connected to both M_1 and M_2, which discharge during execution of a precision grip and a power grip, respectively. The kind of learning that produces mirror neurons (B) occurs when there is correlated (i.e. contiguous and contingent, Table 2) activation of sensory neurons and motor neurons that are each responsive to similar actions. For example, when an adult imitates an infant's facial movements (a), there might be correlated activation of neurons that are responsive to the observation (S_n) and execution (M_n) of lip protrusion. Correlated activation of S_n and M_n increases the strength of the connection between them, so that activation of S_n is propagated to M_n. Therefore, after learning (C), M_n is active, not only during execution of lip protrusion, but also, via its connection with S_n, during observation of lip protrusion, i.e. M_n has become a lip protrusion mirror neuron. Correlated activation of sensory and motor neurons encoding the same property of action occurs not only when we are imitated (a), but also when we use optical mirrors (b), watch our own actions (c), and observe others during the kind of synchronous activities involved in sports and dance training (d) (Ray and Heyes, in press). Correlated activation of visual and motor neurons can also be produced indirectly by accompanying sounds. When the same sound (e.g. smacking) has been heard during observation and execution of an action (e.g. lipsmacking), hearing the sound will activate both visual and motor neurons encoding that action (Heyes and Ray, 2000; Keysers et al., 2003).

actor's behaviour, and, because the observer knows the typical outcome of this motor plan, enables the observer to understand what the actor is doing (Di Pellegrino et al., 1992; Rizzolatti et al., 1996; Rizzolatti and Sinigaglia, 2008).

The associative hypothesis explains the matching property of mirror neurons with reference to the monkey's past experience. For example, whenever a monkey performs a grasping action with visual guidance, the activation of motor neurons (involved in the performance of grasping) and visual neurons (involved in the visual guidance of grasping) is correlated. Through associative learning, this correlated activation gives the grasping motor neurons additional, matching properties; they become mirror neurons, firing not only when grasping is executed, but also when it is observed (Fig. 1).

1.1. Action types

Research has focussed on mirror neurons encoding hand actions, but 'mouth mirror neurons' have also been reported (Ferrari et al., 2003). These cells discharge when, for example, the monkey sucks juice from a syringe and when it sees an experimenter sucking juice from a syringe. The adaptation hypothesis accounts for the existence of mouth as well as hand mirror neurons by assuming that, in the ancestors of extant monkeys, reproductive fitness was enhanced by mirror neuron-mediated understanding of both orofacial and manual actions. The associative hypothesis suggests that, like hand mirror neurons, mouth mirror neurons acquire their matching properties through correlated experience of observing and executing similar actions.

However, in the case of mouth mirror neurons, the monkey gets this experience, not merely from watching his own actions, but from social sources; for example, when mouth movements such as grasping, lip protrusion and tongue protrusion are copied by a conspecific or a human trainer (Paukner and Anderson, 2005; Paukner et al., 2009; Voelkl and Huber, 2007), and when similar sounds accompany the observation and execution of sucking, breaking and lipsmacking (see Fig. 1).

1.2. Context effects

Two studies show that the firing of mirror neurons can be modulated by the context in which an action is observed. These context effects are important because they have been interpreted as evidence that mirror neurons are precisely adapted for action understanding. In the first study (Umilta et al., 2001), monkeys saw a human hand reaching and grasping an object (the visible condition), or they saw the hand reaching but then disappearing behind a screen (the hidden condition). Some 'grasping' mirror neurons fired in the hidden condition, but only if, before the reaching movement started, the monkey viewed an object at the location hidden behind the screen. The adaptation hypothesis explains this effect of object viewing by noting that an efficient system for action understanding must be able to function in a visually noisy environment, where interposition might obscure the final part of a grasping action. The experimental situation simulates this kind of natural environment only when the monkey knows there is an object behind the screen. Under the adaptation hypothesis, the fact that half of the mirror neurons that fired in the visible condition did not fire in the hidden condition could be regarded as a positive feature, yielding a probabilistic estimate of what the agent is doing. However, this feature has been attributed to unavoidable weaknesses in the system (Umilta et al., 2001).

The associative hypothesis offers a unified explanation for the modulating effect of object viewing and the failure of half of the mirror neurons to respond in the hidden condition. It attributes both of these effects to stimulus generalisation (Pearce, 1987)—the tendency of conditioned responses (mirror neuron firing) to be proportional in magnitude to the physical similarity between the current stimulus (the action sequence observed in the experiment) and the learning stimuli (the action sequences observed before the experiment, while the neurons were acquiring their properties through associative learning). It is likely that the learning stimuli included visible grasping and a visible object. Therefore, hiding the grasping component reduces the similarity between the learning and test stimuli, resulting in less mirror neuron activity in the hidden than in the visible condition. Removing the object reduces the similarity yet further, explaining why there is minimal mirror neuron activity when the object is not viewed in the hidden condition (Heyes, 2005).

In the second study of context effects, it was found that different populations of parietal neurons fire when a monkey grasps an object that is subsequently eaten and when it grasps an object that is subsequently placed in a container (Fogassi et al., 2005). Some of these neurons have matching visual properties, e.g. they respond when the monkey observes and executes grasping before eating, but not when it observes or executes grasping before placing. The adaptation hypothesis takes this to indicate that the firing of mirror neurons can be modulated by events that occur afterwards – by placement of a grasped object in the mouth or a container – and explains this by suggesting that the evolutionary function of mirror neurons is to enable an observing animal to infer an actor's intentions.

