

Mirror neurons: From origin to function

Richard Cook

Department of Psychology, City University London, London EC1R 0JD, United Kingdom

richard.cook.1@city.ac.uk

<http://www.city.ac.uk/people/academics/richard-cook>

Geoffrey Bird

MRC Social, Genetic and Developmental Psychiatry Centre, Institute of Psychiatry, King's College London, London SE5 8AF, United Kingdom

geoff.bird@kcl.ac.uk

<https://sites.google.com/site/geoffbirdlab/>

<http://www.iop.kcl.ac.uk/staff/profile/default.aspx?go=13152>

Caroline Catmur

Department of Psychology, University of Surrey, Guildford, Surrey GU2 7XH, United Kingdom

c.catmur@surrey.ac.uk

http://www.surrey.ac.uk/psychology/people/dr_caroline_catmur/

<http://sites.google.com/site/carolinecatmur/>

Clare Press

Department of Psychological Sciences, Birkbeck College, University of London, London WC1E 7HX, United Kingdom

c.press@bbk.ac.uk

<http://www.bbk.ac.uk/psychology/actionlab/>

<http://www.bbk.ac.uk/psychology/our-staff/academic/dr-clare-press>

Cecilia Heyes

All Souls College, University of Oxford, Oxford, OX1 4AL, and Department of Experimental Psychology, University of Oxford, Oxford OX1 3UD, United Kingdom

cecilia.heyes@all-souls.ox.ac.uk

<http://www.all-souls.ox.ac.uk/users/heyes/>

Abstract: This article argues that mirror neurons originate in sensorimotor associative learning and therefore a new approach is needed to investigate their functions. Mirror neurons were discovered about 20 years ago in the monkey brain, and there is now evidence that they are also present in the human brain. The intriguing feature of many mirror neurons is that they fire not only when the animal is performing an action, such as grasping an object using a power grip, but also when the animal passively observes a similar action performed by another agent. It is widely believed that mirror neurons are a genetic adaptation for action understanding; that they were designed by evolution to fulfill a specific socio-cognitive function. In contrast, we argue that mirror neurons are forged by domain-general processes of associative learning in the course of individual development, and, although they may have psychological functions, they do not necessarily have a specific evolutionary purpose or adaptive function. The evidence supporting this view shows that (1) mirror neurons do not consistently encode action “goals”; (2) the contingency- and context-sensitive nature of associative learning explains the full range of mirror neuron properties; (3) human infants receive enough sensorimotor experience to support associative learning of mirror neurons (“wealth of the stimulus”); and (4) mirror neurons can be changed in radical ways by sensorimotor training. The associative account implies that reliable information about the function of mirror neurons can be obtained only by research based on developmental history, system-level theory, and careful experimentation.

Keywords: action understanding; associative learning; contextual modulation; contingency; genetic adaptation; imitation; mirror neuron; poverty of the stimulus; sensorimotor experience.

1. Introduction

Mirror neurons (MNs) were discovered serendipitously in 1992 and given their brilliant name four years later (di Pellegrino et al. 1992; Gallese et al. 1996). The striking feature

of many MNs is that they fire not only when a monkey is performing an action, such as grasping an object using a power grip, but also when the monkey passively observes a similar action performed by another. Neurons with this capacity to match observed and executed actions, to code

both “my action” and “your action,” were originally found in area F5 of the ventral premotor cortex (PMC) (di Pellegrino et al. 1992; Gallese et al. 1996) and the inferior parietal lobule (IPL) (Bonini et al. 2010; Fogassi et al. 2005) of

the monkey brain. There is now a substantial body of evidence suggesting that MNs are also present in the human brain (Molenberghs et al. 2012).

MNs have received a great deal of attention from specialists and in the scientific and public media. Hailed as “cells that read minds” (Blakeslee 2006), “the neurons that shaped civilization” (Ramachandran 2009), and a “revolution” in understanding social behavior (Iacoboni 2008), MNs have been ascribed a wide variety of functions. The primary candidates relate to action understanding (Gallese & Sinigaglia 2011; Rizzolatti et al. 1996), imitation (Iacoboni et al. 1999), and language processing (Rizzolatti & Arbib 1998). However, signifying the way in which MNs have captured the attention and imagination of neuroscientists, psychologists, and philosophers, they have also been implicated in: embodied simulation (Aziz-Zadeh et al. 2006b), empathy (Avenanti et al. 2005), emotion recognition (Enticott et al. 2008), intention-reading (Iacoboni et al. 2005), language acquisition (Theoret & Pascual-Leone 2002), language evolution (Arbib 2005), manual communication (Rizzolatti et al. 1996), sign language processing (Corina & Knapp 2006), speech perception (Glenberg et al. 2008), speech production (Kuhn & Brass 2008), music processing (Gridley & Hoff 2006), sexual orientation (Ponseti et al. 2006), and aesthetic experience (Cinzia & Gallese 2009). In addition, it has been suggested that MN dysfunction contributes to a number of disorders, including autism (Dapretto et al. 2006; Nishitani et al. 2004; J. H. Williams et al. 2001), schizophrenia (Arbib & Mundhenk 2005), Down’s syndrome (Virji-Babul et al. 2008), multiple sclerosis (Rocca et al. 2008), cigarette addiction (Pineda & Oberman 2006), and obesity (Cohen 2008).

Thus, much of the first 20 years of MN research has been devoted to theorizing and speculation about their functions. In contrast, the primary focus of this article is the origin of MNs. Our principal questions are not “What do MNs do?” or “What are they for?”, but “What is the process that gives MNs their ‘mirroriness’; their fascinating, cardinal capacity to match observed with executed actions?”

The standard view of MNs, which we will call the “genetic account,” alloys a claim about the origin of MNs with a claim about their function. It suggests that the mirroriness of MNs is due primarily to heritable genetic factors, and that the genetic predisposition to develop MNs evolved because MNs facilitate action understanding. In the sense of “an adaptation” developed by G. C. Williams, and used in Evolutionary Psychology, the genetic account casts MNs as an adaptation for action understanding. In contrast, we argue in this article that the balance of evidence currently favors an “associative account” of MNs, which separates questions about their origin and function. It suggests that MNs acquire their capacity to match observed with executed actions through domain-general processes of sensorimotor associative learning, and that the role of MNs in action understanding, or any other social cognitive function, is an open empirical question. The associative account is functionally permissive; it allows, but does not assume, that MNs make a positive contribution to social cognition. Thus, there are three critical differences between the genetic and associative accounts: (1) The former combines, and the latter dissociates, questions about origin and function. (2) The genetic account suggests that natural selection has acted directly on MNs,

RICHARD COOK completed his Ph.D. in the Cognitive, Perceptual and Brain Sciences Department at University College London in 2012, and took up a lectureship at City University London in the same year. His research, which combines the methods of social cognitive neuroscience and visual psychophysics, examines the mechanisms of imitation and the perceptual processes recruited by faces, bodies, and actions. His work has been published in the *Journal of Experimental Psychology*, the *Journal of Vision*, *Proceedings of the Royal Society*, and *Psychological Science*.

GEOFF BIRD is a Senior Lecturer at the MRC (Medical Research Council) Social, Genetic, and Developmental Psychiatry Centre at the Institute of Psychiatry, King’s College London. After being taught how to think by Cecilia Heyes during his Ph.D. studies at University College London, Geoff now studies typical and atypical social cognition in collaboration with (among others) his excellent co-authors on this target article.

CAROLINE CATMUR is an ESRC (Economic and Social Research Council) Future Research Leader and Senior Lecturer in Cognitive Psychology at the University of Surrey. She received her B.A. in 2002 from the University of Oxford and her Ph.D. in 2009 from University College London. Her research investigates the cognitive and neural mechanisms underlying social cognition. Her current work, which combines behavioural studies with neuroimaging and transcranial magnetic stimulation, focuses on the development and control of social cognitive processes, including imitation, perspective-taking, and empathy.

CLARE PRESS is a Lecturer in Psychological Sciences at Birkbeck, University of London, UK. She received her B.Sc. and Ph.D. from University College London, and subsequently worked as an ESRC/MRC Postdoctoral Research Fellow at the Wellcome Trust Centre for Neuroimaging. Her research uses behavioural and neuroimaging techniques, with a focus on electrophysiological methods, to examine the mechanisms underlying perception-action mapping for social processes such as imitation and action perception, as well as action selection and control functions.

CECILIA HEYES is a Senior Research Fellow in Theoretical Life Sciences and full Professor of Psychology at All Souls College, University of Oxford, and a Fellow of the British Academy. Her work concerns the evolution of cognition. It explores the ways in which natural selection, learning, and developmental and cultural processes combine to produce the mature cognitive abilities found in adult humans. Her experimental and theoretical research, initially in animal cognition and later in human cognitive neuroscience, has focused on imitation, social learning, self-recognition, and theory of mind. She is currently asking to what extent these processes of cultural learning are, like literacy, culturally inherited.

whereas the associative account suggests that natural selection has played a background role; for example, acting on domain-general mechanisms of associative learning. (3) The genetic account assigns a relatively minor, facilitative role to sensory and/or motor experience in the development of MNs, whereas the associative account suggests that sensorimotor experience plays a major, instructive role in their development.

We begin, in section 2, with some basic information about the ways in which MNs have been defined and investigated in both monkeys and humans.¹ In the third section we present the genetic and associative accounts, and introduce four kinds of evidence that have the potential to favor one of these hypotheses over the other. Sections 4–7 discuss each of these types of evidence in turn. In section 8 we survey recent theories that are, or appear to be, alternatives to the genetic and associative accounts, and suggest that the associative account is stronger. Finally, in section 9 we argue that the associative account has major methodological implications for research investigating the functions of MNs. Unlike the genetic account, the associative account doesn't claim to tell us what MNs do or what they are for, but it does tell us how we can find out.

2. Mirror neuron basics

2.1. Locations and definitions

MNs have been found in the monkey brain (*Macaca nemistrina* and *Macaca mulatta*), not only in “classical” areas – ventral PMC and IPL – but also in “non-classical” areas, including primary motor cortex (Dushanova & Donoghue 2010; Tkach et al. 2007) and dorsal PMC (Tkach et al. 2007). There is also evidence of single neurons, or circumscribed populations of neurons, with sensorimotor matching properties in classical areas of the human brain, including posterior regions of the inferior frontal gyrus (IFG; considered the human homologue of the monkey F5) (Kilner et al. 2009) and inferior parietal cortex (Chong et al. 2008), and non-classical areas of the human brain, including dorsal PMC, superior parietal lobule, and cerebellum (Molenberghs et al. 2012), supplementary motor area, and medial temporal lobe (Mukamel et al. 2010).

