WHAT HAPPENED TO MIRROR NEURONS?

Cecilia Heyes & Caroline Catmur

All Souls College & Department of Experimental Psychology
University of Oxford
Oxford OX1 4AL
United Kingdom
cecilia.heyes@all-souls.ox.ac.uk

Department of Psychology
Institute of Psychiatry, Psychology & Neuroscience
King’s College London
De Crespigny Park
London SE5 8AF
United Kingdom

Acknowledgements
This work was supported by All Souls College, Oxford, and the Leverhulme Trust (grant number PLP-2015-019 to CC).
Abstract

Ten years ago, *Perspectives on Psychological Science* published the Mirror Neuron Forum, debating the role of mirror neurons in action understanding, speech, imitation and autism, and asking whether mirror neurons are acquired through visual-motor learning. Subsequent research on these themes has made significant advances, which should encourage further, more systematic research:

- **Action understanding** - Multivoxel pattern analysis, patient studies, and brain stimulation suggest that mirror neuron brain areas contribute to low-level processing of observed actions (e.g. distinguishing types of grip), but not to high-level action interpretation (e.g. inferring actors’ intentions).
- **Speech perception** – Although it remains unclear whether mirror neurons play a specific, causal role in speech perception, there is compelling evidence for the involvement of the motor system in the discrimination of speech in perceptually noisy conditions.
- **Imitation** – There is strong evidence from patient, brain stimulation and brain imaging studies that mirror neuron brain areas play a causal role in copying of body movement topography.
- **Autism** – Studies using behavioural and neurological measures have tried and failed to find evidence supporting the “broken mirror” theory of autism. Furthermore, research on the origin of mirror neurons has confirmed the importance of domain-general visual-motor associative learning, rather than canalised visual-motor learning, or motor learning alone.
Ten years ago, mirror neurons were everywhere. In 2011, when *Perspectives on Psychological Science* published a Forum discussing the functions and origins of these fascinating cells (Gallese, Gernsbacher, Heyes, Hickok & Iacoboni, 2011; Glenberg, 2011a; Glenberg, 2011b), mirror neurons featured in *Time* magazine and the *New York Times*; programs about mirror neurons were broadcast by CNN and the BBC; and more than 200 articles were published in academic journals implicating mirror neurons in, among other functions: action understanding, alexithymia, autism, business management, empathy, imitation, language comprehension, language production, literary mimesis, post-traumatic stress disorder and schizophrenia. Measured by number of academic publications, interest in mirror neurons peaked two years later in 2013 and then began to decline (Figure 1). Of course, the numbers in Figure 1 are not an infallible measure of scientific interest in mirror neurons. Since 2013 researchers may have begun to use other terms for the same targets of investigation. However, Figure 1 suggests that cognitive scientists are no longer working as actively in this field, or that the mirror neuron “brand” is losing its appeal, or both, and therefore raises the question: What happened to mirror neurons?

![Figure 1. Number of papers published per year from 1996-2019 including the term “mirror neuron” in the title, abstract, or keywords. Data from Scopus, 12th May 2020.](image)

Given the extent of public interest in mirror neurons, and the liveliness of the controversy they provoked among scientists and philosophers, this question could be fruitfully interpreted in an
historical and sociological way. We could ask about the currents in Western society and in contemporary cognitive science that first made mirror neurons “mesmerising” (Heyes, 2010) and then weakened their appeal. But this article takes a more straightforward, natural science approach. After a brief introduction to mirror neurons, we use the questions discussed in the Mirror Neuron Forum (Gallese et al., 2011), concerning the functions and origins of mirror neurons, to structure a succinct survey of research published in the last 10 years. We then consider whether these recent findings have taken the shine off mirror neurons and, if so, whether that reaction is appropriate. We conclude that, although the results of careful empirical research were bound to be disappointing relative to the more grandiose claims, recent work on mirror neurons should encourage further systematic investigation.

**Mirror neuron basics**

Mirror neurons were discovered by chance in monkeys in 1992 and given their evocative name four years later (di Pellegrino, Fadiga, Fogassi, Gallese, & Rizzolatti, 1992; Gallese, Fadiga, Fogassi, & Rizzolatti, 1996). Early studies of the field properties of mirror neurons – the sensory and motoric conditions in which they fire – revealed three basic types: “Strictly congruent” mirror neurons discharge during execution and observation of the same action; for example, when the monkey performs a precision grip and when it passively observes a precision grip performed by another agent. “Broadly congruent” mirror neurons 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” mirror neurons 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 (di Pellegrino et al., 1992). Strictly and broadly congruent mirror neurons were, from the beginning, of primary interest, and they are what we and most other researchers mean when they use the term “mirror neuron”. These cells are intriguing because, like a mirror, they match observed and executed actions; they code both “my action” and “your action”.

Monkey mirror neurons 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, Gallese, Rizzolatti, & Fogassi, 2003).
Mirror neurons were originally found using single cell recording in area F5 of the ventral premotor cortex (di Pellegrino et al., 1992; Gallese et al., 1996) and the inferior parietal lobule (Bonini et al., 2010; Fogassi et al., 2005) of the monkey brain. Subsequently they were found, not only in these “classical” areas, but also in non-classical areas, including primary motor cortex (Dushanova & Donoghue, 2010; Tkach, Reimer, & Hatsopoulos, 2007) and dorsal premotor cortex (Tkach et al., 2007).