The associative hypothesis interprets the same findings as showing that, as a result of conditional learning (Bonardi, 1998), mirror neuron firing can be modulated by events that occur with,

or before, the primary eliciting stimulus (grasping). In this study, a container was always presented in trials involving grasping before placing, and never in trials involving grasping before eating. Therefore, the presence or absence of a container could become a conditional cue differentially activating two groups of grasping mirror neurons.

2. Is there a human 'mirror neuron system'?

Research using transcranial magnetic stimulation (TMS) and functional magnetic resonance imaging (fMRI) has been interpreted by many researchers as evidence that humans also have mirror neurons or, more broadly, a 'mirror neuron system'. TMS studies show that passive observation of arm, hand and finger movements results in selective activation of the muscles involved in producing the observed movement (Aziz-Zadeh et al., 2002; Catmur et al., 2007; Fadiga et al., 1995; Gangitano et al., 2001, 2004; Strafella and Paus, 2000). In addition, a number of fMRI experiments have compared the cortical areas activated by (1) execution of action without visual feedback, and (2) passive observation of similar actions, and found that they overlap in regions broadly homologous with those where mirror neurons have been found in monkeys (Aziz-Zadeh et al., 2006; Buccino et al., 2004a; Jonas et al., 2007; Gazzola et al., 2006, 2007a,b; Grèzes et al., 2003; Iacoboni et al., 1999; Molnar-Szakacs et al., 2005; Shmuelof and Zohary, 2006; Vogt et al., 2007). This evidence of mirror activation in humans has been augmented by studies indicating observation–execution overlap at the level of individual voxels in single subjects (Gazzola and Keysers, 2009), and cross-modal repetition suppression in two classical mirror areas: the inferior frontal gyrus (Kilner et al., 2009) and inferior parietal lobe (Chong et al., 2008). For example, blood oxygen level dependent (BOLD) responses to action observation were attenuated following execution of a similar, relative to dissimilar, actions.

However, the existence of a human mirror neuron system has recently been challenged. Examining the evidence closely, critics argue that (1) mirror activation in humans does not have key features in common with mirror neuron activity in monkeys, and (2) brain areas identified with the mirror neuron system do not play a major, specialised role in action understanding. The next two sections examine each of these arguments in turn, and suggest that they provide reasons to doubt the existence of a mirror neuron system only if one assumes that this system is an adaptation for action understanding. These arguments are based on evidence that is entirely compatible with the associative hypothesis.

2.1. Monkeys versus humans

Four putative differences between monkeys and humans have raised questions about the existence of a mirror neuron system. First, due to the weak spatial resolution of fMRI, it has not been established that humans, like monkeys, have single neurons that discharge during observation and execution of similar actions (Dinstein et al., 2008), and some repetition suppression studies addressing this problem have failed to find cross-modal effects (Dinstein et al., 2007; Lingnau et al., 2009). Second, human mirror activation occurs, not only in homologous areas, but also in regions outside those where mirror neurons have been reported in monkeys (Aziz-Zadeh et al., 2006; Gazzola et al., 2007a,b; Grèzes et al., 2003; Molnar-Szakacs et al., 2005). (This difference may be exaggerated by the methods currently used to study humans and monkeys; fMRI samples the whole cortex while single-unit recording necessarily targets small areas.) Third, the vast majority of monkey mirror neurons are responsive to actions-on-objects, whereas human mirror activation occurs readily in response to gestures as well as actions-on-objects (Hickok, 2009; Turella et al.,

2008). Finally, for a mirror neuron to fire, it appears to be necessary for the monkey to be viewing the actor's whole body (Nelissen et al., 2005), but mirror activation commonly occurs when a human subject is observing the movements of an isolated hand (Turella et al., 2008).

To establish conclusively whether humans have mirror neurons it will be necessary to use single cell recording in human subjects (Iacoboni, 2008), and to resolve methodological issues relating to the use of repetition suppression techniques (Bartels et al., 2008). The adaptation and associative hypotheses both anticipate that the existence of mirror neurons in humans will be confirmed by these methods. However, to make this prediction plausible, the adaptation hypothesis would have to argue that monkeys and humans have been subject to different selection pressures, leading to divergent evolutionary trajectories. In contrast, the associative hypothesis provides a simple, testable explanation for the reported differences between monkeys and humans. It suggests that the human and monkey mirror neuron systems differ because humans receive a great deal more correlated experience of observing and executing similar actions. The gestures of human infants are constantly imitated by adults (Jones, 2009) and, unlike monkeys, humans observe their gestures in optical mirrors, and as they are simultaneously performed by others in the context of dance and sports (Fig. 1; Heyes and Ray, 2000; Ray and Heyes, in press). Human children are also trained by adults to focus their attention on potentially informative acts (Gergely et al., 2007), and, in many cultures, observe the movements of isolated body parts in video displays. If this associative account of the differences between humans and monkeys is correct, it should be possible to make the responsivity of the monkey mirror neuron system more like that of humans through training; by giving monkeys more of the correlated sensorimotor experience that children typically receive during development.

