Connecting mirror neurons and forward models

R. C. Miall,

University Laboratory of Physiology, Parks Road, Oxford OXI 3PT, UK

Two recent developments in motor neuroscience are promising the extension of theoretical concepts from motor control towards cognitive processes, including human social interactions and understanding the intentions of others. The first of these is the discovery of what are now called mirror neurons, which code for both observed and executed actions. The second is the concept of internal models, and in particular recent proposals that forward and inverse models operate in paired modules. These two ideas will be briefly introduced, and a recent suggestion linking between the two processes of mirroring and modelling will be described which may underlie our abilities for imitating actions, for cooperation between two actors, and possibly for communication via gesture and language. *NeuroReport* 14:000–000 © 2003 Lippincott Williams & Wilkins.

Key words:

INTRODUCTION

You reach out and grasp a piece of apple, you perform some impressive and complex neural processes to transform visual information into an internal plan and into a successfully executed action. If you first see me perform the action, and then imitate it, my action can provide a template for you, but the neurophysiological basis for this is still unclear. However, the mirror neuron system seems to provide an important representation of these actions in the premotor system. Mirror neurons were first found in ventral premotor cortex of the monkey (area F5) [1–3] and have also been reported in posterior parietal cortex (PF, or BA7b) [4,5]. In ventral premotor cortex, many cells code for visually guided actions, responding before or during the monkeys' own actions, but these are not mirror neurons, as they do not code for observed actions. Other cells (canonical neurons) respond during object-related actions and to the sight of the target objects. They tend to be specific for particular hand actions (e.g. precision *vs* whole hand grip) and for sight of the corresponding objects (raisons *vs* bananas). Hence they seem to code the affordance of an object; but these cells also are not mirror neurons. Mirror neurons uniquely show similar responses during execution of a movement and during the animal's observation of the same action performed by another monkey or by the human experimenter. They do not respond to the object alone. Furthermore, mirror neurons can fire during a reach to an object placed out of sight, as long as the intention of the reach and grasp action is clear [6]. Thus they are not driven simply by visual input, but are a representation of an object-directed action. It is still unclear what the typical response characteristic of these mirror neurons is, in part because they seem to respond to the sight (and sound [7,8]) of real actions, and are not easily driven by televised images of actions. Thus careful parametric analysis of their receptive characteristics is difficult. However, two key points are, first, that while their motor responses can be quite action-specific, they tend to respond to a broader class of observed actions (a grasp by either a human or a monkey, but not by a robot or by a human using pliers [3]). Second, their responses are object-related movements, typically towards food (the reach for an apple piece, but not for a metal screw, for example; for an excellent review see [5]). Mirror neurons in areas PF of the posterior parietal cortex (PPC) are very similar, but are thought to code more specifically for the kinaesthetic and somatosensory components of an action [9,10]. This may have relevance to neurons in area BA5 of the PPC that code kinematics and not dynamics [11]. The last group of mirror-like cells to mention are in the superior temporal sulcus, STS, where cells respond to biological motion, to view of body parts, especially faces and hands, and where some respond specifically to observation of reach and grasp movements [12,13]. Thus they have rather similar visual response properties to mirror neurons [5], but these STS cells do not fire during execution of the unseen action.

Mirror neurons are, therefore, cells with extraordinarily complex response characteristics, closely linked to visual observation of goal-directed actions. The excitement of their discovery is compounded by the detection of functional brain imaging activation patterns that suggest similar neurons are found in human ventral premotor and parietal cortices [9,10,14], and by the fact that the language area, Broca's area, in human ventral premotor cortex is the homologue of the monkey's area F5. Mirror neurons could therefore be an important neural stepping stone between coding for observed goal-directed actions, to understanding gestures, to spoken language [5,15].