Some researchers apply the term “mirror neuron” only to neurons found in classical areas (Brown & Brune 2012; Molenberghs et al. 2012), whereas others, like us, use the term to refer to neurons in both classical and non-classical areas (Gallese & Sinigaglia 2011; Keysers & Gazzola 2010). In addition to this variation in anatomical specificity, some researchers reserve the term “mirror neuron” for units that discharge during the observation and execution of precisely (Dinstein et al. 2008b; Keysers 2009) or broadly similar actions (Kilner et al. 2009), whereas others use the term, at least on occasions, to refer to any neuron that is responsive to both the observation and execution of action, regardless of whether the observed and executed actions are even broadly similar to one another (Gallese et al. 1996; Rizzolatti & Craighero 2004). In accord with the majority of researchers in the field, and the meaning of the word “mirror,” we take it to be a cardinal feature of MNs that they are responsive to observation and execution of similar actions. However, following common usage, we also refer to “logically related” MNs (see sect. 2.2), which

fire during observation and execution of dissimilar actions, as “mirror neurons.”

2.2. Monkeys

Early studies of the field properties of monkey MNs – the sensory and motoric conditions in which they fire – revealed three basic types: “Strictly congruent” MNs discharge during observation and execution of the same action, for example, precision grip. “Broadly congruent” MNs are typically active during the execution of one action (e.g., precision grip) and during the observation of one or more similar, but not identical, actions (e.g., power grip alone, or precision grip, power grip, and grasping with the mouth). “Logically related” MNs respond to different actions in observe and execute conditions. For example, they fire during the observation of an experimenter placing food in front of the monkey, and when the monkey grasps the food in order to eat it. MNs do not respond to the presentation of objects alone (di Pellegrino et al. 1992). However, “canonical neurons,” which are active during object observation and performance of an action that is commonly performed on that object, are co-located with MNs both in area F5 (Murata et al. 1997) and in the anterior intraparietal sulcus (Murata et al. 2000).

To date, monkey MNs have been found that are responsive to the observation and execution of hand and mouth actions. The hand actions include grasping, placing, manipulating with the fingers, and holding (di Pellegrino et al. 1992; Gallese et al. 1996). The mouth actions include ingestive behaviors such as breaking food items, chewing and sucking, and communicative gestures such as lip-smacking, lip-protrusion, and tongue-protrusion (Ferrari et al. 2003).

2.3. Humans

Only one study purports to offer direct evidence – from single cell recording – of MNs in the human brain (Mukamel et al. 2010). However, there is a considerable body of indirect evidence – from neuroimaging, transcranial magnetic stimulation (TMS), and behavioral studies – suggesting that human brains contain MNs or comparable “mirror mechanisms”; circumscribed cortical areas involved in both action production and observation (Glenberg 2011).

2.3.1. Neuroimaging. Functional magnetic resonance imaging (fMRI) has identified regions of PMC (both classic BA6 and BA44) and inferior parietal areas that are active during both action observation and execution (Aziz-Zadeh et al. 2006a; Buccino et al. 2004; Carr et al. 2003; Grèzes et al. 2003; Iacoboni et al. 1999; Leslie et al. 2004; Tanaka & Inui 2002; Vogt et al. 2007). Overlapping responses to action observation and execution have been found in single-subject analyses of unsmoothed data (Gazzola & Keysers 2009), confirming that the foregoing reports are not artifacts of group averaging. Most recently, repetition suppression protocols have been used to provide evidence of mirror populations encoding visual and motor representations of the same action. These paradigms exploit the logic that repeated stimulus presentation or action execution causes a decrease in neural responses (Grill-Spector et al. 2006). “Cross-modal” repetition

suppression effects have been reported, whereby action observation followed by execution of the same action, or vice versa, elicits repetition suppression in inferior parietal regions (Chong et al. 2008; Lingnau et al. 2009) and in PMC (Kilner et al. 2009; Lingnau et al. 2009).

2.3.2. Mirror pattern of MEPs. A human mirror mechanism is further suggested by “mirror” motor evoked potentials (MEPs) elicited during action observation (Fadiga et al. 1995). When TMS is applied to M1 during passive action observation, the amplitude of the MEPs recorded from the muscles required to execute that action is greater than the amplitude of the MEPs recorded when observing a different action. For example, observing index and little finger abduction movements selectively facilitates the amplitude of MEPs recorded from the first dorsal interosseus and abductor digiti minimi muscles, responsible for index and little finger movements, respectively (Catmur et al. 2011). That action observation selectively increases corticospinal excitability to action relevant muscles is suggestive of “mirror” sensorimotor connectivity.

2.3.3. Automatic imitation. Automatic imitation is said to occur when observation of an action involuntarily facilitates performance of a topographically similar action (body parts make the same movements relative, not to external frames of reference, but to one another) and/or interferes with performance of a topographically dissimilar action (Brass et al. 2001; Stürmer et al. 2000). Humans show robust automatic imitation when they observe hand, arm, foot, and mouth movements (Heyes 2011). This is regarded by many researchers as evidence of a human mirror mechanism (Blakemore & Frith 2005; Ferrari et al. 2009a; Iacoboni 2009; Kilner et al. 2003; Longo et al. 2008; van Schie et al. 2008). Supporting this view, several studies have shown that application of disruptive TMS to the IFG—a classical mirror area—interferes with automatic imitation (Catmur et al. 2009; Newman-Norlund et al. 2010).

3. The mirrorness of mirror neurons: Genetic or associative?

This section presents the standard, genetic account of the origin of MNs and the alternative associative account.

3.1. Genes for mirroring?

The genetic account assumes: (1) Among common ancestors of extant monkeys and humans, some individuals had a stronger genetic predisposition to develop MNs, and (2) these individuals were more reproductively successful than those with a weaker genetic predisposition because the development of MNs enhanced their capacity to understand the actions of other agents. Consequently, (3) a genetic predisposition to develop MNs became universal, or nearly universal, in monkeys and humans. (4) Motor experience (the performance of actions) and/or sensory experience (the observation of actions) plays a facilitative (Gottlieb 1976) or permissive (Gilbert 2003) role in the development of MNs, but their matching properties are primarily due to this genetic predisposition.

The term “action understanding” was introduced by Rizzolatti and colleagues to characterize the function of MNs (Rizzolatti & Fadiga 1998; Rizzolatti et al. 1996). As far as we are aware, it had not previously been used in research on animal or human cognition. The term plays a key role in the genetic account; it describes the adaptive function of MNs, the effects that made them a target of positive selection pressure. However, there is still no consensus about exactly what is meant by “action understanding,” or how it differs from cognate functions such as “action perception,” “action recognition,” and “action selection” (Gallese et al. 2011). Attempts to clarify have emphasized that, in comparison with purely visual processing of action, MN activity relates to the “meaning” of an action and yields a “richer understanding,” “real understanding,” or “understanding from within” (Gallese et al. 2011; Rizzolatti & Sinigaglia 2010). As we discuss further in section 8, these descriptions do not provide an operational definition of action understanding, that is, a definition that would allow behavior based on (this kind of) action understanding to be distinguished empirically from behavior based on other processes.

Until recently, the genetic account was largely implicit in discussions of the “evolution” of MNs (Gallese & Goldman 1998; Rizzolatti & Arbib 1998; Rizzolatti & Craighero 2004; M. J. Rochat et al. 2010). For example, it has been suggested that “the mirror neuron mechanism is a mechanism of great evolutionary importance through which primates understand actions done by their conspecifics” (Rizzolatti & Craighero 2004, p. 172) and that “in their basic properties, MNs constitute a relatively simple action-perception mechanism that could have been exploited several times in the course of animal evolution” (Bonini & Ferrari 2011, p. 172). A number of discussions have also expressed the view that MNs are present at birth (Ferrari et al. 2009; Gallese et al. 2009; Lepage & Theoret 2007; Rizzolatti & Fadiga 1998), a feature commonly associated with traits for which there is strong genetic predisposition (Mameli & Bateson 2006). For example, Casile and colleagues have suggested that “both face processing and the mirror neuron system, or at least the part involved in facial movements, rely on a brain network that is present already at birth and whose elements are probably genetically predetermined” (Casile et al. 2011, p. 531).

In its starkest form, the genetic hypothesis would suggest that gene-based natural selection has provided each individual—monkey and human—with MNs that code the mapping between a fixed set of observed and executed actions, and that experience plays a minimal role in the development of the observation-execution matching properties of these neurons. However, the genetic hypothesis does not necessarily assume that experience plays a minimal role. For example, in a recent explicit statement of the genetic account, Gallese et al. (2009) suggested that links form during gestation between motor regions and “to-become-visual” regions that will subsequently mediate sensorimotor matching abilities in young infants. They implied that these projections are genetically predisposed to target certain visual areas, and therefore that the matching properties of MNs are produced by information encoded in the genome. However, they also suggested that motor experience plays a part in preparing motor regions to send projections to visual areas, and that visual

experience may also facilitate the maturation of fully functioning MNs.

3.2. A product of associative learning

The associative hypothesis assumes that gene-based natural selection has played a significant background role with respect to the development of MNs; for example, in shaping the anatomy of visual and motor cortex for visual guidance of action, and in producing the capacity for associative learning in neural tissue. However, it suggests that the cardinal matching properties of MNs are a product, not of a specific genetic predisposition, but of domain-general processes of associative learning—the same kind of learning that produces Pavlovian and instrumental conditioning phenomena (Catmur et al. 2009; Heyes 2010; Ray & Heyes 2011). Associative learning is found in a wide range of vertebrate and invertebrate species, indicating that it is an evolutionarily ancient and highly conserved adaptation for tracking predictive relationships between events (Heyes 2012b; Schultz & Dickinson 2000).

Figure 1 is a schematic representation of how MNs could acquire their matching properties through sensorimotor associative learning. Before associative learning, sensory neurons in the superior temporal sulcus (STS), responsive to different high-level visual properties of observed action (Oram & Perrett 1994; 1996), are weakly connected, directly or indirectly, to motor neurons in PMC (Rizzolatti et al. 1988) and parietal cortex (Gallese et al. 2002). Some of these connections may be stronger than others, but the links between sensory and motor neurons coding similar

actions are not consistently stronger than other, non-matching links. The kind of learning that produces MNs occurs when there is correlated (i.e., contiguous and contingent) excitation of sensory neurons and motor neurons that code similar actions. For example, when an adult imitates an infant's facial movements, there might be correlated excitation of neurons that are responsive to the observation and execution of lip protrusion. Correlated excitation of the sensory and motor neurons increases the strength of the connection between them, so that subsequent excitation of the sensory neuron propagates to the motor neuron. Thereafter, the motor neuron fires, not only during execution of lip protrusion, but also, via its connection with the sensory neuron, during observation of lip protrusion; what was originally a motor neuron has become a lip protrusion MN. Correlated excitation of sensory and motor neurons encoding the same property of action occurs not only when humans are imitated, but also when we observe our own actions (directly or using an optical mirror); observe others during the kind of synchronous activities involved in sports and dance training; and as a consequence of “acquired equivalence” experience, for example, when the same sound (a word, or a sound produced by an action) is paired sometimes with observation of an action and sometimes with its execution (Catmur et al. 2009; Ray & Heyes 2011).