The associative hypothesis is also consistent with evidence that motor mirror activation occurs in brain areas other than the ventral premotor and posterior parietal cortex. If mirror neurons are forged by associative learning, any area that processes both sensory and motor information has the potential to show mirror effects, given sufficient correlated experience of observing and executing similar actions (Catmur et al., in press).

2.2. Action understanding

Recent reviews emphasise that there has been very little empirical research examining the contribution of the mirror neuron system to action understanding, and argue that, at best, the available data indicate a minor, non-specialised, rather than a major, specialised contribution (Chaminade, 2009; Hickok, 2009; Kilner et al., 2007; Mahon and Caramazza, 2008; Scott et al., 2009). For example, a minor role is indicated by studies showing that, in patients with brain damage, impairments in action production are sometimes, but not consistently, associated with impairments in action understanding (Mahon, 2008), and that mirror activation appears to be dissociated from action understanding when humans observe the actions of animals, or when the responsivity of the mirror neuron system is altered by training (Buccino et al., 2004b; Catmur et al., 2007; Hickok, 2009). A non-specialised role is indicated by experiments using repetitive TMS to disrupt mirror neuron system activity, and showing that this can interfere with performance in both action understanding (Podric and Hamilton, 2006) and imitation (Catmur et al., 2009; Heiser et al., 2003) tasks.

These findings have many important implications, but they raise questions about the existence of a human mirror neuron system only when one assumes that, if it exists, this system is an adaptation for action understanding rather than a byproduct of associative learning (see Table 1). An adaptation tends to be tailored precisely to fulfil a particular function (Williams, 1966). In contrast, while a byproduct may have a variety of effects or uses, one would not normally expect it to be either necessary or sufficient for the fulfilment of any particular adaptive function. Therefore, the associative hypothesis is consistent with both the suggestion that the mirror neuron system contributes in some way to a variety of social cognitive functions (Iacoboni, 2009), and studies showing that, in each case, mirror activation provides just one source of input to a complex system (Hickok, 2009; Mahon and Caramazza, 2008).

3. The development of mirroring

3.1. Imitation in newborns?

Newborn infants have had minimal opportunity for sensorimotor learning. Therefore evidence of mirror neuron system activity in newborn monkeys or humans would provide strong support for the adaptation hypothesis over the associative hypothesis (Meltzoff and Decety, 2003; Rizzolatti et al., 2002). No direct evidence of this kind is available. However, it has been argued that imitation depends on the mirror neuron system, and therefore that studies reporting imitation in neonates provide indirect evidence that the mirror neuron system is present at birth (Lepage and Theoret, 2007). The logic of this argument is sound, but the evidence is highly controversial. Recent reviews suggest that human neonates reliably match only one action, tongue protrusion; that this effect is transitory; and that it is due, not to imitation, but to a non-specific arousal mechanism (Anisfeld, 1996; Jones, 2009; Ray and Heyes, in press). Similar results were obtained in a recent study of newborn monkeys: of five actions tested at four postpartum intervals, a behavioural matching effect occurred only on Day 3 and only for tongue protrusion and a related behaviour, lipsmacking (Ferrari et al., 2006). Thus, the data on imitation in newborns do not provide compelling evidence that mirror neurons are present at birth.

3.2. Effects of experience

Studies involving expertise and training in various action domains show that the activity of the mirror neuron system is modulated by experience. The primary evidence of experiential modulation in monkeys comes from a study in which 'tool-responding mirror neurons' were found in animals that had received many opportunities to observe humans using a stick or pliers as a tool, and subsequently to grasp the object themselves (Ferrari et al., 2005). These neurons discharged when the monkey observed the use of a stick or pliers, and when the monkey grasped the object with its own hand or mouth. Earlier experiments, involving monkeys with little sensory or sensorimotor experience of tool-use, did not find these neurons (Rizzolatti and Arbib, 1998; Rizzolatti et al., 1996).

Studies involving musicians and dancers indicate that experience also modulates the activity of the human mirror neuron system (D'Ausilio et al., 2006; Cross et al., 2006; Margulis et al., 2009). For example, there is more mirror activation in pianists than in non-pianists during observation of piano-playing finger movements (Haslinger et al., 2005), and in classical ballet dancers than in capoeira dancers during observation of ballet movements (Calvo-Merino et al., 2005).

Both the adaptation and associative hypotheses are consistent with this evidence of experiential modulation of the mirror neuron system. Both accounts allow that repeatedly observing (sensory experience; Ferrari et al., 2005) or performing (motor experience; Calvo-Merino et al., 2006) an action might contribute to mirror neuron development, but the associative hypothesis makes the

Table 3
Counter-mirror neurons.

Research on associative learning in animals has indicated that it depends exclusively on the temporal features of the relationship between two events (contiguity and contingency; see Table 2), and not on their 'similarity' (Hall, 1994). Therefore, the associative hypothesis suggests that there is nothing about the internal process that produces mirror neurons – associative learning – that makes it more likely to generate mirror neurons than 'counter-mirror neurons', i.e. neurons that fire during observation of one action and execution of a different action. It predicts that the prevalence of counter-mirror neurons will depend on the kind of experience provided by the developmental environment: correlated experience of observing and executing similar actions will produce mirror neurons, and correlated experience of observing and executing dissimilar actions will produce counter-mirror neurons.