Turning now to the concept of internal models, it is currently thought that the motor system uses two forms [16]. Inverse models map the relationship between intended actions or goals and the motor commands to reach those goals. The term inverse is due to the inverse relationship between the properties of the motor system and the properties of the internal model. Any form of accurate

Susan Ashok/ WNR lww.wnr.1121 1–3

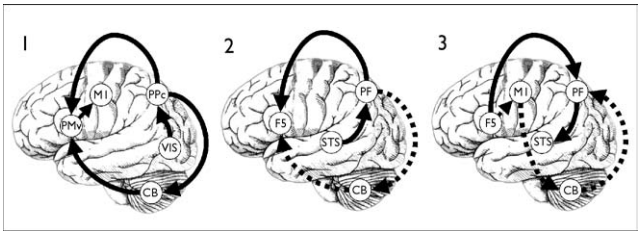

Fig. 1. (**1**) A schematic diagram of pathways activated during a visually guided reach: information from visual areas is processed by the posterior parietal cortex, and feeds through to the premotor and motor cortex by cortico-cortical and cortico-cerebellar routes. (**2**) Activity during observation of actions. The circuit linking STS, PF and F5 (solid arrows) may act as an inverse model [9,19]. However, the cerebellum could also perform this function (dashed arrows). (**3**) Activity during the execution of imitated actions. The circuit linking F5, PF and STS (solid arrows) may act as a forward model, to generate a prediction of movement outcome [9,19]. Again, the cerebellum could perform this function (dashed arrows).

pre-programming implicitly invokes this inverse relationship, but the motor system probably develops an explicit inverse model. When you reach for the piece of apple, the inverse model would transform a sensory representation of the relative positions of apple and arm into motor commands to perform the action (Fig. 1).

Forward models map the relationship between motor commands and the resultant change in the state of the motor system, which is monitored by reafferent sensory inflow. Hence a forward model estimates the next (sensory) state of the motor system based upon information about its current state, its dynamics, and the current motor commands being issued to it. Forward models can support sensorimotor control in many ways [17]. Their estimates of the state of the system can supplement noisy and incomplete sensory signals; they can be used in advance of delayed reafference from the periphery; they can cancel out reafference during movement; they can generate appropriate error signals for modifying future control; and they can help identify the current context within which control is being attempted. It is this last function that was recently used in a computational model, MOSAIC [18], whereby multiple forward models predicted in parallel changes in motor state. The accuracy of each forward model prediction to the actual state allowed the current sensory-motor context to be estimated, as only the right model for a given context would accurately predict movement outcome. So if your lift of the apple matches your prediction, you have confidence that the context (lifting apple) was correct, and you have not mistakenly grabbed the table instead. Moreover, in MOSAIC, this matching process also leads to a weighting of the output of the inverse model paired to each forward model, such that the right inverse model is selected to control the motor system in any particular context.

So how do mirror neurons and forwards models bear on each other? Marco Iacoboni has suggested [9,19] that the F5 mirror cells lie at a crucial interface between forward and inverse models, which are instantiated by cells in STS, PF and F5. STS is responsible for the visual representation of observed actions, and Iacoboni suggests that the connections from STS to PF and onwards to the mirror cells in F5 form an inverse model, converting this visual representation into a motor plan (solid arrows, Fig. 1, part 2). The reverse connection from mirror cells in F5 to PF, and back to STS (solid arrows, Fig. 1, part 3), then forms a forward model [20], converting the motor plan back into a predicted visual representation (a sensory outcome of action). Thus these two streams could underpin imitation, in which actions are first observed, then transformed by the inverse model into potential motor commands, and the visual consequence of these movements are then predicted by the forward model, for comparison with the exemplar visual images [9,10]. This two-way process could also allow an observer to track another's hand actions with predictive eye movements, the gaze shifts anticipating the other's hand motion with the same advance as seen when tracking one's own actions [21]. It could underpin imitative learning, as demonstrated by robotic studies [22,23]. It could allow cooperative actions such as shaking hands or dancing, or it could allow successful competitive actions, such as when we both attempt to grab the same bit of apple.