There are several important things to note about the associative hypothesis:

1. *Strong experience-dependence*—It suggests that correlated sensorimotor experience plays an inductive

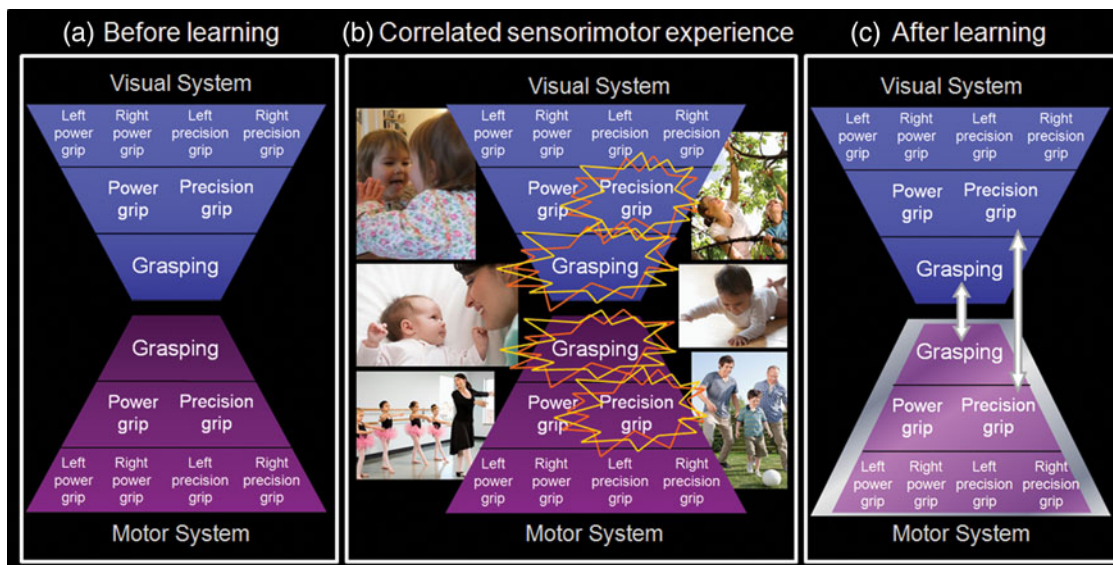


Figure 1. Mirror neurons from associative learning. (a) Before learning, sensory neurons in STS, encoding visual descriptions of observed action, are not systematically connected to motor neurons in premotor and parietal areas involved in the production of similar actions. (b) Through social interaction and self-observation in the course of typical development, agents receive correlated sensorimotor experience; they see and do the same action at about the same time (contiguity), with one event predicting the other (contingency). This experience produces correlated activation of sensory and motor neurons coding similar actions, and, through associative learning, (c) strengthens connections between these neurons. Due to these connections, neurons that were once involved only in the execution of action will also discharge during observation of a similar action; motor neurons become MNs (see sect. 3.2). Because the visual system and motor system are organised hierarchically, some types of sensorimotor experience produce correlated activation of sensory and motor neurons coding relatively low-level features of action (e.g., left or right hand, power or precision grip), and thereby generate strictly congruent, hand- and direction-sensitive MNs. Other types produce correlated activation of neurons coding relatively high-level features (e.g., grasping) and generate broadly congruent MNs (see sect. 5.1).

(Gottlieb 1976) or instructive (Gilbert 2003) role; without this kind of experience, MNs would not develop at all.

2. *Social construction* – It proposes that much of the sensorimotor experience required for MN development comes from being imitated, synchronous action, and exposure to action words (Ray & Heyes 2011), and therefore that MNs are to a very large extent built through social interaction.

3. *Contingency* – Following contemporary learning theory, the associative account specifies that MN development requires, not just that sensory and motor neurons “fire together” in a Hebbian way (contiguity), but that the event provoking firing of one predicts the event provoking firing of the other (contingency; Cook et al. 2010).

4. *Testability* – The associative account makes novel predictions about the development and mature properties of MNs, many of which have already been tested and supported by experiments using a variety of methods (see sect. 7).

Thus, the associative hypothesis implies that the characteristic, matching properties of MNs result from a genetically evolved process, associative learning, but this process was not “designed” by genetic evolution to produce MNs. Rather, it just happens to produce MNs when the developing system receives correlated experience of observing and executing similar actions. When the system receives correlated experience of observing objects and executing actions, the same associative process produces canonical neurons. When the system receives correlated experience of observing one action and executing a different action, the same associative process produces logically related MNs.

3.3. Not nature versus nurture

The contrast between the genetic and associative hypotheses does not represent a dichotomous nature–nurture debate. It has been recognized for decades that the development of all phenotypic characteristics depends on the interaction of nature *and* nurture, genes *and* the environment, evolution *and* learning (Elman et al. 1996; Oyama 1985). Rather, the two accounts differ in the specific roles they assign to genetic evolution and to learning, and in the types of experience they take to be important, in producing the characteristic matching properties of MNs. The genetic hypothesis says that genetic evolution has played a specific and decisive role, and learning – based on sensory and/or motor experience – plays a merely facilitative role, in the development of MNs. In contrast, the associative hypothesis says that genetic evolution has played a non-specific background role, and that the characteristic matching properties of MNs are forged by sensorimotor learning.

Regarding the function of MNs, the genetic account assumes that they play a fundamental role in action understanding, and that this is why a specific genetic predisposition to develop MNs was favored by natural selection. In other words, it proposes that action understanding is the “adaptive function” of MNs, or that MNs are “an adaptation” for action understanding. In this way, the genetic account offers a hypothesis about the function of MNs as an explanation for their origins. In contrast, the associative account separates questions about the origin and function of MNs. It suggests that MNs develop through associative learning, and that further research is needed to find out

how they contribute to social cognition (see sect. 9). If this research reveals that MNs make positive contributions to social cognition, these would be “psychological uses” or “psychological functions,” but not necessarily “adaptive functions”; they may not have enhanced reproductive fitness, nor resulted in the evolution of mechanisms specifically designed to foster the development of MNs (see sect. 8). Rather, it is possible that MNs are constructed by domain-general processes of associative learning, and are recruited in the course of development to contribute to one or more psychological functions, without either the construction or the recruitment processes having become a specific target of gene-based selection (Elman et al. 1996).

In this respect, MNs may be like beak morphology in Neotropical woodcreepers, which has been selected for foraging and food manipulation (a non-social function, analogous to visuomotor capability) but also has effects on song production (a social function, analogous to action understanding; Derryberry et al. 2012). Another more closely related example comes from honeybees, which are able to use associative learning to discriminate among human faces (Dyer et al. 2005). Given the taxonomic- and domain-generality of associative learning, and the fact that human faces were not part of the environment in which honeybee nervous systems evolved, we can be sure that associative learning is not an adaptation for face discrimination in honeybees. However, when they are put in an environment where faces are important, honeybees can use associative learning about faces to optimize their foraging behavior. Another example, which may be closely related in a different way, is the area of the human occipito-temporal cortex known as the “visual word form area” (VWFA; Petersen et al. 1990). This area plays an important role in reading, but, given the recent emergence of literacy in human history, the VWFA is very unlikely to be a genetic adaptation for reading. Rather, the reading-related properties of the VWFA are forged in the course of development, by literacy training, from a system adapted for generic object recognition.

3.4. Four kinds of evidence

Four evidence-based arguments are crucial in deciding between the genetic and associative accounts. The first provided the inspiration and foundation for the genetic hypothesis. It suggests that the field properties of MNs indicate that they were designed for action understanding. The terms “design” and “purpose” are used here as they were by G. C. Williams in his seminal work on *Adaptation and Natural Selection* (Williams 1966). Williams described adaptations as designed by natural selection to fulfill a particular purpose, and emphasized that the mark of an adaptation is that it has features making it peculiarly apt to achieve a specific end in a highly efficient way. For example, “An examination of the legs and feet of the fox forces the conclusion that they are designed for running and walking, not for the packing or removal of snow” (p. 13). In a similar way, supporters of the genetic hypothesis argue that examination of the field properties of MNs – and, in particular, their “goal” coding – forces the conclusion that MNs are designed for action understanding. In section 3 we examine the field properties of MNs and suggest that this argument is not compelling.

The second argument is complementary to the first; it has a similar form but advocates the associative hypothesis. It suggests that research using conditioning procedures shows associative learning to be the right kind of learning to produce MNs. Specifically, as we discuss in section 5, the ways in which associative learning tracks contingent relationships, and enables contextual modulation of these connections, makes it apt to produce MNs (and non-matching visuomotor neurons) in typical developmental environments.

The second pair of arguments draws on research examining the development of MNs and their modification through sensorimotor experience. Section 6 discusses research with infants and adults that has been used to support a “poverty of the stimulus” argument (Chomsky 1975); to suggest that MNs emerge too early in development or, more generally, after too little sensorimotor experience, to have been forged by associative learning. In contrast, we offer a “wealth of the stimulus” argument.

Finally, section 7 focuses on evidence that, even in adulthood, the properties of MNs can be changed in radical ways by relatively brief periods of sensorimotor experience. We argue, against various objections, that this evidence is sound and therefore supports the associative hypothesis by showing that it has produced novel, testable predictions which have been confirmed by experiment.

4. Designed for action understanding

Supporters of the genetic hypothesis argue that examination of the field properties of MNs shows that they encode “goals,” and this characteristic indicates that they were designed by genetic evolution to mediate action understanding (Bonini & Ferrari 2011; Rizzolatti & Craighero 2004; Rizzolatti & Sinigaglia 2010). We therefore begin our survey of the evidence by considering how well the neurophysiological data accord with this view. The term “goal” affords numerous interpretations (Hickok 2009). We will consider two definitions commonly adopted, assuming that MNs encode “goals” if they encode object-directed actions (sect. 4.1) or high-level action intentions (sect. 4.2).