Early research with human participants used functional magnetic resonance imaging (fMRI) to show spatial overlap in the areas of ventral premotor cortex and inferior parietal lobule that are active when people observe and execute movements (Rizzolatti et al., 1996; Decety et al., 1997; Buccino et al., 2001; Grezes and Decety, 2001). This was not conclusive evidence of the existence of human mirror neurons because the spatial overlap could have been due, not to neurons that each respond both to observation and execution of action (mirror neurons), but to clusters of neurons each responding either to action observation or to action execution (Dinstein et al., 2007, 2008). By 2011, doubts about the presence of mirror neurons in the human brain had been assuaged by studies using single cell recording in pre-surgical patients (Mukamel, Ekstrom, Kaplan, Iacoboni, & Fried, 2010), and the repetition suppression fMRI procedure in healthy volunteers (Kilner et al., 2009). Like monkey mirror neurons, evidence consistent with the existence of human mirror neurons has been found in both classical areas – ventral premotor cortex and inferior parietal lobule – and non-classical areas, including dorsal premotor cortex, superior parietal lobule, cerebellum (Molenberghs et al., 2012), supplementary motor area, and medial temporal lobe (Mukamel et al., 2010).

Research using single cell recording (and, to some extent, repetition suppression) suggests that mirror neurons are typically present in adult human brains. However, this research does not licence the inference that mirror neurons are always or usually responsible for spatial overlap in fMRI responses during the observation and execution of action. Consequently, it has become common to use terms such as “the mirror neuron system” and “mirror neuron brain areas” to refer to regions of the brain that are active during action observation and execution and/or for which there is evidence of the presence of mirror neurons. As reviewers, we have adopted the latter of these conventions, but it should be noted that these terms are unsatisfactory in at least two respects. First, it is not clear in what sense the areas containing mirror neurons constitute a “system”. Second, it is likely that only a small proportion of the neurons in each of these areas have mirror properties: for example, fewer than 10% of the neurons studied in di Pellegrino et al.’s (1992) seminal paper showed “strict” or “broad” congruence in their firing patterns to observed and executed actions, and although some single unit studies report higher proportions of mirror
neurons, the nature of the single-unit technique makes it problematic to estimate the true prevalence of mirror neurons in any brain area (see also Kilner & Lemon, 2013). These issues should be borne in mind when considering the results of neuroimaging and neurostimulation studies discussed in this review.

From their discovery in 1992, theorizing about the function of mirror neurons was dominated, not by computational modelling and experimental intervention, but by consideration of their field properties. Defining functions broadly and in everyday language, researchers reflected on what neurons responsive to similar observed and executed actions would be “good for”; what kinds of psychological task they might be able to fulfil. In some cases, this strategy produced hyperbole. Mirror neurons were hailed as “cells that read minds” (Blakesee, 2006), “the neurons that shaped civilization” (Ramachandran, 2009), and a “revolution” in understanding social behavior (Iacoboni, 2008). But most researchers, including the group in Parma that discovered mirror neurons, focussed on four realistic possibilities: action understanding, speech perception, imitation, and (their potential dysfunction in) autism. Each of these four hypotheses about the function of mirror neurons was debated in the Mirror Neuron Forum (Gallese et al., 2011; henceforward “the Forum”), alongside a key question about their origins: Do mirror neurons get their characteristic visual-motor matching properties from learning?

Functions

Action understanding

In his lucid summary analysis of the Forum, Glenberg (2011b) concluded in relation to action understanding that there was broad agreement that mirror neurons, or a “mirror neuron system”, “plays some role in action processing” but no consensus about what that role might be. It could be relatively low-level; mirror neurons may contribute to action selection or to action recognition, helping to distinguish one type of action from another (e.g. precision grip from power grip). Alternatively, mirror neurons may have a high-level function in action processing, enabling “understanding from within” (Rizzolatti & Sinigaglia, 2010) or inferences about actors’ mental states. Since 2011, advances in addressing this issue have come mainly from two broad lines of evidence: the use of multivoxel pattern analysis in fMRI to “decode” the information represented within and across brain areas; and the use of patient studies and neurostimulation to investigate the causal role of mirror neuron brain areas for action understanding.

Multivoxel pattern analysis has revealed that “mirror” areas including premotor cortex encode concrete representations of observed actions (e.g. the action involved in opening a particular
bottle), rather than abstract, higher-level representations (e.g. the goal “to open”; Wurm & Lingnau, 2015; Wurm & Caramazza, 2019). These findings are consistent with the involvement of mirror neurons in lower-level processing of observed actions.

However, any such involvement does not entail that mirror neurons play a causal role in action processing. The evidence in this case is still rather mixed. Studies of individuals born without upper limbs indicate that action recognition can take place without motor representations of the relevant effectors (Vannuscorps & Caramazza, 2016), and some stroke patients can identify actions despite damage to mirror neuron brain areas (Tarhan, Watson & Buxbaum, 2015); although a meta-analysis of previous findings indicates that such patients do show impairments in action identification (Urgesi, Candidi & Avenanti, 2014).

Patient studies can be hard to interpret, however, due to heterogeneity in the damage incurred and the acquisition over time of compensatory strategies. Temporary disruption of brain function using neurostimulation can therefore provide convergent evidence that a particular brain area plays a causal role in relation to a particular cognitive function. One prominent neurostimulation study indicated that premotor cortex was necessary for identification of the intentions underlying observed actions (Michael et al., 2014). However, in that study, premotor cortex stimulation also disrupted perceptual matching of the observed actions. It is possible therefore that the disruption to intention identification was the result of disruption to low-level action processing, and that premotor cortex does not play a direct, causal role in intention reading (Catmur, 2014).

