The ontogenetic origins of mirror neurons: evidence from ‘tool-use’ and ‘audiovisual’ mirror neurons

Since their discovery, mirror neurons—units in the macaque brain that discharge during action observation and execution—have attracted considerable interest. Whether mirror neurons are an innate endowment or acquire their sensorimotor matching properties ontogenetically has been the subject of intense debate. It is widely believed that these units are an innate trait; that we are born with a set of mature mirror neurons because their matching properties conveyed upon our ancestors an evolutionary advantage. However, an alternative view is that mirror neurons acquire their matching properties during ontogeny, through correlated sensorimotor experience. The present article re-examines frequently overlooked neurophysiological reports of ‘tool-use’ and ‘audiovisual’ mirror neurons within the context of this debate. It is argued that these findings represent compelling evidence that mirror neurons are a product of sensorimotor experience, and not an innate endowment.

Keywords: mirror neurons; associative learning; audiovisual mirror neurons; tool-use mirror neurons

1. INTRODUCTION
Mirror neurons (MN) are single units identified in the ventral premotor [1–3] and inferior parietal [4,5] cortices of the macaque brain, which respond to both the observation and execution of transitive and communicative actions. Approximately, 25–30% of the MNs reported are strictly congruent; that is, they respond selectively to the observation and execution of the same action. The remaining MNs (so-called broadly congruent, logically related and non-congruent MNs) respond to similar, related or different actions in observe and execute conditions. Since their discovery in monkeys, considerable indirect evidence has accumulated suggesting that humans also have an MN system [6–8]. Whether MNs are an innate endowment or acquire their properties ontogenetically has been the subject of intense debate [9]. Crucially, while few now doubt that independent sensory and motor experience can fine-tune the response profiles of MNs [9,10], there continues to be considerable disagreement as to how these units acquire their fundamental sensorimotor matching properties [11]. This study contributes to this debate by considering the insights afforded by ‘tool-use’ and ‘audiovisual’ MNs. It is argued that despite being frequently overlooked, the existence and properties of these units provide compelling evidence that MNs acquire their matching properties during ontogeny, as a consequence of correlated sensorimotor experience.

2. THE ORIGENS OF MIRROR NEURONS
Where do MNs come from? One possibility is that MNs are an innate endowment; that we are born with a set of mature MNs because their matching properties conveyed upon our ancestors an evolutionary advantage [3,12–14]. Several authors have argued that early selection pressure favoured MNs because they afforded ‘action understanding’ [3,12]. According to this view, congruent MNs mediate the covert simulation of observed actions; a process that yields first person insights into the intentions and goals of conspecifics [15]. At subsequent stages in primate evolution, MNs may have conveyed further adaptive benefits, including theory of mind [14], imitation learning [12] and language development [13]. Innate MN theory appears to receive some support from reports that neonates ‘imitate’ certain mouth gestures [16–18] (but see Ray & Heyes [19], for an alternative interpretation).

A different view is that MNs acquire their sensorimotor properties ontogenetically, through the same domain-general associative mechanisms that mediate conditioning [11,19,20]. Where visual and motor representations of actions are predictive of one another, the two may become associated. Thereafter, action observation may excite associated motor programmes. Sources of correlated sensorimotor experience likely to promote the emergence of congruent MNs include visual monitoring of one’s own actions either directly or in mirrors; being imitated by others; synchronous activity in response to a common stimulus (e.g. a crowd cheering victory in a sporting arena [19]). Sources of non-matching sensorimotor experience likely to cause the emergence of non-congruent or logically related MNs include coordinated instrumental action (e.g. when an object is passed between interactants, the sight of object-releasing predicts the performance of object-grasping [21]) and control behaviours (the observation of dominant expansive gestures predicts the execution of submissive contractive movements [22]). The associative account is consistent with evidence that neuroimaging, electrophysiological and behavioural markers of the human MN system may be readily modified through correlated sensorimotor experience [23–25].

3. TOOL-USE AND AUDIOVISUAL MIRROR NEURONS
Despite this ongoing debate, direct evidence that macaque MNs acquire their properties through correlated sensorimotor experience exists within the neurophysiological literature, but continues to be frequently overlooked. MNs have been reported in the ventral premotor area F5 of the macaque, which discharge both during observation of actions performed by an experimenter with tools (pliers or a stick) and during manual execution (i.e. performed with the hands) of the same actions by the macaque [26].
Testing was conducted after a two-month training period during which the tools were used to pass food items to the monkeys. According to an associative account, this sort of sensorimotor experience is likely to cause motor representations for grasping food items to become associated with the visual representations of actions made with sticks and pliers, because the former was reliably predicted by the latter. Reports of tool-use MNs, therefore, accord well with the associative account of MN origins, and appear to challenge the view that the sensorimotor matching properties of MNs are an innate endowment [11].