In human developmental environments there are at least two major sources of the kind of experience that would produce counter-mirror neurons: coordinated instrumental action (e.g. I grasp an object while you release it, or push while you pull), and social control behaviour (e.g. I duck when you punch, cringe while you preen). Consistent with the predictions of the associative hypothesis, there is now evidence of counter-mirror activation in humans, not only after laboratory training (Catmur et al., 2007, 2008, in press) and atypical development (Gazzola et al., 2007a,b), but also in each of these contexts following typical development (Newman-Norlund et al., 2007; Tiedens and Fragale, 2003; Van Schie et al., 2008).

distinctive claim that a third type of experience, sensorimotor experience, plays an inductive role; that mirror neurons are created by experience of observing and executing similar actions. In each of the studies reviewed above, experts had more sensorimotor experience of the action domain than non-experts, but they also had more sensory experience and/or more motor experience. Therefore, although the results of these studies are entirely consistent with the associative hypothesis, they do not uniquely favour that account.

3.3. Effects of sensorimotor experience

Evidence favouring the associative over the adaptation hypothesis comes from a series of experiments showing specifically that sensorimotor experience can enhance (Press et al., 2007), abolish (Heyes et al., 2005) and even reverse (Catmur et al., 2007, 2008, in press) mirror activation in human subjects. For example, the mirror neuron system is usually more responsive to hand than foot movements. However, watching foot movements while performing hand movements, and vice versa, can reverse this dominance relationship in the premotor and parietal cortex during subsequent action observation (Catmur et al., 2008). Similarly, incompatible sensorimotor training, in which index finger abduction was paired with little finger abduction, and vice versa, reversed TMS-induced muscle-specific activation during action observation (Catmur et al., 2007). This training induced 'counter-mirror' activation; for example, the observation of index finger movement produced more activity in little finger than in index finger muscles, implying that observation of index finger movement had activated neurons in the premotor cortex that represent the alternative, nonmatching action (Catmur et al., in press; see Table 3).

These training effects do not provide conclusive evidence that the mirror neuron system normally acquires its matching properties through sensorimotor learning. However, they are exactly the effects one would expect if, as the associative hypothesis suggests, mirror neurons are forged by sensorimotor experience, and they are hard to reconcile with the adaptation hypothesis for two reasons. First, one would expect the development of an adaptation to be buffered against naturally occurring variations in the environment (Cosmides and Tooby, 1994; Pinker, 1997), but in these studies there was no sign of resistance to perturbation. Incompatible sensorimotor experience occurs naturally during coordinated instrumental action and in social control situations (see Table 3), and, in the experiments described above, a brief period of this kind of experience (0.5–2.5 h) induced changes in mirror system properties that lasted for at least 24 h. Second, these training studies indicate that sensorimotor experience can transform rather than merely deform the mirror neuron system. Unlike the effects of wearing a cast on the development of normal gait, or the effects of visual deprivation on the development of orientation cells, the effect of incompatible sensorimotor experience was not just to weaken, or to reduce the selectivity of, the mirror neuron

system, but to make it operate in an entirely different way. The degree of selectivity was maintained, while the pattern of selectivity was reversed.

3.4. Sources of sensorimotor experience

Human infants like to watch their own hands in motion, and it has been suggested that this visual bias promotes the development of mirror neurons through sensorimotor learning (Del Giudice et al., 2009). This interesting suggestion is wholly compatible with the associative hypothesis. If the visual bias, or related aspects of spontaneous grasping behaviour, evolved specifically to promote mirror neuron development via sensorimotor learning, then it could be argued that the mirror neuron system is both an adaptation and a product of associative learning. In that case, the matching properties of mirror neurons would be due to both the monkey's past experience of observing and executing actions and to their role in action understanding; the former would provide a proximal, and the latter a distal or ultimate, explanation. However, there is no evidence that the tendency of human infants to watch their own hands evolved specifically to promote mirror neuron development. Indeed, if this bias is an adaptation at all, it is likely to be an adaptation for acquiring precise visuomotor control over hand and finger movements; a capacity that is necessarily more primitive than the ability to understand such actions when they are observed. Therefore, any beneficial effect of the visual bias on mirror neuron development would be consistent with both aspects of the associative account—the claim that mirror neurons are a product of associative learning, and the denial that mirror neurons have been a specific focus of selection pressure, and therefore that their matching properties can be explained by their function.

By appealing to 'exaptation' (Gould and Vrba, 1982) or 'Baldwin effects' (Weber and Depew, 2003), it may be possible to formulate other adaptation-association hybrid hypotheses, but this is unlikely to be a fruitful direction for future research. In a system that had already evolved for precise visuomotor control, associative learning would be a flexible and reliable source of mirror neurons. It would yield mirror neurons coding for currently important actions, in a range of environments. Therefore, any mutation that had an impact on mirror neuron development would be unlikely to enhance reproductive fitness.

4. Conclusions and future directions

The associative hypothesis could be tested more decisively by examining the effects of incompatible sensorimotor experience on mirror neurons in monkeys. This hypothesis predicts that if monkeys were given experience in which, for example, observation of a precision grip is reliably correlated with execution of a power grip, then some of the mirror neurons that were previously responsive to observation and execution of a precision grip would become counter-mirror neurons, discharging during observation of a precision grip and execution of a power grip.