On the basis of evidence including that summarized above, Thompson and colleagues’ (Thompson, Bird & Catmur, 2019) recent review of the putative contribution of mirror neurons to action understanding concluded that any involvement of mirror neurons appears to be confined to lower-level processing of observed actions (e.g. aiding action discrimination or recognition). In particular, they found no compelling evidence for the involvement of mirror neurons, or mirror neuron brain areas, in higher-level processes such as inferring others’ intentions from their observed actions.

Speech perception

Four contributors to the Forum (Gallese et al., 2011) - Gernsbacher, Gallese, Hickok, and Iacoboni - agreed that the motor system has some role in speech perception, but they disagreed about the type and magnitude of the motor system’s role, and about whether mirror neurons in particular are important (Glenberg, 2011b).
Neuroimaging data indicate that mirror neuron brain areas respond during speech perception. For example, Callan and colleagues (2014) demonstrated that responses in ventral premotor cortex during a vowel identification task were enhanced when signal-to-noise was reduced, suggesting that such responses may improve speech discrimination in perceptually noisy conditions. Measurement of motor cortex excitability has also been used to investigate the involvement of motor areas in speech perception. Motor cortical representations of speech effectors (e.g. lips or tongue) are enhanced during the perception of speech in noise (Nuttall et al., 2016, 2017; but see Panouilleres et al., 2018, for a contrasting result). This enhancement may have functional implications for speech perception ability: participants with greater motor mirroring of perceived speech showed better ability to discriminate speech in noise (d’Ausilio et al., 2014).

However, evidence from patient studies casts doubt on whether the motor system is causally involved in speech perception. If speech perception requires perceived speech to be matched with motor commands for the production of speech, then one should find speech perception impairments in patients who, due to brain lesions, have impairments in speech production. In contrast to this prediction, a series of studies have demonstrated intact speech sound discrimination in patients with speech production difficulties (Hickok et al., 2011; Rogalsky et al., 2011; Stasenko et al., 2015).

A final line of evidence comes from brain stimulation studies. Restle and colleagues (2012) demonstrated that facilitatory brain stimulation to the inferior frontal gyrus – a classical mirror neuron brain area - improved participants’ accuracy in repeating unfamiliar foreign speech sounds. They argue that this indicates the role of the inferior frontal gyrus in matching perceived to produced speech; but this result cannot tell us whether the crucial role of this brain area is in speech perception, speech production, or the matching process itself. Furthermore, as this study did not include a control task, it is unclear whether stimulation of this brain area had a specific effect on speech processing, or whether any complex sensorimotor task might have been improved by such stimulation. However, a series of subsequent studies have shown that stimulation of motor (Rogers et al., 2014; Smalle et al., 2015) or premotor cortex (Nuttall et al., 2018) affects speech perception ability, in particular for distorted speech (Nuttall et al., 2018); and one of these studies included a control task permitting the conclusion that the stimulation had a specific effect on perception of speech but not non-speech sounds (Rogers et al., 2014).

In summary, there appears to be reasonably strong evidence for the involvement of the motor system (including premotor mirror neuron brain areas, as well as motor cortex) in the discrimination of speech in perceptually noisy conditions. However, this conclusion is not yet
supported by the patient data. A priority for future research, therefore, is to test whether patients with premotor lesions are impaired at discrimination of speech from noise.

*Imitation*

Of the functions discussed in the Forum, imitation attracted the strongest consensus. It was agreed that, although early work on the relationship between mirror neurons and imitation had involved some dubious definitions and inferences, “when imitation [is] defined in terms of action topography [how body parts move relative to one another], most agree mirror neurons contribute” (Glenberg, 2011b). This consensus was due in large measure to two studies showing that repetitive transcranial magnetic stimulation (rTMS, a disruptive intervention) of the inferior frontal gyrus, a mirror neuron brain area, selectively impaired imitative behaviour (Catmur, Walsh, & Heyes, 2009; Heiser, Iacoboni, Maeda, Marcus, & Mazziotta, 2003).

Causal methodologies, including brain stimulation and patient studies, have continued to support the consensus that mirror neuron brain areas contribute to imitation. Two studies using facilitatory brain stimulation to inferior frontal gyrus demonstrated improvements in vocal imitation and naturalistic mimicry (Restle et al., 2012; Hogeveen et al., 2015), while inhibitory stimulation to the inferior parietal lobule slowed participants in an instructed imitation task (Reader et al., 2018). Inhibitory stimulation of the inferior frontal gyrus also disrupted automatic imitation (Newman-Norlund et al., 2010) but this effect was not specific to human body movements, with similar disruption being found for non-biological stimuli.

Binder and colleagues (2017) demonstrated that apraxic patients were impaired on an instructed imitation task, and that this impairment was associated with lesions to a set of brain areas thought to contain mirror neurons, including the left postcentral gyrus, intraparietal sulcus, and inferior frontal cortex. A similar result was reported by Frenkel-Toledo et al. (2016) with imitation impairments associated with lesions to the left inferior and superior parietal lobules and postcentral gyrus.

Data from causal studies such as these have been complemented by a series of fMRI studies over the last decade which demonstrate greater responses in mirror neuron brain areas during imitation than during other closely matched tasks (e.g. Mengotti et al., 2012; Ocampo et al., 2011; Mainieri et al., 2013; Campbell et al., 2018).