4.1. Goals as object-directed actions

Early descriptions of MN field properties reported that pantomimed actions (e.g., miming a precision grip in the absence of an object) and intransitive actions (e.g., tongue-protrusion) did not elicit MN responses (di Pellegrino et al. 1992; Gallese et al. 1996). In contrast, robust responses were reported when monkeys observed object-directed actions. This pattern raised the possibility that MNs encode “goals” in the sense that they are responsive only to object-directed actions (di Pellegrino et al. 1992; Gallese et al. 1996).

However, a close reading of the single-cell data suggests that only a small subset of MNs appeared to have been “designed for” encoding action goals in these terms. A subset of the MNs described in the early reports continued to respond, albeit less strongly, to pantomimed or intransitive actions (di Pellegrino et al. 1992; Gallese et al. 1996, Figure 5b), and subsequent studies confirmed that sizable proportions, perhaps the majority, of MNs exhibit robust

responses to the observation of object-free body movement. Kraskov et al. (2009) reported that 73% of MN responses modulated by observation of object-directed grasping showed similar modulation during observation of pantomimed grasping. Also, substantial proportions of MNs respond to intransitive mouth movements such as lip-smacking, lip-protrusion, and tongue-protrusion (Ferrari et al. 2003).

Single-unit data also show that, even when they are responding to object-directed actions, MNs have field properties suggesting that they were not tuned to do this by genetic evolution. For example, after training in which tools were used to pass food items to monkeys, MNs were discovered that respond to the observation of actions such as grasping with pliers (Ferrari et al. 2005). Similarly, “audiovisual” MNs respond to unnatural sounds associated with actions; for example, the sound of metal striking metal, plastic crumpling, and paper tearing (Keyes et al. 2003; Kohler et al. 2002). Importantly, large numbers of tool-use and audiovisual MNs respond more to the sight of tool-actions and to action sounds than to the sight of gripping or tearing executed with the hands. The fact that these MNs respond maximally to unnatural stimuli—that is, stimuli to which the evolutionary ancestors of contemporary monkeys could not possibly have been exposed—is hard to reconcile with the genetic hypothesis (Cook 2012; see sect. 7).

4.2. Goals as high-level intentions

The term “goal” has also been used to refer to what, at a high level of generality, the actor intends to achieve through their behavior—for example, “grasp in order to eat” (Fogassi et al. 2005) or “taking possession of an object” (M. J. Rochat et al. 2010). Rizzolatti and Sinigaglia (2010, p. 269) state: “only those [neurons] that can encode the goal of the motor behavior of another individual with the greatest degree of generality can be considered to be crucial for action understanding.” The suggestion that MNs encode high-level action intentions is made plausible by reports that MN responses to grasping can be modulated by the final outcome of the motor sequence (Bonini et al. 2010; Fogassi et al. 2005). It is also consistent with reports that some broadly congruent MNs respond to the observation of multiple actions; for example, any “grasping” action executed with the hand or mouth (Gallese et al. 1996).

However, the single-cell data again suggest that relatively few MNs have the field properties one would expect of a system designed by genetic evolution to represent high-level action intentions. For example, Gallese et al. (1996) reported that during action observation 37.5% of MNs responded differently depending on whether the action was executed with the left or right hand, and 64% showed direction sensitivity, preferring either left-to-right or right-to-left grasping actions. Similarly, many MNs (53%) respond selectively to the observation of actions executed within (“peripersonal” MNs) or beyond (“extrapersonal” MNs), not the actor’s, but the observing monkey’s reach (Caggiano et al. 2009). The majority (74%) of MNs also exhibit view-dependent responses; some MNs are tuned to egocentric (first-person) presentation, while others respond maximally to allocentric (third-person) perspectives (Caggiano et al. 2011). Each of these classes of

MN is sensitive to features of action that fall well below the “greatest degree of generality” of intentions such as “grasping in order to eat” or “taking possession of an object.”

The field properties of logically related MNs are perhaps the hardest to reconcile with the idea that MNs were designed by genetic evolution to mediate action understanding by activating in the observer the same “goal” that is guiding the actor’s behavior. Logically related MNs fire when a monkey observes an action with one goal (e.g., placing food items on a table, with the intention of giving food to the monkey) and when the monkey executes an action with a different goal (e.g., grasping the food with a precision-grip and bringing it to the mouth, with the intention of eating). Thus, these MNs respond to different object-directed actions, with different intentions, in observe and execute conditions.

5. The right kind of learning

The previous section argued that many MNs have field properties incompatible with the hypothesis that they were designed by genetic evolution to mediate action understanding via goal coding. In complementary fashion, this section argues that research on the roles of contingency and contextual modulation in associative learning enables the associative hypothesis to provide a unified account of all the MN field properties reported to date.

5.1. Predictive relationships

It has long been recognized that associative learning depends, not only on contiguity—events occurring close together in time—but also on contingency—the degree to which one event reliably predicts the other. Where the predictive relationship between two events is weak—that is, where one event is equally likely to occur in the presence and absence of the other event—contiguous pairings produce little or no learning (Elsner & Hommel 2004; Rescorla 1968; Schultz & Dickinson 2000). The associative account therefore predicts that MNs will acquire sensorimotor matching properties only when an individual experiences contingencies between sensory events and performed actions (Cooper et al. 2013b). This feature of associative learning ensures that the matching properties of MNs reflect, not just chance co-occurrences, but sensorimotor relationships that occur reliably in the individual’s environment. Evidence that the human mirror mechanism is modified by contingent but not by non-contingent sensorimotor experience has been reported by Cook et al. (2010).

Contingency sensitivity explains the mix of strictly congruent MNs, sensitive to the low-level features of observed actions (type of grip, effector used, direction of movement, viewpoint, proximity to the observer), and broadly congruent MNs, responsive to multiple related actions irrespective of the manner of their execution. Both visual and motor systems are known to be organized hierarchically (Felleman & Van Essen 1991; Giese & Poggio 2003; Jeanerod 1994; Perrett et al. 1989), comprising different populations encoding relatively low-level representations (e.g., descriptions of particular “precision” or “power” grips) and more abstract representations (e.g., descriptions of

“grasping”). Therefore, contingencies can be experienced between both low- and high-level sensory and motor representations. When a monkey observes itself performing a precision grip, the excitation of sensory and motor populations encoding a specific grip are correlated. However, during group feeding, a monkey might observe and perform a range of grasping actions, thereby causing correlated excitation of higher-level visual and motoric descriptions of grasping. Contingency sensitivity therefore explains the existence of both strictly congruent MNs, tuned to a particular sensory representation (e.g., a right-to-left precision grip executed with the right hand viewed egocentrically), and broadly congruent MNs, responsive to the observation of a number of related actions.

Contingency sensitivity also explains the existence of logically related, audiovisual, and tool-use MNs. According to the associative hypothesis, MNs acquire sensorimotor properties whenever individuals experience a contingency between “seeing” and “doing.” There is no requirement that contingencies are between action performance and the observation of the same action, or indeed of natural action-related stimuli, such as the sight of animate motion or sounds that could have been heard by ancestors of contemporary monkeys. Both monkeys and humans frequently experience non-matching sensorimotor contingencies, where the observation of one action predicts the execution of another; for example, you release and I grasp (Newman-Norlund et al. 2007; Tiedens & Fragale 2003). The associative account therefore explains in a very straightforward way why logically related MNs respond to different actions in observe and execute conditions. Equally, the associative account explains in a simple way why “tool-use” MNs (Ferrari et al. 2005) develop when action performance is reliably predicted by the sight of actions performed with tools (e.g., food items being gripped with pliers), and why “audiovisual” MNs (Keysers et al. 2003; Kohler et al. 2002) develop when action performance predicts characteristic action sounds (e.g., paper tearing or plastic crumpling; Cook 2012).

5.2. Contextual modulation

Studies of conditioning indicate that learned responses are often subject to contextual control; if a stimulus is associated with two responses, each in a different context, then the context determines which response is cued by the stimulus (Bouton 1993; 1994; Peck & Bouton 1990). For example, Peck and Bouton (1990) initially placed rats in a conditioning chamber with a distinctive scent (e.g., coconut) where they learned to expect electric shock following a tone. The rats were then transferred to a second chamber with a different scent (e.g., aniseed) where the same tone predicted the delivery of food. The rats quickly learned the new contingency and conditioned foraging responses replaced conditioned freezing. However, learning in the second phase was context dependent. When the rats were returned to the first chamber, or transferred to a third chamber with a novel scent, the tone once again elicited freezing. By drawing on the components of associative learning theory that explain this kind of effect, the associative account of MNs can explain contextual modulation of MN firing (Cook et al. 2012a).

Many, possibly all, of the findings cited as evidence that monkey MNs code action goals can also be interpreted

within an associative framework as evidence that MNs are subject to contextual control. For example, some MNs show stronger visual responses to object-directed grasping than to pantomimed grasping in object-absent contexts (Gallese et al. 1996), and in some cases, the modulating influence of the object-context can be seen even when the target object is occluded prior to contact with the hand (Umiltà et al. 2001). Similarly, MN responses during the observation of grasping may be modulated by the type of object being grasped (Caggiano et al. 2012), with some MNs responding more strongly in the presence of high-value (food, or non-food objects predictive of reward), and some in the presence of low-value (non-food objects not associated with reward) stimuli. In the clearest example, the same motor act, grasping with a precision grip, elicits different MN responses dependent on whether the action is observed in the presence (“grasp to place”) or absence (“grasp to eat”) of a plastic cup (Bonini et al. 2010; Fogassi et al. 2005). Rather than the plastic cup providing a cue to the actor’s intention, it may act as a contextual cue modulating the operation of two associations. In the same way that the sound of the tone elicited different behaviors when presented in the coconut and aniseed contexts (Peck & Bouton 1990), observing a precision grip may excite different MNs in the cup-present and cup-absent contexts.

Thus, many of the field properties cited as evidence of goal (intention) coding by MNs can also be explained by contextual modulation within an associative framework. Under the “goal” interpretation, these field properties constitute direct evidence that MNs mediate action understanding. Under the associative interpretation, they are very interesting but not decisive. The flexibility apparent in the field properties of MNs gives them the potential to make a useful contribution to social behavior. However, further research, examining the behavior of whole organisms, not only of neurons, is needed to find out how this potential is realized (see sect. 9).

6. Wealth and poverty of the stimulus

Research involving infants (sect. 6.1) and adults (sect. 6.2) has been used to support a poverty argument suggesting that MNs emerge too early in development or, more generally, after too little sensorimotor experience, to have been forged by associative learning.