The associative hypothesis also predicts that mirror neurons will be found in other nonhuman species. Animals from a wide range of taxa are capable of associative learning (see Table 2). Therefore, if the associative account is correct, one would expect to find naturally occurring mirror neurons in animals who receive appropriate sensorimotor experience via self-observation or synchronous action, or as a result of vocalisations that accompany both the observation and execution of specific actions. In addition, it should be possible, via sensorimotor training in the laboratory, to produce mirror neurons in animals that would not normally have them.

Another clear priority for future research is to establish, using single-cell recording, whether humans not only show mirror activation, but also have mirror neurons—single cells that fire during the observation and execution of similar actions. The associative hypothesis predicts that if humans have mirror neurons, they will also have counter-mirror neurons (see Table 3). Instead of putting the observer in the actor's shoes, counter-mirror neurons would activate in the observer a motor representation that contrasts with the one generating the actor's behaviour. Neurons of this kind could, in principle, play some role in action understanding, but not by 'direct matching', i.e. putting the observer in the same causal state as the actor.

Ever since mirror neurons were discovered, there has been a great deal of discussion about their potential effects on social cognition, but very little empirical work investigating these effects. Drawing attention to this gap between theory and evidence, and to differences between human and monkey mirror activation, some researchers have questioned the existence of a human mirror neuron system. Consequently, a gulf has opened between mirror neuron 'believers' and 'sceptics'; a difference of opinion that could obstruct further research on both their fundamental properties and social cognitive effects. The associative hypothesis bridges this gulf, embracing both positive evidence and critical analyses. It explains why there are differences between human and monkey mirror activation; why this activation is found outside the original mirror areas of the premotor and parietal cortex; and why we should not necessarily expect the effects of mirror activation to be specialised for a particular social cognitive function.

The associative hypothesis is wholly compatible with the existence of mirror neurons in monkeys and humans, but it suggests that we should be more careful in using the term 'mirror neuron system'. 'System' often refers to a set of interacting parts organised by an external agent or process for a specific purpose. If mirror neurons are an adaptation – if they were 'organised' by natural selection – they constitute a system in this sense. However, if mirror neurons are a byproduct of associative learning, they may have a variety of beneficial effects, but they do not have a specific purpose (see Table 1), and their combined effects could be due to self-organisation. Therefore, use of the term 'mirror neuron system' should not obscure the fact that further work is needed to discover how, and to what extent, mirror neuron activity is organised.

Research on the effects of mirror neurons is just beginning. If mirror neurons are produced by associative learning, they are unlikely to be 'the driving force behind the "great leap forward" in human evolution' (Ramachandran, 2000), or the kind of unsplitable atoms that might provide a key to understanding what makes humans special (Blakeslee, 2006; Heyes, in press). However, careful empirical work could show that mirror neurons, and counter-mirror neurons, contribute to a variety of social cognitive functions, including action understanding, action prediction, imitation, language processing and mentalising. The challenge will be to discover exactly what they do and how they do it; to characterise the parts played by mirror neurons in the complex cognitive systems supporting human sociality.

If the associative hypothesis is correct, mirror neurons not only support, but are supported by, human sociality. They come, not from evolution, but from sensorimotor experience, and much of this experience is obtained through interaction with others. Therefore, the associative account implies that mirror neurons are a product, as well as a process, of social interaction.

Acknowledgements

References

Anisfeld, M., 1996. Only tongue protrusion modelling is matched by neonates. Developmental Review 16, 149–161.

Aziz-Zadeh, L., Koski, L., Zaidel, E., Mazziotta, J., Iacoboni, M., 2006. Lateralization of the human mirror neuron system. Journal of Neuroscience 26, 2964–2970.

Aziz-Zadeh, L., Maeda, F., Zaidel, E., Mazziotta, J., Iacoboni, M., 2002. Lateralization in motor facilitation during action observation: a TMS study. Experimental Brain Research 144, 127–131.

Bartels, A., Logothetis, N.K., Moutoussis, K., 2008. fMRI and its interpretations: an illustration on directional selectivity in area V5/MT. Trends in Neurosciences 31, 444–453.

Blakeslee, S., 2006. Cells that read minds. The New York Times, January 10. http://www.nytimes.com/2006/01/10/science/10mirr.html.

Bonardi, C., 1998. Conditional learning: an associative analysis. In: Holland, P.C., Schmajuk, N.A. (Eds.), Associative Learning and Cognition in Animals: Occasion Setting. American Psychological Association, Washington, DC, pp. 37–67.

Buccino, G., et al., 2004. Neural circuits underlying imitation learning of hand actions: an event-related fMRI study. Neuron 42, 323–334.

Buccino, G., et al., 2004. Neural circuits involved in the recognition of actions performed by nonconspecifics: an fMRI study. Journal of Cognitive Neuroscience 16, 114–126.

Calvo-Merino, B., Glaser, D.E., Grezes, J., Passingham, R.E., Haggard, P., 2005. Action observation and acquired motor skills: an fMRI study with expert dancers. Cerebral Cortex 15, 1243–1249.