*Autism*
In marked contrast with imitation, the Forum unearthed “huge disagreement on autism” (Glenberg, 2011b). Although Gallese defended the view that “impaired motor cognition”, rather than mirror neurons, contributes to the autistic1 phenotype, Iacoboni stood by the more specific “broken mirror” theory (Iacoboni & Dapretto, 2006; Oberman & Ramachandran, 2007; Williams, Whiten, Suddendorf & Perrett, 2001; but see Southgate & Hamilton, 2008), citing a range of brain imaging studies suggesting that people with autism have abnormal activity in mirror neuron areas of the brain.

Gernsbacher contested all of this evidence, highlighting methodological problems and replication failures, while Heyes, also sceptical about the broken mirror theory, focussed on evidence that people with autism show intact (Bird et al., 2007; Cook & Bird, 2012; Gowen et al., 2008; Press et al., 2010), and sometimes exaggerated (Spengler, Bird & Brass, 2010), automatic imitation.

A systematic review in 2013 of neuroscientific evidence concluded that there was “little evidence for a global dysfunction of the mirror system in autism” (Hamilton, 2013). Studies published since that review, using a range of techniques to investigate neural responses to observed actions, and imitation (the cognitive function thought to rely most strongly on such responses), support this conclusion.

A popular technique to measure motor system responses during action observation is “mu” suppression, an electroencephalographic (EEG) measure of the reduction in coherence in certain frequency bands which occurs both when performing and when observing actions. However, the use of this technique to index mirror neuron responses has been criticised on the grounds that it measures attentional (Hobson & Bishop, 2016, 2017) and somatosensory (Coll et al., 2015, 2017), rather than motor, responses. Notwithstanding these critiques, recent studies of mu suppression during action observation in autism have shown no differences from neurotypical controls (Bernier et al., 2013; Ruysschaert et al., 2014).

In contrast, fMRI of neural responses during action observation indicates some differences between autistic and neurotypical participants. A recent meta-analysis of six fMRI studies of action observation and imitation indicated that participants with autism showed greater responses in bilateral fronto-parietal regions than those without (Yang & Hofmann, 2016), although another study showed no differences in neural responses during action observation (Pokorny et al., 2015). These data suggest some differences in responses during action observation, but they are not consistent with a broken mirror account of autism; if anything they point to greater neural responses in mirror neuron brain areas during action observation in those with, compared to those without, autism. Furthermore, although these responses are in brain areas thought to contain mirror neurons.

---

1 The use of the term ‘autistic’ is endorsed by many individuals with autism (see Kenny et al., 2016). We therefore use this term as well as ‘person-first’ language (e.g., ‘people with autism’) to respect the wishes of autistic individuals.
neurons, none of these papers used techniques that permit the definitive conclusion that such responses are due to mirror neurons, as opposed to other neurons co-located in these brain areas.

Techniques that measure motor representations of specific actions (e.g. grasping) are potentially more informative in this respect. Recent studies using motor-evoked potentials to measure motor cortical excitability during action observation provide mixed evidence for differences between autistic and neurotypical participants. Enticott and colleagues (2012) reported a smaller increase in motor cortical excitability in participants with autism compared to neurotypical controls when viewing hand grasps. However, a later study found no differences on the same measure when observing socially relevant hand actions (Enticott et al., 2013). Previous claims of disrupted mirror responses in autism based on electromyographic evidence (Cattaneo et al., 2007) have also come under scrutiny, with a series of methodological critiques casting doubt on the interpretation of previous data (Pascolo & Cattarinussi, 2012; Ruggiero & Catmur, 2018).

Finally, behavioural measures of imitation have also provided little evidence for a broken mirror account of autism. Although one study reported some differences in instructed imitation (Cossu et al., 2012), a task with many demands unrelated to mirror neurons, the majority of recent studies have found either no difference in imitation between autistic and neurotypical participants, or greater imitation in people with autism (Sowden et al., 2016; Schunke et al., 2016; Schulte-Ruther et al., 2017; Gordon et al., 2020). Overall, therefore, the last ten years of research have produced no compelling evidence for the claim that autism is associated with mirror neuron dysfunction.

Origins

Contributors to the Forum also debated the origins of mirror neurons, addressing the question: Do mirror neurons get their characteristic visual-motor matching properties from learning? Heyes argued that mirror neurons get their matching properties via standard mechanisms of sensorimotor associative learning. They start out as motor neurons, active only during the performance of action. Then, through correlated experience of seeing and doing the same actions – in the context of self-observation (e.g. an infant watches her hand own in motion) and social interactions in which the same movements are repeatedly both observed and executed (e.g. pat-a-cake; Heyes 2001) – these motor neurons become strongly connected to visual neurons tuned to similar actions. Consequently, what was once a motor neuron becomes a mirror neuron – responsive to both the sight and performance of an action. Iacoboni agreed that visual-motor learning is likely to be important but saw signs that it is “canalized” by a genetic pre-disposition to develop mirror neurons (Del Guidice et al., 2009). Gallese went further, arguing that there is an “innate” or “genetically pre-
determined” propensity to develop mirror neurons which is facilitated, not primarily by visual-motor learning, but by motor experience before and after birth (Gallese et al., 2009). Summarising the debate about origins, Glenberg (2011b) identified as a key question “the degree to which neonatal imitation is a reliable phenomenon”.