6.1. Mirroring in infancy

It has been argued that: (1) imitation is mediated by MNs (or a mirror mechanism); (2) both human and monkey infants are able to imitate observed actions when they have had minimal opportunity for visuomotor learning; and (3) therefore, the associative account of the origin of MNs must be wrong (Gallese et al. 2011). The structure of this argument is valid, but the evidence supporting the second assumption (e.g., Heimann et al. 1989; Meltzoff & Moore 1977; Nagy et al. 2005) has been challenged in two respects. Building on previous analyses (e.g., Anisfeld 1996), a recent review found evidence that human neonates copy only one action – tongue-protrusion – and that this copying does not show the specificity characteristic of imitation or of MNs (Ray & Heyes 2011). Figure 2

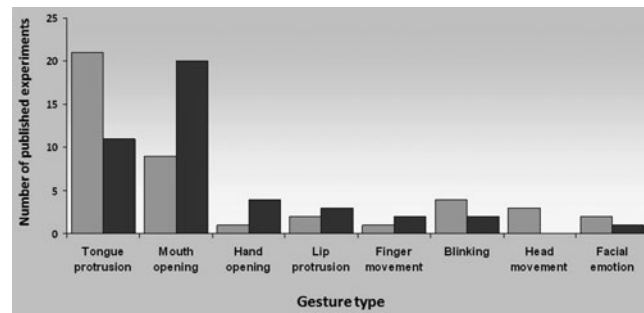


Figure 2. Summary of experiments seeking evidence of gesture imitation in human infants (adapted from Ray & Heyes 2011). “Gesture type” refers to the target or modelled movement. Positive frequencies (lighter bars) indicate the number of published experiments reporting positive cross-target comparisons (i.e., infants performed the target action more often after observing the target action than after observing an alternative action). Negative frequencies (darker bars) indicate the number of experiments reporting failure to find a significant difference in cross-target comparison.

illustrates the first of these points. For each of the action types tested in young infants, it shows the number of published studies reporting positive evidence of imitation and the number reporting negative evidence. This is a highly conservative measure of how often young infants have failed imitation tests, because it is much harder to publish negative than positive results (Fanelli 2012). Nonetheless, Figure 2 shows that the number of positive reports substantially exceeds the number of negative reports only for tongue-protrusion. Evidence that even the tongue-protrusion effect lacks the specificity characteristic of imitation and MNs – that it is an exploratory response, rather than an effect in which action observation is met with performance only of a similar action – comes from research showing that tongue-protrusion can be elicited by a range of arousing stimuli, including flashing lights and lively music (Jones 1996; 2006), and that it is greater when infants observe a mechanical “tongue” or disembodied mouth (Soussignan et al. 2011).² More broadly, evidence that the development of imitation is crucially dependent on learning is provided by a study of 2-year-old twins showing that individual differences in imitation were a result predominantly of environmental rather than genetic factors (McEwen et al. 2007), and by a recent study of infants indicating that individual differences in associative learning ability at 1-month predicted imitative performance eight months later (Reeb-Sutherland et al. 2012).

Turning from human to monkey infants, Ferrari et al. (2006) reported immediate imitation of tongue-protrusion and lip-smacking in 3-days-old monkeys. However, the effects were not present on days 1, 7, and 14 postpartum, and it is not clear whether they were replicated in a subsequent study using a similar procedure (Paukner et al. 2011). The later study did report imitation of lip-smacking in monkeys less than one week old, but this effect seems to have been due to a low frequency of lip-smacking in the control condition, when infants were observing a static neutral face, rather than to an elevated frequency of lip-smacking when the infants were observing lip-smacking. Therefore, in common with the data from human infants, studies of imitation in newborn monkeys do not currently support a poverty argument.

A similar poverty argument suggests that the associative account must be wrong because (1) suppression of alpha band (~6–13 Hz) oscillations over central scalp locations during action observation (and execution) reflects the operation of MNs or a mirror mechanism, and (2) electroencephalographic (EEG) studies indicate that both human and monkey infants show alpha suppression when they have had minimal opportunity for sensorimotor learning (Gallese et al. 2011; Nyström et al. 2011). In this case, both of the assumptions are questionable. First, the functional significance of alpha band oscillatory activity is poorly understood even in human adults, and is yet more difficult to interpret in human and monkey infants where, for example, less information is available about the source (Marshall & Meltzoff 2011). Second, human adult studies have traced the likely source of central alpha suppression during action execution to the somatosensory cortex (Hari & Salmelin 1997), suggesting that it may not index motor processing at all.³ Third, even if alpha suppression does index motor processing, it does not show that the motor activation matches or mirrors the actions observed (Marshall & Meltzoff 2011; Pfurtscheller et al. 2000). For example, alpha suppression during observation of lip-smacking, which has been reported in neonatal monkeys (Ferrari et al. 2012), may reflect a generalized readiness to act, or arousal-related motor activation of tongue-protrusion, rather than motor activation of lip-smacking, and thereby the occurrence of MN or mirror mechanism activity. Furthermore, studies of human infants, which provide superior source information, have not shown that central alpha suppression occurs when infants have had insufficient correlated sensorimotor experience to build a mirror mechanism through associative learning. Indeed, studies of human infants suggest an age-related trend consistent with the associative hypothesis (see Marshall et al. [2011] for a review).

Sound evidence of MN activity in newborns—which, we suggest, has not been provided by research to date on imitation and alpha suppression—would be inconsistent with the associative model. However, it is important to note that the associative account is predicated on a “wealth of the stimulus” argument and therefore anticipates MN activity in young infants (Ray & Heyes 2011). This wealth argument points out that human developmental environments typically contain multiple sources of the kind of correlated sensorimotor experience necessary to build MNs; each of these sources is rich; and the mechanisms of associative learning can make swift and efficient use of these sources. The range of sources available to young human infants includes self-observation, being imitated by adults, being rewarded by adults for imitation, and the kind of experience in which, for example, lip movements make the same smacking or popping sound when observed and executed. Evidence of the richness of these sources comes from studies showing that infants spend a high proportion of their waking hours observing their own hands in motion (P. Rochat 1998; White et al. 1964); in face-to-face interaction with a caregiver, they are imitated on average once every minute (Jones 2009; Pawlby 1977; Uzgiris et al. 1989); and “noisy actions,” which provide an early source of acquired equivalence experience, are among the first that infants imitate (Jones 2009). A common misconception about associative learning is that it always occurs slowly. Directly relevant evidence that this is not

the case comes from studies showing that, when the contingency is high, infants can learn action-effect associations in just a few trials (Paulus et al. 2012; Verschoor et al. 2010).

6.2. Motor training in adulthood

It has been claimed that the associative account “cannot explain why motor experience obtained without visual feedback can affect perception of human biological motion related to that experience” (Gallese et al. 2011, p. 383). This claim assumes that the perception of human biological motion is mediated by MNs or a mirror mechanism, and appeals to a subtle poverty argument; it suggests that the fundamental properties of MNs—the way in which they map observed with executed actions—can be changed by motor experience alone, that is, in the absence of correlated sensorimotor experience.

Two types of evidence, from studies that were not designed to investigate MNs, have been cited in support of this subtle poverty argument (Gallese et al. 2011). First, when observing point-light displays of whole body movements such as walking, from a third party perspective, people are better able to recognize themselves than to recognize their friends (Beardsworth & Buckner 1981; Loula et al. 2005). Second, practice in executing actions can improve visual discrimination of those actions, even when actors are prevented from observing their movements during execution (Casile & Giese 2006; Hecht et al. 2001). These motor training effects, and the self-recognition advantage, are interesting and important phenomena in their own right. If they were mediated by a mirror mechanism—that is, a mechanism in which there is a direct, unmediated connection between visual and motor representations of action—they would also support a poverty argument. However, a recent study provides evidence that these effects depend on an indirect mechanism representing temporal cues. It shows, using avatar facial motion stimuli, that the self-recognition advantage is maintained despite gross distortion of the kind of spatial cues that characterize biological motion, but is abolished by even relatively minor disturbance of domain-general temporal cues (Cook et al. 2012b). In the absence of appropriate visual experience, actors appear to be able to use their considerable knowledge of the rhythmic characteristics of their own actions to recognize and better represent allocentric movement displays. Thus, motor training effects and the self-recognition advantage are of independent interest, but they do not support a poverty argument because current evidence suggests that they do not depend on a mirror mechanism.

7. Sensorimotor learning changes mirror neurons

7.1. Testing the predictions of the associative account

The associative account has been explicitly tested in experiments examining the effects of laboratory-based sensorimotor training on mirror mechanisms in human adults. Building on the results of more naturalistic studies (Calvo-Merino et al. 2005; 2006; Cross et al. 2006; Ferrari et al. 2005; Haslinger et al. 2005; Jackson et al. 2006; Keysers et al. 2003; Kohler et al. 2002; Margulis et al. 2009; Vogt et al. 2007), these experiments have isolated the effects of sensorimotor experience from those of purely visual and

purely motor experience. Using all of the measures of mirror mechanism activity commonly applied to humans (sect. 2.3), they have shown that relatively brief periods of sensorimotor experience can enhance (Press et al. 2007; Wiggett et al. 2012), abolish (Cook et al. 2010; 2012a; Gillmeister et al. 2008; Heyes et al. 2005; Wiggett et al. 2011), reverse (Catmur et al. 2007; 2008; 2011), and induce (Landmann et al. 2011; Petroni et al. 2010; Press et al. 2012a) mirror mechanism activity. Each of these findings confirms a novel prediction of the associative account: it reveals flexibility of exactly the kind one would expect if MNs/mechanisms are forged by sensorimotor associative learning. In contrast, this kind of flexibility does not provide any support for the genetic hypothesis. Indeed, if MNs were a genetic adaptation, some evolutionary frameworks would predict that the development of MNs would be protected or “buffered” against environmental perturbations that could interfere with their adaptive function (Cosmides & Tooby 1994; Pinker 1997). In the case of a genetic adaptation for action understanding, this would include perturbations with the potential to divert MNs from coding properties of action, rather than of inanimate stimuli, and from coding similar, rather than dissimilar, observed and executed actions.

Evidence that MNs/mechanisms are not resistant to coding inanimate stimuli comes from studies showing that arbitrary sound, color, and shape stimuli can induce mirror MEP (D’Ausilio et al. 2006; Petroni et al. 2010), fMRI (Cross et al. 2009; Landmann et al. 2011; Press et al. 2012a) and behavioral effects (Press et al. 2007) following sensorimotor training (Press 2011). For example, Press and colleagues gave participants approximately 50 minutes of sensorimotor training in which they repeatedly opened their hand when seeing a robotic pincer open, and closed their hand when seeing the robotic pincer close (Press et al. 2007). Prior to this training, the pincer movement elicited less automatic imitation (see sect. 2.3) than human hand movement, but 24 hours after training, the automatic imitation effect was as strong for the pincer movement as for the human hand.