Calvo-Merino, B., Grezes, J., Glaser, D.E., Passingham, R.E., Haggard, P., 2006. Seeing or doing? Influence of visual and motor familiarity in action observation. Current Biology 16, 1905–1910.

Catmur, C., Gillmeister, H., Bird, G., Liepelt, R., Brass, M., Heyes, C.M., 2008. Through the looking glass: counter-mirror activation following incompatible sensorimotor learning. European Journal of Neuroscience 28, 1208–1215.

Catmur, C., Mars, R., Rushworth, M., Heyes, C.M., in press. Making mirrors: premotor cortex stimulation enhances mirror and counter-mirror motor facilitation effects.

Catmur, C., Walsh, V., Heyes, C.M., 2007. Sensorimotor learning configures the human mirror system. Current Biology 17, 1527–1531.

Catmur, C., Walsh, V., Heyes, C.M., 2009. Associative sequence learning: the role of experience in the development of imitation and the mirror system. Philosophical Transactions of the Royal Society B 364, 2369–2380.

Chaminade, T., 2009. From mirror neurons to human cognition: the missing links. In: Hatfield, G. (Ed.), The Evolution of Mind, Brain, and Culture. Penn Museum Press.

Chong, T., Cunnington, R., Williams, M., Kanwisher, N., Mattingley, J., 2008. fMRI adaptation reveals mirror neurons in human inferior parietal cortex. Current Biology 18, 1576–1580.

Cook, R., Press, C., Heyes, C.M., in press. Is the acquisition of automatic imitation sensitive to sensorimotor contingency?

Cosmides, L., Tooby, J., 1994. Beyond intuition and instinct blindness: toward an evolutionarily rigorous cognitive science. Cognition 50, 41–77.

Cross, E.S., Hamilton, A.F., Grafton, S.T., 2006. Building a motor simulation de novo: observation of dance by dancers. Neuroimage 31, 1257–1267.

D'Ausilio, A., Altenmuller, E., Olivetti, B.M., Lotze, M., 2006. Cross-modal plasticity of the motor cortex while listening to a rehearsed musical piece. European Journal of Neuroscience 24, 955–958.

Del Giudice, M., Manera, V., Keysers, C., 2009. Programmed to learn? The ontogeny of mirror neurons. Developmental Science 12, 350–363.

Dinstein, I., Thomas, C., Behrmann, M., Heeger, D.J., 2008. A mirror up to nature. Current Biology 18, 13–18.

Dinstein, I., Hasson, U., Rubin, N., Heeger, D.J., 2007. Brain areas selective for both observed and executed movements. Journal of Neurophysiology 98, 1415–1427.

Di Pellegrino, G., Fadiga, L., Fogassi, L., Gallese, V., Rizzolatti, G., 1992. Understanding motor events: a neurophysiological study. Experimental Brain Research 91, 176–180.

Fadiga, L., Fogassi, L., Pavasi, G., Rizzolatti, G., 1995. Motor facilitation during action observation: a magnetic stimulation study. Journal of Neurophysiology 73, 2608–2611.

Ferrari, P.F., Gallese, V., Rizzolatti, G., Fogassi, L., 2003. Mirror neurons responding to the observation of ingestive and communicative mouth actions in the monkey ventral premotor cortex. European Journal of Neuroscience 17, 1703–1714.

Ferrari, P.F., Rozzi, S., Fogassi, L., 2005. Mirror neuron responding to observation of actions made with tools in monkey ventral premotor cortex. Journal of Cognitive Neuroscience 17, 212–226.

Ferrari, P.F., et al., 2006. Neonatal imitation in rhesus macaques. PLoS Biology 4, 1501–1508.

Fogassi, L., Ferrari, P.F., Gesierich, B., Rozzi, S., Chersi, F., Rizzolatti, G., 2005. Parietal lobe: from action organization to intention understanding. Science 308, 662–667.

Gallese, V., Fadiga, L., Fogassi, L., Rizzolatti, G., 1996. Action recognition in the premotor cortex. Brain 119, 593–609.

Gallese, V., Fadiga, L., Fogassi, L., Rizzolatti, G., 2002. Action representation and the inferior parietal lobule. In: Prinz, W., Hommel, B. (Eds.), Common Mechanisms in Perception and Action: Attention and Performance XIX. Oxford University Press, Oxford, pp. 247–266.

Gallese, V., Goldman, A., 1998. Mirror neurons and the simulation theory of mind-reading. Trends in Cognitive Sciences 2, 493–500.

Gangitano, M., Mottagy, F.M., Pascual-Leone, A., 2001. Phase-specific modulation of cortical motor output during movement observation. Neuroreport 12, 1489–1492.

Gangitano, M., Mottaghy, F.M., Pascual-Leone, A., 2004. Modulation of premotor mirror neuron activity during observation of unpredictable grasping movements. European Journal of Neuroscience 20, 2193–2202.

Gazzola, V., Aziz-Zadeh, L., Keysers, C., 2006. Empathy and the somatotopic auditory mirror system in humans. Current Biology 16, 1824–1829.

Gazzola, V., van der Worp, H., Mulder, T., Wicker, B., Rizzolatti, G., Keysers, C., 2007. Aplasics born without hands mirror the goal of hand actions with their feet. Current Biology 17, 1235–1240.