**Sensorimotor learning**

In the last decade, more evidence has emerged that learning plays an important role in the development of mirror neurons (Brunsdon et al., 2019; Catmur et al., 2011; Copete et al., 2016; de Klerk et al., 2015; Fitzgibbon et al., 2016; Furukawa et al., 2017; Guidali et al., 2020; Hou et al., 2017; McKyton et al., 2017; Orlandi et al., 2017; Press et al., 2012a; Wiggett et al., 2012; Zazio et al., 2019). Some of the recent studies report greater activity in mirror neuron brain areas in pianists and dancers than in people who lack such expertise during observation of musical performance and dance, respectively (Furukawa et al., 2017, Hou et al., 2017; Orlandi et al., 2017). These studies are of interest because they indicate that activity in mirror neuron brain areas is affected by long-term learning under naturalistic conditions, but they do not tell us what kind of learning is important. For example, dancers may show greater mirror neuron brain area activity than controls during dance observation because the dancers have watched more dance movements (sensory learning), performed more dance movements (motor learning), and/or watched more dance movements while performing similar dance movements (sensorimotor learning).

Experiments that were designed to isolate the kind of learning involved in mirror neuron development have suggested that sensorimotor learning is crucial (Catmur et al., 2011; de Klerk et al., 2015; Press et al., 2012a; Fitzgibbon et al., 2016; Guidali et al., 2020; Wiggett et al., 2012). For example, replicating and extending earlier work with a similar design (e.g. Catmur et al., 2007), Wiggett et al. (2012; see also Brunsdon et al., 2020) found using fMRI that mirror neuron brain areas were more strongly activated by observation of hand movement sequences in participants who had simultaneously observed and executed the movements (sensorimotor learning) than in participants who had either observed the movements without performing them (sensory learning), or performed the movements without observing them (motor learning). Furthermore, using paired pulse TMS, Catmur et al. (2011) confirmed that novel sensorimotor experience of the kind given by Wiggett et al. acts on mirror responses. They showed that “counter-mirror” training (in which participants performed index finger movements while observing little finger movements, and vice versa) reversed mirror responses (e.g. resulted in greater activation of an index finger muscle during observation of little finger movement than during observation of index finger movement) via the
same connections between premotor and motor cortex that were responsible for mirror effects prior to training (e.g. greater activation of an index finger muscle during observation of index than little finger movement).

Responding to a study in which counter-mirror training yielded later effects on motor excitability than mirror training (Barchiesi & Cattaneo, 2013; see also Ubaldi et al., 2015), Cavallo et al. (2013), like Catmur et al. (2011), found that mirror responses and counter-mirror responses followed the same time course. Another TMS study found that counter-mirror responses can be induced by instruction alone (Bardi et al., 2015). However, it is unlikely that instructional learning, rather than sensorimotor learning, was responsible for the training effects reviewed above because 1) in the study by Bardi et al. (2015) participants were tested immediately after instruction, whereas in the earlier studies they were tested 24 hours after both instruction and sensorimotor training, and 2) training effects have been observed in uninstructed infants (De Klerk et al., 2015).

The training studies described above involved adult participants, but there is also now evidence that sensorimotor learning is important in the early development of mirror responses. Using EEG recordings of sensorimotor alpha suppression as an index of mirror neuron activity, de Klerk et al. (2015) found in 7-month-old infants, who could not yet walk, that mirror responses during observation of stepping movements increased with the amount of sensorimotor experience they had received in earlier training sessions. Infants who had frequently seen their own stepping movements while performing those movements showed greater alpha suppression than infants who had relatively little correlated experience of seeing and doing the stepping movements. De Klerk et al. did not find similar effects of sensory experience (observing stepping) or motor experience (making stepping movements) on alpha suppression, suggesting that, at least in this study, alpha suppression indexed the sensorimotor matching function of mirror neurons rather than purely attention or arousal.

The idea that motor learning alone is sufficient to change the properties of mirror neurons (Gallese et al., 2009) has not been supported. The study by de Klerk et al. (2015) failed to find a relationship between the frequency with which infants performed stepping movements during training and the extent of sensorimotor alpha suppression during observation of stepping movements after training. Yet more striking, using imitation as a behavioural index of mirror neuron activity, McKyton et al. (2017) found reduced automatic imitation in newly sighted children who suffered from dense bilateral cataracts from early infancy and were surgically treated only years later. These children, who had been deprived of sensory and sensorimotor experience of action, but not of motor experience, were less inclined than control children to imitate task-irrelevant hand actions.
In addition to showing that sensorimotor learning is important for mirror neuron development, recent research suggests that, in everyday life, much of this learning occurs in the context of social interactions between infants and their caregivers. The extent to which mothers imitate infant facial expressions at two months postpartum predicts EEG alpha suppression at nine months during observation of the same facial expressions (Rayson et al., 2017; see also Markodimitraki & Kalpidou, 2019; Murray et al., 2018). Furthermore, this relationship is action specific. De Klerk et al. (2019) found that parental imitation of facial expressions predicted infant imitation of facial but not hand movements, implying that parental imitation supports mirror neuron development through learning of specific sensorimotor associations (e.g., between the sight and performance of mouth opening) rather than by enhancing attention to body movements or social motivation.