Evidence that MNs/mechanisms are not resistant to coding dissimilar actions comes from studies showing that non-matching (or “counter-mirror”) sensorimotor training abolishes automatic imitation (Cook et al. 2010; 2012a; Gillmeister et al. 2008; Heyes et al. 2005; Wiggett et al. 2011), and reverses both fMRI (Catmur et al. 2008) and MEP mirror responses (Catmur et al. 2007). For example, Catmur and colleagues gave participants approximately 90 minutes of non-matching sensorimotor training in which they repeatedly made an index finger movement while observing a little finger movement, and vice versa (Catmur et al. 2007). Before this training the participants showed mirror MEP responses, for example, observation of index finger movement elicited more activity in an index finger muscle than observation of little finger movement, and vice versa for the little finger muscle. After training, this pattern was reversed, for example, observation of index finger movement elicited more activity in the little finger muscle than observation of little finger movement.

7.2. Objections to sensorimotor training evidence

Objections to this evidence suggest, in various ways, that it does not show that sensorimotor experience can change

MNs/mechanisms. For example, it has been suggested that the evidence comes only from studies of object-free actions and yet MNs code only object-directed actions (Rizzolatti & Sinigaglia 2010). However, a study of pianists has shown that experience modulates mirror responses to object-directed actions (Haslinger et al. 2005) and, as discussed in sect. 4.1, monkey studies of communicative gestures (Ferrari et al. 2003) and pantomimed reaching movements (Kraskov et al. 2009) have identified MNs that code object-free actions.

A related concern is that, because they use indirect measures (fMRI, MEPs, and automatic imitation), rather than single-cell recording, sensorimotor learning experiments may not be measuring MN responses. Section 2.3 reviewed the evidence for human MNs from a range of experimental techniques. These include conjunction of neural responses during action observation and performance (Gazzola & Keysers 2009; Iacoboni et al. 1999; Vogt et al. 2007), suppression of neural responses to cross-modally (perceptual-motor or motor-perceptual) repeated actions (Kilner et al. 2009; Press et al. 2012b), muscle-specific MEPs (Catmur et al. 2011; Fadiga et al. 1995), and automatic imitation (Brass et al. 2001; Stürmer et al. 2000). In isolation, each of these measures is imperfect (Caggiano et al. 2013), but together they provide strong converging evidence for human MNs. Sensorimotor learning effects have been demonstrated for all these measures of mirror responses: fMRI conjunction (Catmur et al. 2008; Landmann et al. 2011); repetition suppression (Press et al. 2012a); MEPs (Catmur et al. 2007; 2011; D’Ausilio et al. 2006; Petroni et al. 2010); and automatic imitation (Cook et al. 2010; 2012a; Gillmeister et al. 2008; Heyes et al. 2005; Press et al. 2007; Wiggett et al. 2011). Thus, converging evidence using multiple techniques strongly suggests that sensorimotor learning experiments are measuring – and changing – MN responses. Furthermore, although experiments specifically testing sensorimotor learning (in which sensory, motor, and sensorimotor experience are compared and/or controlled) have not been performed using single-unit recording, this conclusion is supported by single-unit data showing that experience with tools creates MN responses to observed tool use (Ferrari et al. 2005; M. J. Rochat et al. 2010; see sect. 4).

Considerations regarding anatomical specificity raise another possible objection to the sensorimotor training evidence: Sensorimotor experience may only affect neurons in non-classical mirror areas (e.g., dorsal PMC). However, while recordings of monkey MNs have mostly been confined to ventral PMC and IPL, measurements in humans using single-unit recording and fMRI conjunction suggest that MNs are more widespread (e.g., Arnstein et al. 2011; Gazzola & Keysers 2009; Landmann et al. 2011; Mukamel et al. 2010; Vogt et al. 2007). Furthermore, paired-pulse TMS indicates that functional connections from dorsal (as well as ventral) PMC to primary motor cortex enhance muscle-specific MEP responses to action observation (Catmur et al. 2011). Thus, several sources of evidence suggest that MNs are not restricted to classical mirror areas. Therefore, even if sensorimotor experience were altering neuronal responses only outside ventral PMC and inferior parietal cortex, it could still be affecting MNs. However, there is also evidence that sensorimotor learning affects classical mirror areas. Many studies have demonstrated effects of sensorimotor experience on classical

mirror areas (Cross et al. 2009; Haslinger et al. 2005; Landmann et al. 2011; Margulis et al. 2009; Vogt et al. 2007); counter-mirror sensorimotor learning reverses ventral PMC and inferior parietal cortex responses to observed actions (Catmur et al. 2008); and such learning is supported by ventral PMC-M1 connections (Catmur et al. 2011). Furthermore, localizing the effects of sensorimotor learning to specific neural populations using repetition suppression suggested that sensorimotor experience affects MNs in classical mirror areas (Press et al. 2012a). Therefore, it appears that MNs are not restricted to classical mirror areas and that sensorimotor experience has effects both on classical mirror areas and elsewhere.

A final possibility is that counter-mirror training changes relatively late neural responses to action observation, leaving earlier responses, mediated by MNs, intact (Barchiesi & Cattaneo 2013). Such a finding might indicate that counter-mirror responses result from a more indirect route (e.g., via prefrontal areas for rule retrieval) than mirror responses. Barchiesi and Cattaneo (2013) tested this hypothesis using a task that is likely to have provoked coding of domain-general spatial cues, rather than action-specific topographic cues, and therefore to have failed to index mirror responses at any time-point. A more recent study, using a more specific test of mirror responses, found effects of counter-mirror training on MEPs from 200 msec, the earliest time-point at which mirror responses have been observed in monkeys and humans (Cavallo et al. 2013; see also Catmur et al. 2011).⁴ Thus, effects of counter-mirror training occur at the time when complex information about the observed action has just reached PMC, making it improbable that mirror and counter-mirror effects occur at different times. It is likely that a prefrontal route is involved *during* the training session, when participants retrieve a rule in order to implement task instructions (e.g., “if index, do little”). However, the finding that *after* counter-mirror training, effects of training are present in MEPs from 200 msec suggests that such rule-based responding merely initiates associative learning; after learning, action observation activates counter-mirror responses as quickly as the original mirror responses.

In summary: Although there are currently no studies systematically testing the effects of sensorimotor learning on MN responses in monkeys, a substantial body of evidence from studies of training and expertise in humans has confirmed the predictions of the associative account, showing that mirror responses can be changed in radical ways by sensorimotor learning. Furthermore, these studies have provided no evidence that MNs/mechanisms are buffered or protected against sensorimotor experience of a kind that makes them code inanimate stimuli and dissimilar actions.

8. Other models: Canalization and exaptation

This article focuses on the genetic and associative accounts of the origins of MNs because these were the first models to be proposed, and the associative hypothesis is the most fully developed alternative to the standard, genetic view. For example, unlike other alternatives, it has been used to generate and test novel empirical predictions. However, two other alternatives, which have been motivated in part by the data generated in these tests (see sect. 7) should be considered. One raises the possibility

that the development of MNs is “canalized,” and the other that it represents an “exaptation” for action understanding. These are interesting possibilities but, we argue, they are not supported by the evidence reviewed in sections 4–7.

8.1. Canalization

It has been suggested that MNs are acquired through “Hebbian learning” (Keyser & Perrett 2004) and that their development is supported or “canalized” by genetically predisposed features of the perceptual-motor system, including the tendency of infants to look at their own hands in motion (Del Giudice et al. 2009). On one reading, this canalization hypothesis is identical in substance to the associative hypothesis; it is helpful in providing a more detailed neuronal model of how sensorimotor experience makes MNs out of motor neurons, and, in contrast with the associative hypothesis, it emphasizes self-observation over social interaction as a source of relevant sensorimotor experience in development, but otherwise the canalization hypothesis is identical to the associative account. On this reading, the term “Hebbian learning” is understood to be a synonym for “associative learning,” and the canalization hypothesis suggests that if the infants’ tendency to look at their own hands in motion is an adaptation (Clifton et al. 1994; Meer et al. 1996) – if this attentional bias evolved “for” anything – it was to promote the development of precise visuomotor control, rather than MNs and action understanding.

On another reading, which we think is less likely to represent the authors’ intentions, “Hebbian learning” differs from “associative learning” in depending on contiguity alone, rather than both contiguity and contingency (see sect. 3.2), and the infant preference for manual self-observation evolved specifically to promote the development of MNs and action understanding. If this reading is correct, the canalization hypothesis is a hybrid of the associative and genetic accounts; it claims that MNs develop through (Hebbian) sensorimotor learning *and* constitute a genetic adaptation for action understanding. However, this hybrid model would not be supported by current evidence for three reasons. First, there is no evidence that the tendency of infants to look at their own hands evolved to promote the development of MNs or action understanding rather than visuomotor control (Del Giudice et al. 2009). Second, experimental data and modeling work have indicated that the sensorimotor learning which changes MNs depends on contingency as well as on contiguity (Cook et al. 2010; Cooper et al. 2013b). Third, if MNs are forged by contingency-based sensorimotor learning, there is no problem for evolution (or scientists) to solve through canalization for MN development. If it was based on contiguity alone, there is a risk that sensorimotor learning would produce lots of “junk associations” – visuomotor neurons mapping observed and executed actions that happen to have co-occurred by chance. However, contingency-based (i.e., associative) learning could produce the observed distribution of strictly congruent, broadly congruent, and non-matching MNs all by itself, without MN-specific canalization (see sect. 5.1).