Gazzola, V., Keysers, C., 2009. The observation and execution of actions share motor and somatosensory voxels in all tested subjects: single-subject analyses of unsmoothed fMRI data. Cerebral Cortex 19, 1239–1255.

Gazzola, V., Rizzolatti, G., Wicker, B., Keysers, C., 2007. The anthropomorphic brain: the mirror neuron system responds to human and robotic actions. Neuroimage 35, 1674–1684.

Gergely, G., Egyed, K., Király, I., 2007. On pedagogy. Developmental Science 10, 139–146.

Gould, S.J., Vrba, E.S., 1982. Exaptation—a missing term in the science of form. Paleobiology 8, 4–15.

Grèzes, J., Armony, J.L., Rowe, J., Passingham, R.E., 2003. Activations related to "mirror" and "canonical" neurones in the human brain: an fMRI study. Neuroimage 18, 928–937.

Hall, G., 1994. Pavlovian conditioning: laws of association. In: Mackintosh, N.J. (Ed.), Animal Learning and Cognition. Academic Press, San Diego, pp. 15–43.

Haslinger, B., et al., 2005. Transmodal sensorimotor networks during action observation in professional pianists. Journal of Cognitive Neuroscience 17, 282–293.

Hebb, D., 1949. The Organization of Behavior. Wiley.

Heiser, M., Iacoboni, M., Maeda, F., Marcus, J., Mazziotta, J.C., 2003. The essential role of Broca's area in imitation. European Journal of Neuroscience 17, 1123–1128.

Heyes, C.M., 2001. Causes and consequences of imitation. Trends in Cognitive Sciences 5, 245–261.

Heyes, C.M., 2005. Imitation by association. In: Hurley, S., Chater, N. (Eds.), Perspectives on Imitation: From Mirror Neurons to Memes. MIT Press.

Heyes, C.M., in press. Mesmerizing mirror neurons.

Heyes, C., Bird, G., Johnson, H., Haggard, P., 2005. Experience modulates automatic imitation. Cognitive Brain Research 22, 233–240.

Heyes, C.M., Ray, E.D., 2000. What is the significance of imitation in animals? Advances in the Study of Behavior 29, 215–245.

Hickok, G., 2009. Eight problems for the mirror neuron theory of action understanding in monkeys and humans. Journal of Cognitive Neuroscience 21, 1229–1243.

Iacoboni, M., 2008. Mesial frontal cortex and super mirror neurons. Behavioral and Brain Sciences 31, 30.

Iacoboni, M., 2009. Imitation, empathy and mirror neurons. Annual Review of Psychology 60, 653–670.

Iacoboni, M., Woods, R.P., Brass, M., Bekkering, H., Mazziotta, J.C., Rizzolatti, G., 1999. Cortical mechanisms of human imitation. Science 286, 2526–2528.

Jonas, M., et al., 2007. Do simple intransitive finger movements consistently activate frontoparietal mirror neuron areas in humans? Neuroimage 36 (Suppl. 2), T44–T53.

Jones, S.S., 2009. The development of imitation in infancy. Philosophical Transactions of the Royal Society B 364, 2325–2335.

Keysers, C., et al., 2003. Audiovisual mirror neurons and action recognition. Experimental Brain Research 153, 628–636.

Keysers, C., Perrett, D.I., 2004. Demystifying social cognition: a Hebbian perspective. Trends in Cognitive Sciences 8, 501–507.

Kilner, J.M., Neal, A., Weiskopf, N., Friston, K.J., Frith, C.D., 2009. Evidence of mirror neurons in human inferior frontal gyrus. Journal of Neuroscience 29, 10153–10159.

Kilner, J.M., Friston, K.J., Frith, C.D., 2007. The mirror-neuron system: a Bayesian perspective. NeuroReport 18, 619–623.

Lepage, J.-F., Theoret, H., 2007. The mirror neuron system: grasping others actions from birth? Developmental Science 10, 513–529.

Lingnau, A., Gesierich, B., Caramazza, A., 2009. Asymmetric fMRI adaptation reveals no evidence for mirror neurons in humans. Proceedings of the National Academy of Sciences of the United States of America 106, 9925–9930.

MacPhail, E.M., 1996. Cognitive function in mammals: the evolutionary perspective. Cognitive Brain Research 3, 279–290.

Mahon, B.Z., 2008. Action recognition: is it a motor process? Current Biology 18, 1068–1069.

Mahon, B.Z., Caramazza, A., 2008. A critical look at the embodied cognition hypothesis and a new proposal for grounding conceptual content. Journal of Physiology 102, 59–70.

Margulis, E.H., Mlsna, L.M., Uppunda, A.K., Parrish, T.B., Wong, P.C., 2009. Selective neurophysiologic responses to music in instrumentalists with different listening biographies. Human Brain Mapping 30, 267–275.

Meltzoff, A.N., Decety, J., 2003. What imitation tell us about social cognition: a rapprochement between developmental psychology and cognitive neuroscience. Philosophical Transactions of the Royal Society London B, Biological Sciences 358, 491–500.