**Is the sensorimotor learning genetically canalized?**

The importance of sensorimotor learning for the development of mirror neurons is now well established, but questions remain about the character of this learning. The “associative account” maintains that the sensorimotor learning that builds mirror neurons is of exactly the same kind as the learning that produces Pavlovian and instrumental conditioning; it is a computationally undemanding, domain-general process that forges excitatory and inhibitory links between simple event representations (Cook et al., 2014; Heyes, 2001; Keysers & Perrett, 2004). In contrast, the “canalization account” (Del Guidice et al., 2009; Gallese et al., 2009), supported by Iacoboni and Gallese in the Forum, suggests that monkeys and humans genetically inherit a specific propensity to acquire mirror neurons. On this view, the sensorimotor learning that contributes to mirror neuron development is domain-specific – it involves computations distinct from those involved in standard conditioning – and/or the learning is primed for the development of mirror neurons; given a head start by genetically inherited behavioural mechanisms.

A few studies in the last 10 years have addressed the domain-generality of the sensorimotor learning involved in mirror neuron development. Consistent with the associative account, these have indicated that, like standard conditioning, mirror neuron learning depends on contingency as well as contiguity (Cooper et al., 2013), and shows a distinctive pattern of contextual modulation (Cook et al., 2012a). However, as Glenberg (2011b) predicted, the majority of research bearing on the associative and canalization accounts has focussed on neonatal imitation. Given the evidence that mirror neurons contribute to imitation (see above), reliable evidence that newborns can imitate, before they have had the opportunity for relevant sensorimotor learning, would suggest that the development of mirror neurons is canalized or genetically pre-determined.
A set of 10 studies from one research group, eight of them published since the Forum, have claimed to provide evidence of imitation in newborn monkeys (Ferrari, Visalberghi, Paukner, Fogassi, Ruggiero & Suomi, 2006; Ferrari, Paukner, Ruggiero, Darcey, Unbehagen & Suomi, 2009; Paukner, Ferrari & Suomi, 2011; Simpson, Paukner, Sclafani, Suomi & Ferrari, 2013; Paukner, Simpson, Ferrari, Mrozek & Suomi, 2014; Simpson, Paukner, Suomi, & Ferrari, 2014; Simpson, Miller, Ferrari, Suomi & Paukner, 2016; Kaburu, Paukner, Simpson, Suomi & Ferrari, 2016; Paukner, Pedersen & Simpson, 2017; Wooddell, Simpson, Murphy, Dettmer, Paukner, 2019). These studies did not use the “cross-target” procedure which has been agreed, by both enthusiasts and sceptics, to be necessary to detect imitation in newborns (e.g. Redshaw, 2019; Meltzoff & Moore, 1977; Ray & Heyes, 2011; Meltzoff, 1996; Oostenbroek et al., 2014; Whiten, 2002). For example, when testing for imitation of tongue-protrusion and lip-smacking, they did not use these behaviours as controls for one another. Instead of looking for a higher frequency of tongue protrusion than of lip-smacking in infants who had just observed tongue protrusion, and a higher frequency of lip-smacking than of tongue protrusion in infants who had just observed lip-smacking, they reported, for example, a higher frequency of tongue protrusion after observation of tongue protrusion than after observation of a rotating disk. An effect of this kind could be due, not to imitation of tongue protrusion, but to a biological, social stimulus eliciting more behaviour of all kinds than a nonbiological, asocial stimulus. Pointing out this problem alongside a number of others (e.g. multiple comparisons without correction), Redshaw (2019) re-analysed the data from the full corpus of 10 neonatal monkey studies. Applying the cross-target methodology, the re-analysis found no evidence whatever of imitation in newborn monkeys.

Recent work with human neonates points in the same direction. In a study with unprecedented power, conducted in Brisbane, Oostenbroek et al. (2016) tested more than 100 infants longitudinally at 1, 3, 6 and 9 weeks of age, in a cross-target procedure involving a wide range of targets. They recorded the frequencies of nine target actions — tongue protrusion, mouth opening, happy expressions, sad expressions, index finger protrusion, grasping, MMM sound, EEE sound, tongue click — while infants observed eleven movement stimuli — an adult performing each of the nine actions, and two object movements (spoon protruding through a tube and box opening). The results of the Brisbane study were wholly negative: in no case did the infants consistently perform a target action more often while observing the same action than while observing all of the alternative actions.

Previous failures to find neonatal imitation have been attributed to methodological factors – for example, to the use of an inappropriate model, an inadequate response interval, or sub-optimal statistical procedures. A recent meta-analysis of neonatal imitation research by the Brisbane group,
encompassing 336 effect sizes dating back to 1977, sought and did not find a modulating influence of 13 methodological factors previously cited as reasons for replication failure (Davis et al., in press). However, the meta-analysis did find a modulating effect of “researcher affiliation”; a small number of laboratories are more likely than others to find large positive effects. Furthermore, across the whole data set there was a relationship between standard error and effect size indicative of publication bias; i.e. suggesting that smaller studies have been conducted, found no evidence of neonatal imitation, and not been published.

Re-analysing the data from the Brisbane study, using a more liberal statistical method, Meltzoff et al. (2018; see also Oostenbroek et al., 2018) found evidence of imitation for one of the nine target actions – tongue protrusion. Although consistent with other reviews and meta-analyses of neonatal imitation data (e.g. Anisfeld, 1996; Jones, 2006; Ray & Heyes, 2011), this result does not uphold the historical claim that newborns are capable of voluntary imitation of a range of actions (Meltzoff & Moore, 1977; see also Keven & Akins, 2017), or support the view that mirror neurons are learned via a canalised or genetically pre-determined process.