So-called audiovisual MNs have also been identified in the F5 region of the macaque premotor cortex [27,28]. In addition to the sight and execution of actions, these neurons also respond to the sounds associated with the actions. A range of ripping and tearing sounds cause F5 MNs to discharge, including the sound of a peanut breaking, paper ripping, plastic crumpling, metal striking metal and paper shaking. This finding is again entirely consistent with an associative view. Action execution is frequently predictive of both action observation and characteristic 'action sounds'. Repeated exposure to these sensorimotor contingencies will cause the motor representations for ripping and tearing to become associated with both the auditory and visual sensory consequences. Consistent with the reports of tool-use MNs, audiovisual MNs also suggest that the linkage between sensory and motor representations appears to be determined by the correlated sensorimotor experience to which individuals are exposed.

Reports of tool-use and audiovisual MNs appear to argue against the nativist account: evidently, MNs may emerge which respond to seemingly arbitrary stimuli provided they have been paired contingently with the execution of an action. However, ‘mediated activation’ accounts may be advanced to sustain the innate MN hypothesis, if it is assumed that the sight of tool actions, or action sounds, become associated, not with motor programmes directly, but rather with hardwired visual descriptions of hand actions [18,26] or hardwired representations of ‘action goals’ (cf. [29]). The observation of grasping with pliers or the sound of paper tearing might thereby excite motor representations indirectly, via innate representations of grasping or tearing (figure 1), rather than via direct sensorimotor associations.

Nevertheless, while logically plausible, mediated activation accounts cannot explain all of the neuronal responses observed. Crucially, tool-use MNs discharged significantly less often, if at all, to the sight of actions performed with biological effectors, despite robust responses to the sight of the same actions performed with tools [26, p. 214]. Similarly, several audiovisual MNs showed no response to the sight of their effective action alone [28, p. 847], or responded...
more strongly to the sound of actions than to the combined sight and sound of actions [27, p. 633]. These observations are inconsistent with mediated activation accounts, as they imply that the receptive fields of tool-use and audiovisual MNs are tuned to the sensory inputs of tool actions and action sounds, rather than to (i) the sight of actions executed with biological effectors, or (ii) to the ‘goals’ of actions. Mediated activation accounts predict the opposite pattern; that MNs ought to respond maximally to the sight of hand actions executed with biological effectors, indicative of tuning, and weaker responses to any associated sensory inputs. These observations suggest that the sight of tool actions and sensory representations of action sounds excite motor representations directly and not via intermediate hardwired representations.

4. CONCLUSION
Despite being frequently overlooked within the literature, the existence and properties of tool-use and audiovisual MNs argue against the view that the sensorimotor matching properties of MNs are an innate endowment; a product of natural selection [3,12–14]. These reports indicate that the receptive fields of MNs may be tuned to sensory inputs to which the subjects’ ancestors could not possibly have been exposed—e.g. the sight of actions performed with pliers or to the sound of a plastic crumpling. Instead, such findings accord well with the view that all MNs acquire their sensorimotor matching properties ontogenetically, through correlated sensorimotor experience [11,19,20].

To account for the evidence provided by tool-use and audiovisual MNs, nativist MN theory needs to posit that these units are somehow qualitatively distinct from the MNs that could become hardwired through natural selection [18,30]. However, delineating different classes of MNs on the basis of which units accord with a nativist account, and which do not, may be construed as fitting data to theory and not theory to data. Attempts to distinguish audiovisual and tool-use MNs from those units that respond to the observation and execution of actions made with biological effectors appear redundant when an associative framework [11,19,20] offers a single comprehensive account of the existence and properties of all of these sensorimotor units.

I thank Cecilia Heyes, Geoff Bird and Clare Press for useful discussions and comments on an earlier version of this manuscript.

Richard Cook*
Department of Psychology, City University London,
London EC1V 0HB, UK
*Richard.Cook.1@city.ac.uk


Cogn. Brain Res. 3, 131–141. (doi:10.1016/0926-6410(95)00038-0)