8.2. Exaptation

Another interesting hybrid of the genetic and associative hypotheses has been developed by Arbib and colleagues

(e.g., Arbib 2005; Oztop et al. 2006). They propose that MNs are produced, not by domain-general mechanisms of Hebbian or associative sensorimotor learning, but by a special kind of sensorimotor learning which receives input from self-observation of hand motion. This special kind of learning is an “exaptation” for action understanding: It evolved from more domain-general mechanisms, such as those producing canonical neurons, specifically to promote action understanding through the production of MNs. This exaptation hypothesis does not specify, in psychological or neurological terms, exactly what is distinctive about the kind of sensorimotor learning that produces MNs. However, it suggests that “some extra structure is required, both to constrain the variables relevant for the system, and to track trajectories of those relevant variables,” and that the function of this extra structure is to ensure coding of goals or “hand-object relationships” (Oztop et al. 2006, p. 269). Bonini and Ferrari (2011) recently advanced a similar exaptation hypothesis, also motivated by the need to explain why MNs consistently encode goals. However, as we have argued in section 4, the evidence from single-unit recording in monkeys suggests that MNs do not consistently encode goals. Therefore, the primary motivation for invoking exaptation is not compelling. Furthermore, there is no evidence that the sensorimotor learning involved in MN development is modified or constrained relative to the associative learning that occurs in standard conditioning experiments. On the contrary, there is experimental evidence that it is sensitive to contingency, subject to contextual modulation, and open to the encoding of both animate and inanimate stimuli in exactly the same way as standard associative learning (see sects. 5 and 7).

A recent article (Casile et al. 2011) adds another element to the hybrid model advanced by Arbib and colleagues. It suggests that a special, exapted form of sensorimotor learning underwrites the development of hand-related MNs, but the development of facial MNs is minimally dependent on experience. This suggestion is designed to accommodate evidence from studies of imitation and EEG suppression in newborns, which some authors have interpreted as showing that facial MNs are present at or shortly after birth. As we reported in section 6, this evidence has been challenged on a number of counts. Independent motivation and support for the idea that hand and face MNs have different origins would be provided by evidence that face MNs are less susceptible than hand MNs to modification by sensorimotor experience. However, as far as we are aware, this novel prediction of the hand/face hybrid model has not been tested, and a recent study of improvement in facial imitation suggests that face MNs are as susceptible to modification by sensorimotor experience as hand MNs (Cook et al. 2013). Thus, until it is used to generate and test novel predictions, the hand/face hybrid model stands as an intriguing but essentially ad hoc hypothesis.

Hybrid modelling is a promising direction for future research. However, to preserve predictive power, it is essential to check not only that hybrid models are consistent with existing data, but also that they have independent support. We have argued that both of these conditions are met by the associative account, and that neither is currently fulfilled by canalization and exaptation models.

9. A new approach to the function of mirror neurons

We have argued that, at present, there is no positive evidence that MNs are a genetic adaptation or exaptation, or that their development has been canalized, for action understanding. However, the associative hypothesis is functionally permissive; it does not deny that MNs make a positive – possibly even an adaptive – contribution to social cognition. Rather, the associative hypothesis implies that a new approach is required to find out what MNs contribute to social behavior.

9.1. From reflection to theory-based experimentation

In the 20 years since MNs were discovered, theories relating to their function have been inspired by a method which (if you like a pun) could be called “reflection.” This method focuses on the field properties of the MNs found in a sample of laboratory monkeys with unreported developmental histories. It asks, usually without reference to pre-existing computational or psychological theory, what neurons with these field properties would be “good for”; that is, what they might enable the animal to do. For example, early reports that MNs discharged when monkeys saw and produced object-directed actions inspired the theory that MNs mediate “action understanding” via “motor resonance,” when neither of these was an established category of psychological functioning. Even now, opposition to the idea that MNs mediate action understanding tends to be answered by stressing their field properties (Gallese et al. 2011). The associative account suggests that the reflection method needs to be changed and extended in three principal ways.

9.1.1. Developmental history. If MNs were a genetic adaptation, it is likely that their properties would be relatively invariant across developmental environments. Therefore, it would be possible to make valid inferences about species-typical properties of MNs based on a relatively small and developmentally atypical sample of individuals. If MNs are instead a product of associative learning, this kind of inference is not valid. Whether or not an individual has MNs, which actions are encoded by their MNs, and at what level of abstraction, will all depend on the types of sensorimotor experience received by the individual in the course of their development. Therefore, the associative account implies that it is crucial for studies of laboratory monkeys to report, and ideally to control, the animals’ developmental history; that is, the kinds of sensorimotor experience to which they have been exposed. It also suggests that, if we want to know the species-typical properties of monkey MNs, it will be necessary to test monkeys that have received all and only the types of sensorimotor experience typically available to them under free-living conditions. A corollary of this is that we cannot assume that the mirror mechanisms found in the members of one human culture are representative of the whole human species. With its emphasis on the role of social practices – such as the imitation of infants by adults, sports and dance training, and mirror self-observation – in driving the development of MNs, the associative account provides specific, theory-driven motivation for cross-cultural studies of mirroring.

9.1.2. System-level theory. If MNs were a genetic adaptation, one could argue that new categories of psychological functioning—such as “action understanding” and “motor resonance”—are necessary to characterize what they do. It could be argued that, since they were “specially created” by evolution, MNs are likely to have a highly distinctive, largely independent, and previously unrecognized psychological function. In contrast, by showing that established psychological theory—associative learning theory—can cast light on the origin of MNs, the associative account underlines the value of embedding research on MN function within system-level psychological and computational theories of how the brain produces behavior (Giese & Poggio 2003; Kilner 2011; Kilner et al. 2007a). This implies that hypotheses about MN function should specify a part in a process—a process that goes all the way from peripheral sensory input to overt motor output—that MNs are thought to fulfill. The name assigned to this part is not important in itself. What is important is that the hypothetical function of MNs is distinguished clearly from other components of the same overall process. For example, in this kind of system-level, theory-guided approach, “action understanding” would be distinguished from components that are likely to be more purely perceptual (which might be called “action perception” or “action recognition”), more purely motoric (e.g., “action execution”), or to constitute a higher level of “understanding” (e.g., “mentalizing”). This approach would also make it clear whether the hypothetical function is thought to be optional or obligatory; whether it can be, or must be, done by MNs. The kind of system-level theoretical approach required in research on the functions of MNs is exemplified by studies of their role in speech perception (Lotto et al. 2009; Scott et al. 2009).

A system-level theoretical approach would also overcome a problem that has haunted discussions of the “action understanding” hypothesis since MNs were discovered: Is this hypothesis claiming that MN activity causes or constitutes “action understanding”? The former is an empirically testable hypothesis suggesting that there is a distinctive behavioral competence (the nature of which has not yet been specified, see sect. 3.1), called “action understanding,” to which the activity of MNs contributes. The latter implies that the firing of MNs during action observation is, in itself, a form of “action understanding”; it does not need to have further consequences in order to qualify as “action understanding.” This claim is not subject to empirical evaluation; it is true, or otherwise, by virtue of the meanings of words.

9.1.3. Experimentation. Empirical (rather than constitutive) claims about the function of MNs need to be tested by experiments looking for, at minimum, covariation between MN activity and behavioral competence, and, ideally, testing for effects on behavioral competence of interventions that change MN activity. A brief survey of recent research of this kind—using fMRI, TMS, and the effects of focal brain lesions in human participants—is provided in the next section. At present, this research faces two major challenges. First, because the hypothetical functions of MNs typically are not defined in the context of a system-level theory, it is difficult to design appropriate control tasks. For example, if an experiment is testing the hypothesis that MNs play a causal role in action understanding,

should it control for the possibility that they instead play some role in action perception? If so, what kind of behavioral competence is indicative of action perception rather than action understanding?⁵ To date, only a small number of studies (e.g., Pobric & Hamilton 2006) include control conditions designed to address this issue.

The second major challenge is that, with rare exceptions (Mukamel et al. 2010), MN activity cannot be localized precisely within the human brain. Consequently, many studies assume that activity in the ventral PMC and IPL—areas homologous to those in which MNs have been found in monkeys—is MN activity, and that behavioral changes brought about through interference with the functioning of these areas are due to interference with MNs. The results of such studies are of interest regardless of whether they relate to MNs. However, it is unsatisfactory to assume that they relate to MNs, because, in monkeys, it is likely that fewer than 20% of the neurons in these classical mirror areas are actually MNs, and because there is evidence of MNs in non-classical areas in both monkeys and humans (see sect. 2.1). Techniques such as fMRI repetition suppression and TMS adaptation (Cattaneo et al. 2011; Silvanto et al. 2007) hold some promise as means of overcoming the localization problem with human participants, by isolating behavioral effects to specific populations of neurons. Guided by system-level theory, future studies could use these techniques with a range of tasks to isolate processes in which MNs are involved.

Alongside the development of techniques such as fMRI repetition suppression and TMS adaptation for use with human participants, it would be valuable to conduct animal studies that, not only document the field properties of MNs, but also examine how those properties relate to behavioral competence. For example, are animals with MNs for actions X and Y better than other animals of the same species at behavioral discrimination of X and Y, or at imitating X and Y? Studies of this kind have been dismissed as impractical on the assumption that they would have to involve monkeys, which are demanding and expensive laboratory animals, and that between-group variation in MN activity would have to be induced via lesions or disruptive TMS. However, the associative account suggests that, in the long term, it may be possible to overcome these problems by establishing a rodent model and using sensorimotor training to induce between-group variation in the number and type of MNs present in rodent brains. If the associative account is correct, rodents are likely to have the potential to develop MNs because they are capable of associative learning. Whether or not they receive in the course of typical development the sensorimotor experience necessary to realize this potential, it could be provided by various regimes of laboratory-based sensorimotor training.

9.2. Early signs

Given the theoretical limitations and methodological challenges faced by research to date on the functions of MNs, it is very difficult indeed to form a consistent and potentially reliable picture of where their findings are pointing. Nonetheless, for completeness, we offer the following brief overview of research relating to the two most commonly investigated hypotheses—MN activity contributes to action understanding and to imitation. We first summarize the results of meta-analyses of functional imaging

data, which, by identifying commonalities across multiple studies, highlight candidate brain areas involved in these hypothesized MN functions. Building on the fMRI studies, patient data and studies using disruptive TMS applied to classical mirror areas have the potential to identify causal, rather than purely correlational, relationships between MN activity and behavioral competence.

9.2.1. Action understanding. As discussed in section 3.1, consensus regarding the term “action understanding” has yet to be reached. In terms of the involvement of mirror areas in perceiving others’ actions, a recent meta-analysis of fMRI data on action observation revealed that observation of hand movements produces responses in both premotor and parietal cortex, while face and body movements recruit premotor and parietal cortex, respectively (Grosbras et al. 2012). If, however, “action understanding” is operationalized in terms of understanding others’ intentions, then the role of mirror areas is less clear: A meta-analysis of studies using mentalizing tasks concluded that classical mirror areas are not recruited unless the tasks involve action stimuli (Van Overwalle & Baetens 2009). Even when action stimuli are presented, intention understanding does not necessarily recruit mirror areas (e.g., Brass et al. 2007; de Lange et al. 2008), implying that the recruitment of mirror areas during mentalizing tasks that involve action stimuli is likely to be due to the action observation component of these tasks. Patient data indicate that components of action perception – including action detection, discrimination between observed actions, action recognition (e.g., naming observed actions), and object-action matching – may require classical MN areas (Buxbaum et al. 2005; Fontana et al. 2011; Kalenine et al. 2010; Moro et al. 2008; Pazzaglia et al. 2008; Saygin 2007; Saygin et al. 2004; Serino et al. 2010).