Rather than supporting canalisation, one study of young human infants has provided evidence that the development of imitation (and, by inference, mirror neurons) depends on unspecialised, unconstrained associative learning. Reeb-Sutherland et al. (2012) found that associative learning ability at one-month post-partum, measured using a delay eyeblink conditioning paradigm, predicted performance on a range of imitation tasks at 9 months of age.

**Reflection**

In the last 10 years there has been significant progress in resolving the questions debated in the Mirror Neuron Forum (Gallese et al., 2011). Regarding action understanding: multivoxel pattern analysis, patient studies, and research using TMS now suggests that mirror neuron brain areas contribute to low-level processing of observed actions (e.g. distinguishing types of grip), but not directly to high-level action interpretation (e.g. inferring actors’ intentions). In terms of speech perception, although it remains unclear whether mirror neurons play a specific, causal role in speech perception, there is now compelling evidence for the involvement of the motor system (including premotor mirror neuron brain areas as well as primary motor cortex) in the discrimination of speech in perceptually noisy conditions. In respect of imitation, building on research published before 2011, there is strong evidence from patient, TMS and fMRI studies that mirror neuron brain areas play a causal role in behavioural copying of body movement topography. Finally, concerning autism, studies using behavioural and neurological measures have tried and failed to find evidence for the
broken mirror theory of autism. Instead, there are intriguing signs that, under some conditions, people with autism have stronger mirror responses than neurotypical controls. Alongside these developments, research on the origin of mirror neurons has confirmed the importance of domain-general visual-motor associative learning, rather than canalised visual-motor learning, or motor learning alone. Specifically, major studies assessing “the degree to which neonatal imitation is a reliable phenomenon” (Glenberg, 2011b) have shown that it is not reliable at all.

These findings are disappointing relative to early newspaper headlines. It turns out that mirror neurons are not “Cells that read minds” (Blakesee, 2006); they do not alone explain “what makes humans social” (Scientific American, July 1, 2008), and they have not been able to “do for psychology what DNA did for biology” (Ramachandran, 2009). But it is unlikely that any serious cognitive scientist or neuroscientist would be surprised that mirror neurons have not lived up to these sensational claims. The puzzle is why mirror neurons have so inflamed the popular imagination. We speculate that two factors are important. First, the deep historical pull of atomism. Mirror neurons are small and apparently indivisible; they combine sensory and motor properties in a single unit. From ancient Greece to particle physics, there is a long tradition in which atoms of this kind are understood to be the building blocks of reality. Immersed in this tradition, people may be captivated by the idea that simple, tidy mirror neurons explain the distressing complexities of the social world—including political strife, drug addiction, pornography and responses to media violence (Bocher et al., 2001; Iacoboni, 2008). Second, some descriptions of mirror neurons imply telepathy. If mirror neurons mediate “understanding from within” (Rizzolatti & Sinigaglia, 2010), a “pre-conceptual and pre-linguistic form of understanding,” which can “overcome all linguistic and cultural barriers” (Rizzolatti and Sinigaglia, 2008), they allow an effortless, wordless form of communication that is a lot like telepathy. Given ancient links between mirrors, oracles and divination (e.g. Orofino, 1994), even the name “mirror neurons” may pump the intuition (Dennett, 1984) that these cells give direct, transparent access to other minds (Heyes, 2010).

Figure 1 suggests that, even if public enthusiasm for mirror neurons has been sustained, scientific interest began to decline in 2014. Two high profile reviews were published in that year. One of them, a target article with commentaries in The Behavioral and Brain Sciences, did not contest the existence of mirror neurons, or that they contribute to social behaviour, but marshalled evidence that they are forged by sensorimotor associative learning (Cook et al., 2014). This evidence challenged the view that mirror neurons are a biological adaptation—that they evolved via genetic mechanisms for action understanding, speech, imitation or any other function. The other review, a book entitled The Myth of Mirror Neurons, was more sceptical. It argued that, because mirror neurons are products of associative learning, they could not mediate action understanding or any
other cognitive function (Hickok, 2014). The impact of these two publications should not be overestimated. They distilled and developed insights and concerns that had been emerging over the preceding years. But it is plausible that they precipitated the subsequent decline of mirror neuron research.

Should the decline be resisted, or should it be hoped that the trend apparent in Figure 1 continues until mirror neurons are a thing of the past? We support resistance for two reasons. First, even if the mirror neuron boom was fuelled by less-than-rational currents of thought (e.g. atomism and telepathy), inside as well as outside science, it would not be rational to allow mirror neurons to go bust. This field has produced substantial findings, many summarised in this article, which should not be dismissed just because “Mirror neurons are the most hyped concept in neuroscience” (Jarrett, 2012). Second, in our view, much of the scepticism about mirror neurons is based on a misunderstanding. The discovery that mirror neurons are forged by associative learning does not imply that they are without function. It suggests they are by-products with respect to genetic evolution, but by-products can be very useful indeed - in the oil business and in the brain. Consider the area of the left fusiform gyrus that mediates identification of visual word forms. Literacy emerged late in human history, only 5000-6000 years ago, making it clear that the “Visual Word Form Area” was not designed by genetic evolution to enable reading. It is a by-product of genetic adaptation for the discrimination of visually complex objects – just as mirror neurons are a by-product of genetic adaptation for learning about predictive relationships – but literacy is hugely important as a means of relating to, and learning from, other people; as a social cognitive function (Heyes, 2018).