Molnar-Szakacs, I., Iacoboni, M., Koski, L., Mazziotta, J.C., 2005. Functional segregation within pars opercularis of the inferior frontal gyrus: evidence from fMRI studies of imitation and action observation. Cerebral Cortex 15, 986–994.

Nelissen, K., Luppino, G., Vanduffel, W., Rizzolatti, G., Orban, G., 2005. Observing others: multiple action representation in frontal lobe. Science 310, 332–336.

Newman-Norlund, R.D., van Schie, H.T., van Zuijlen, A., Bekkering, H., 2007. The mirror neuron system is more active during complementary compared with imitative action. Nature Neuroscience 10, 817–818.

Oram, M., Perrett, D.I., 1994. Responses of anterior superior temporal polysensory (STPa) neurons to biological motion stimuli. Journal of Cognitive Neuroscience 6, 99–116.

Oram, M., Perrett, D.I., 1996. Integration of form and motion in the anterior superior temporal polysensory area (STPa) of the macaque monkey. Journal of Neurophysiology 76, 109–129.

Paukner, A., Anderson, J.R., 2005. Video-induced yawning in stumptail macaques. Biology Letters 2, 36–38.

Paukner, A., Suomi, S.J., Visalberghi, E., Ferrari, P.F., 2009. Capuchin monkeys display affiliation toward humans who imitate them. Science 325, 880–883.

Pearce, J.M., 1987. A model of stimulus generalization in Pavlovian conditioning. Psychological Review 94, 61–73.

Pinker, S., 1997. How the Mind Works. The Penguin Press, London.

Podric, G., Hamilton, A., 2006. Action understanding requires left inferior frontal cortex. Current Biology 16, 524–529.

Press, C., Gillmeister, H., Heyes, C.M., 2007. Sensorimotor experience enhances automatic imitation of robotic action. Proceedings of the Royal Society of London, Biological Sciences 274, 2509–2514.

Ramachandran, V.S., 2000. Mirror neurons and imitation learning as the driving force behind "the great leap forward" in human evolution. Edge. http://www.edge.org/3rd_culture/ramachandran/ramachandran_index.html.

Ray, E.B., Heyes, C.M., in press. Imitation in infancy: the wealth of the stimulus.

Rizzolatti, G., Arbib, M.A., 1998. Language within our grasp. Trends in Neurosciences 21, 188–194.

Rizzolatti, G., Craighero, L., 2004. The mirror neuron system. Annual Review of Neuroscience 27, 169–192.

Rizzolatti, G., Fadiga, L., Gallese, V., Fogassi, L., 1996. Premotor cortex and the recognition of motor actions. Cognitive Brain Research 3, 131–141.

Rizzolatti, G., Fadiga, L., Gallese, V., Fogassi, L., 2002. From mirror neurons to imitation: facts and speculations. In: Meltzoff, A., Prinz, W. (Eds.), The Imitative Mind. Cambridge University Press, Cambridge, pp. 247–266.

Rizzolatti, G., Fogassi, L., Gallese, V., 2001. Neurophysiological mechanisms underlying the understanding and imitation of action. Nature Reviews Neuroscience 2, 661–670.

Rizzolatti, G., et al., 1988. Functional organization of inferior area 6 in the macaque monkey. II. Area F5 and the control of distal movements. Experimental Brain Research 71, 491–507.

Rizzolatti, G., Sinigaglia, C., 2008. Mirrors in the Brain. Oxford University Press, Oxford.

Schultz, W., Dickinson, A., 2000. Neuronal coding of prediction errors. Annual Review of Neuroscience 23, 473–500.

Scott, S.K., McGettigan, C., Eisner, F., 2009. A little more conversation, a little less action—candidate roles for the motor cortex in speech perception. Nature Reviews Neuroscience 10, 295–302.

Shmuelof, L., Zohary, E., 2006. A mirror representation of others' actions in the human anterior parietal cortex. Journal of Neuroscience 26, 9736–9742.

Sober, E., 2008. Evidence and Evolution: the Logic Behind the Science. Cambridge University Press.

Strafella, A.P., Paus, T., 2000. Modulation of cortical excitability during action observation: a transcranial magnetic stimulation study. Neuroreport 11, 2289–2292.

Tiedens, L.Z., Fragale, A.R., 2003. Power moves: complementarity in dominant and submissive nonverbal behaviour. Journal of Personality & Social Psychology 84, 558–568.

Turella, L., Pierno, A.C., Tubaldi, F., Castiello, U., 2008. Mirror neurons in humans: consisting or confounding evidence? Brain & Language 108, 10–21.

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

Van Schie, H.T., van Waterschoot, B.M., Bekkering, H., 2008. Understanding action beyond imitation: reversed compatibility effects of action observation in imitation and joint action. Journal of Experimental Psychology: Human Perception and Performance 34, 1493–1500.

Voelkl, B., Huber, L., 2007. Imitation as faithful copying of a novel technique in marmoset monkeys. PLoS ONE 2, e611.

Vogt, S., et al., 2007. Prefrontal involvement in imitation learning of hand actions: effects of practice and expertise. Neuroimage 37, 1371–1383.

Weber, B.H., Depew, D.J., 2003. Evolution and Learning: The Baldwin Effect Reconsidered. MIT Press, Cambridge, MA.

Williams, G.C., 1966. Adaptation and Natural Selection. Princeton University Press.