However, not all patients with impairments in action production are also impaired in action recognition, suggesting that motor ability may not always predict this aspect of “action understanding” (Calder et al. 2000; Negri et al. 2007), or at least that these abilities can dissociate, either prior to brain damage or via subsequent compensatory processes. Several TMS studies have demonstrated that PMC stimulation disrupts components of action perception including detection or discrimination of actions (Candidi et al. 2008; Urgesi et al. 2007b; van Kemenade et al. 2012), configural processing of bodies (Urgesi et al. 2007a), judgment of body aesthetics (Calvo-Merino et al. 2010), and motor-to-visual adaptation for observed actions (Cattaneo et al. 2011). It also impairs the ability to use information from perceived actions to judge the weight of grasped objects (Pobric & Hamilton 2006) and to initiate online predictions about ongoing actions (Stadler et al. 2012). It appears therefore that PMC may be necessary for some components of action perception, and this is an important result. However, given the current uncertainty about what is meant by “action understanding,” it is not clear whether these results consistently demonstrate a role for mirror areas in “action understanding” as opposed to a more perceptual process (see sects. 3.1 and 9.1).

9.2.2. Imitation. A recent meta-analysis of functional imaging studies found consistent responses during imitation tasks in classical MN areas, suggesting a possible functional involvement of MNs in imitation (Caspers et al. 2010; but

see Molenberghs et al. [2009] for inconsistent results in ventral PMC). Research investigating the effects of damage to parietal cortex supports a role for this area in imitation: for example, of mimed tool use (Halsband et al. 2001), of meaningless and object-related gestures (Buxbaum et al. 2005; Goldenberg & Karnath 2006; Tessari et al. 2007), and of phonetic detail in speech (Kappes et al. 2010). Additionally, damage to inferior frontal cortex results in impairments in imitation of finger movements (Goldenberg & Karnath 2006). Patient studies have used intentional imitation tasks that make many demands on the information processing system, in addition to the core imitation requirement to match sensory with motor representations of action (Leighton et al. 2008). Therefore, it is important that TMS studies have demonstrated that stimulation of classical mirror areas can disrupt both intentional and automatic imitation of simple finger and hand actions (Catmur et al. 2009; Heiser et al. 2003; Mengotti et al. 2013; Newman-Norlund et al. 2010).

Thus, the localization problem notwithstanding, insofar as current experimental data provide even early signs, they suggest that MN activity may make some contribution to action perception and imitation. However, the picture for action *understanding* is obscured by the fundamental problem of defining exactly what is meant by “action understanding” and how it differs from action perception. There is thus a pressing need for system-level theory to guide the design of control tasks in studies of action understanding. The picture is somewhat clearer in the case of imitation. The idea that MNs contribute to imitation was originally rejected on the grounds that monkeys do not imitate. However, guided by a leaner and more precise definition of imitation, based on system-level theory, subsequent studies have confirmed, not only that monkeys can imitate (Voelkl & Huber 2000; 2007), but also, as the associative account predicts, that distantly related species, such as budgerigars (Richards et al. 2009) and dogs (Range et al. 2011), are capable of imitation. Consistent with this, there is convergent evidence from meta-analyses, lesion studies, and TMS techniques implicating MN activity in a core component process that translates sensory input from body movement into a matching motor plan.

9.3. Conclusion

Like many other people, we find MNs intriguing. In Google Scholar “mirror neuron” scores some 11,000 hits, whereas “visuomotor neuron” scores 50. Some of this excitement may be ill-founded (Heyes 2010), but that is not the point of this target article. The associative account of the origin of MNs acknowledges that they were a fascinating scientific discovery, and this account is open to the possibility that MNs play one or more important roles in the control of social interaction. It differs from the received view in suggesting that (1) sensorimotor learning plays a crucial, inductive role in the development of MNs, and, because of this, (2) we will get reliable information about the function of MNs only by applying an approach based on developmental history, system-level theory, and rigorous experimentation. The first of these methodological implications underlines the fact that, relative to the genetic hypothesis, the associative account shifts the balance of explanatory power from MNs themselves to

the environments in which they develop. In some ways this is inconvenient because developmental environments are much harder to study in the laboratory, but there are significant potential payoffs. As a rich source of testable predictions about when, where, and how MNs develop, associative learning theory can provide firm guidance for future research on the taxonomic distribution, typical properties, and functional roles of MNs.

NOTES

1. The present article concerns what might be described as “motor” MNs – that is, MNs that are responsive during the observation and performance of actions. There may be neurons with analogous properties involved in empathetic emotional and somatosensory responses. This interesting possibility is beyond the scope of the present article, but for discussion of how an associative framework may be applied to the origins of empathic mirroring, see Heyes and Bird (2007).

2. Infant research suggesting that tongue-protrusion “imitation” improves over trials, in the absence of visual feedback, has been taken as evidence against the view that this “imitation” effect is really a nonspecific exploratory response (Soussignan et al. 2011). However, a recent experiment showing that even adults cannot improve their imitative performance in the absence of visual feedback suggests that the trends observed in the infant data may not have been signs of improvement (Cook et al. 2013).

3. Central alpha suppression is often seen alongside attenuation of beta band (~15–30 Hz) oscillations, and beta effects are thought to reflect motor processing. The sum of the two effects is defined as “mu suppression” (Hari & Salmelin 1997), but this term is often used more liberally in the MN literature, to refer to effects observed solely in the alpha band.

4. Monkey studies reporting this information suggest that MN responses start around 250 msec after observed movement onset (di Pellegrino et al. 1992; Kraskov et al. 2009; Umiltà et al. 2001). In humans, EEG and MEG data indicate that complex visual stimuli, including actions, reach PMC and motor areas in around 300 msec (Nishitani & Hari 2000; 2002; Proverbio et al. 2009; Sitnikova et al. 2003); while muscle-specific patterns of MEP facilitation are present from 200 msec after observed action onset (Cavallo et al. 2013).

5. The need to control for “action perception” is suggested by Rizzolatti and Sinigaglia’s definition of “action understanding” as understanding “‘from the inside’ as a motor possibility, rather than ‘from the outside’ as a mere visual description” (Rizzolatti & Sinigaglia 2010, p. 265). However, the potential circularity of defining perception as “merely visual” underlines the need for system-level theory.

genetics may specify how neurons may learn, not what they learn. Paying more attention to recent work linking mirror neurons to language acquisition and evolution would strengthen Cook et al.’s arguments against a rigid genetic hypothesis.

In the target article, Cook et al. claim that the genetic account of mirror neurons (MNs) is problematic and they propose an associativist alternative, arguing that “sensorimotor learning plays a crucial, inductive role in the development of MNs” (sect. 9.3). However, several models of sensorimotor learning for MNs have addressed neurophysiological findings (e.g., Bonaiuto & Arbib 2010; Bonaiuto et al. 2007; Keyser & Perrett 2004; Oztop & Arbib 2002; Oztop et al. 2013), so it seems the article is directed only to those who take MNs metaphorically and/or are unfamiliar with the primary literature.

It is true that any account proposing a rigid genetic fixation of MNs is incompatible with available evidence. Yet, Cook et al. fail to provide evidence that proponents of the genetic hypothesis are committed to such problematic accounts and some of their criticism seems to attack implausible “straw men” instead of existing accounts. For example, they claim that the “fact that these MNs respond maximally to unnatural stimuli – stimuli to which the evolutionary ancestors of contemporary monkeys could not possibly have been exposed – is hard to reconcile with the genetic hypothesis” (sect. 4.1, para. 3). The fact that “audiovisual” MNs are responding to “unnatural sounds” such as metal striking metal, plastic crumpling, or paper tearing associated with actions is allegedly problematic for the genetic hypothesis because no such sound/action connections existed for our distant ancestors and evolution could not have “acted on them.” The real issue is: “What would genetics specify?” Given that natural selection can only act on present conditions and pass currently beneficial traits on to the offspring, there can be a time lag between what an organism encounters in its environment and what it has been “genetically equipped” to deal with. So if our distant ancestors encountered *any* novel sound-action combination, a rigid mechanism that encodes only correlations as specific as those mentioned by Cook et al. would be of no use. One should reasonably expect then, that it is the ability to acquire MNs adapted to changing circumstances that is genetically specified, not *what* MNs code.

Cook et al. conclude that we can “get reliable information about the function of MNs only by applying an approach based on developmental history, system-level theory, and rigorous experimentation” (sect. 9.3). Given this reasonable conclusion, it is surprising that Cook et al. failed to pay adequate attention to recent work linking MNs to language acquisition and evolution (e.g., Arbib 2005; 2010; 2011; Arbib et al. 2008; Corballis 2010; Corina & Knapp 2008; Gentilucci & Corballis 2006; Ramachandran 2000; for a skeptical view, see Bickerton 2007). Language acquisition seems to offer an excellent opportunity to gather evidence against a rigid genetic hypothesis (e.g., language acquisition accounts defended by: Chomsky 1981; 1995; 2012; Legate & Yang 2002; Lightfoot 1999; McGilvray 2006; Pietroski & Crain 2005; Pinker 1994). According to this framework, all humans possess “some innate mental state common to the species that provides the basis for acquisition of knowledge of grammar” (Chomsky 1981, p. 224). The interesting question is how the intricate details of linguistic knowledge might be genetically encoded. A healthy infant can acquire any human language. Therefore, she needs to be able to imitate both sounds that are the same as, and very different from, those her ancestors have acquired. Further, the English acquired by an infant born in 2013 differs greatly from that of an infant born in the year 848. The differences between those sounds are arguably as great as the difference between the sound of branches breaking (natural) and plastic crumbling (artificial) discussed above. It is indeed implausible that any genetically fixated mechanism could underwrite such highly flexible imitation. Such considerations lead to the abandonment of the proposal that the “innate endowment consists of a

Open Peer Commentary

The role of mirror neurons in language acquisition and evolution

doi:10.1017/S0140525X13002203

Christina Behme

Department of Philosophy, Dalhousie University, Halifax, NS B3H 4R2, Canada.

Christina.Behme@dal.ca

Abstract: I argue that Cook et al.’s attack of the genetic hypothesis of mirror neurons misses its target because the authors miss the point that