Our review of mirror neuron research in the last 10 years, and the critiques discussed in this section, have several implications for future research. In relation to mirror neuron functions, research on low-level action discrimination is more likely to make progress than continuing effort to find a direct role for mirror neurons in high-level action interpretation. Similarly, imitation looks more promising than speech perception, and - given the strength of the evidence that mirror neurons contribute to imitation, and the lack of evidence that autism is due to a “broken mirror” – dyspraxias are a more promising target than autism for research with clinical applications. The idea that mirror neurons contribute to imitation was dismissed at an early stage on the grounds that monkeys, the first species in which mirror neurons were identified, cannot imitate (Rizzolatti & Craighero, 2004). This argument assumed that imitation involves copying an entirely novel action, guided by understanding of the model’s intentions; a definition so rich that it implies imitation is rare even in adult humans (Heyes in Gallese et al., 2011). Defining imitation in a way that is more common in cognitive science, as copying the topography of body movement, the associative account
of the origin of mirror neurons implies that any animal capable of sensorimotor associative learning has the potential to develop mirror neurons and to imitate. Humans are better imitators than other animals, including monkeys, because sociocultural experience (e.g. synchronous dance and sporting activities, being imitated by others) provides us with matching sensorimotor experience for a broader range of actions (Cook et al., 2014). In retrospect, it seems that the rich definition of imitation diverted attention from one of the most important functions of mirror neurons.

As for future research on the origins of mirror neurons, there is still much to be discovered about the sources and developmental timing of the sensorimotor experience that builds mirror neurons, how this varies across cultures, and how this information might be used for clinical and educational interventions. More broadly, investigation of how mirror neurons are woven into computational-neurological systems could provide valuable clues about how other cognitive systems are assembled through learning. When a function is “hardwired” or “innate”, the construction work is done by evolution acting on genetic variants and lost in the mists of time, but when the construction is done by learning, it can be studied in creatures alive today (Heyes, 2018).

Turning from targets of study to methods, it is clear that causal methodologies, such as TMS, have more power to illuminate functions than correlational methodologies, and, if it is concluded that the unitary character of mirror neurons is of crucial importance (see below), the use of more single unit recording would be desirable. The potential for such recording would be greatly expanded by the development of rodent models using sensorimotor training (Heyes, 2013).

Thinking more broadly about methodology, research on mirror neurons would benefit greatly from extension of a system-level, computational approach of the type advanced by Hickok and Poeppel (2015) for language, and by Rushworth and colleagues (e.g. Apps et al., 2016) for other aspects of social cognition. Predictive coding is one potential example of a system-level computational approach to investigating mirror neuron contributions to social cognition (Kilner, Friston & Frith, 2007; Press, Heyes & Kilner, 2011). Ideally, in any such approach, each hypothesis about function would specify a part in a psychological process – a process going all the way from peripheral sensory input to overt motor output – that mirror neurons are thought to fulfil, and do this in a way that is testable using behavioural and neurophysiological methods. The name given to this part is not important in itself. What is important is that the name does not derive its meaning purely from folk psychology, and that the hypothetical function of mirror neurons is distinguished clearly from other components and from the overall process. For example, “movement discrimination” or “action recognition” (component) needs to be distinguished from mentalising (whole process). Similarly, on the neurological side, we need more specific, testable theories of how mirror neurons work with other types of neuron, and the kinds of networks in which they are embedded (including
distinguishing whether any differences in function are due to region- or species-specific differences in mirror neurons themselves, or instead due to the properties of the region or organism in which they are located). The term “mirror neuron system” is commonly used, but proponents of this term need to specify the sense in which mirror neurons constitute a “system” rather than just cells with similar properties, found in interconnected areas of the brain.

At the broadest, conceptual level we need to think hard about why, if at all, it is important that single neurons have mirror properties. Setting aside the historical appeal of atomism, does it really matter, in the context of contemporary psychology and neuroscience, whether individual neurons or small networks of neurons match observed and executed actions? Indeed, in a potentially welcome development, recent single-unit recording studies indicate a move away from considering the properties of individual neurons and instead focus on population-level encoding. For example, individual neurons responsive to the observation of others’ manipulative actions (grasping, dragging, etc.) in anterior intraparietal area show viewpoint-dependent coding, but as a population they provide viewpoint-invariant coding of the observed action (Lanzilotto et al., 2020; see also Livi et al., 2019, for a similar demonstration of population encoding of observed actions in pre-supplementary motor area).

In conclusion: It turns out that mirror neurons contribute to complex control systems rather than dominating such systems or acting alone. Their contributions are at a relatively low level – for example, to body movement discrimination rather than intention-reading. And rather than being immutable units from birth, mirror neurons acquire their mirror properties through sensorimotor learning, and change them by the same route. Although disappointing relative to some early claims, we argue that these discoveries should not discourage further research on mirror neurons. The findings reviewed in this article suggest that, when mirror neurons are studied in the context of system-level theory – as having the potential to fulfill a specified part in a complex process – they can help us to understand the categorization of body movements, aspects of speech perception, and the neurological bases of imitation. Furthermore, the evidence that mirror neurons are forged by sensorimotor experience not only raises important questions about the sources of this experience in everyday life across cultures, but opens up the possibility that other neurocognitive mechanisms, once thought to be genetically inherited, are shaped by cultural learning (Heyes, 2018). Mirror neurons should not be tarnished; they are yet to fulfill their true promise.
References


magnetic stimulation and tissue-Doppler ultrasound imaging. Philosophical Transactions of the Royal Society of London - Series B: Biological Sciences, 369(1644), 20130418.