This elegant scheme linking STS, PF and F5 has some unresolved issues. First, it ignores the cerebellum, strongly favoured to act as an internal model, forward or inverse or both [24,25]. The PPC has strong projections to the cerebellum [26,27], which then projects to the ventral premotor cortex [28], among other areas. This may be the major route for visuo-motor information to reach the premotor cortex (Fig. 1) and contribute to the evolving motor command [27]. Hence, the mirror cells in premotor cortex may code a motoric representation of visuo-motor actions, both during action execution and during observation, driven by the cerebellar inverse model. The output of the motor cortical areas also projects to the cerebellum, as cerebro-pontine fibres from premotor and from primary motor cortex. The dentate nucleus projects, via thalamus, back to the PPC [29,30]. Consequently, a forward model in the cerebellum could use an efferent copy of motor commands to provide a visuo-motor update to the motor representation held in PPC. So both inverse and forward processing pathways might depend on the cerebellum (dashed arrows, Fig. 1 parts 2 and 3) [24,25]. The critical interface in this version of the scheme is the posterior parietal cortex rather than ventral premotor cortex. The PPC codes for the spatial location of a visual goal for action [2], and this would provide appropriate input to the inverse model process. The PPC also codes the spatial relationships defining the current state of the body. It appears to optimally combine multi-modal visual, auditory and proprioceptive sensory afferents [31,32] with predictive updates during actions, the latter based on efference copy of the motor commands [33,34], appropriate from the forward model process.

Second, there seems to be an important gap between the specificity of forward models and the generality of mirror neurons. Forward model predictions need to be rather precise if they are to provide signals on which to base estimates of movement outcome and of current context [17]. For imitation, however, a very close link between observed and executed actions is often inappropriate: I may see you perform the apple-grasping action with your left hand, but execute it with my right hand, if it seems that your intention was a goal directed reach to the target object. Thus some deeper coding of the intention behind the observed action

has been invoked [35]. Wolpert *et al.* [36] have recently suggested a hyper-MOSAIC extension of their model to incorporate these deeper levels. The lowest level (the original MOSAIC model [18]) has paired forward/inverse models that are action specific, estimating and controlling movement to reach specific motor states: the grasp and lift of the apple. The next layer receives from the lowest level and forms a representation that is somewhat independent of specific movement trajectories or states. Higher levels receiving from this level encode actions independent of specific effectors, and might then code for a reach to a target (the apple) with either hand. The very highest layers then code for behavioural goals, at some abstract symbolic level independent of effectors or movements, to satisfy hunger, for example. Quite how this multi-level controller could be generated neurally is not yet clear, but the link with mirror neurons seems appealing. Mirror neurons would appear to code for actions at something like the intermediate levels within HMOSAIC. But a final important question is why mirror neurons are specific for movements towards objects: the internal model theories outlined here have no clear answer.

REFERENCES

1. Di Pellegrino G, Fadiga L, Fogassi L *et al. Exp Brain Res* **91**, 176–180 (1992).
2. Gallese V, Fadiga L, Fogassi L *et al. Brain* **119**, 593–609 (1996).
3. Rizzolatti G, Fadiga L, Gallese V *et al. Brain Res Cogn Brain Res* **3**, 131–141 (1996).
4. Gallese V, Fogassi L, Fadiga L *et al.* Action representation and the inferior parietal lobule. In: Prinz W and Hommel B (eds). *Attention and Performance*, Vol. XIX**. Oxford: Oxford University Press; 2002, pp. 247–266.**
5. Rizzolatti G, Fogassi L and Gallese V. *Nature Rev Neurosci* **2**, 661–670 (2001).
6. Umilta MA, Kohler E, Gallese V *et al. Neuron* **31**, 155–165 (2001).
7. Keysers C, Kohler E, Umilta C *et al. Exp Brain Res* DOI: 10.1007/s00221-003-1603-5, (2003).
8. Kohler E, Keysers C, Umilta MA *et al. Science* **297**, 846–848 (2002).
9. Carr L, Iacoboni M, Dubeau MC *et al. Proc Natl Acad Sci USA* **100**, 5497–5502 (2003).
10. Iacoboni M, Woods RP, Brass M *et al. Science* **286**, 2526–2528 (1999).
11. Kalaska JF, Cohen DAD, Prud'homme M *et al. Exp Brain Res* **80**, 351–364 (1990).
12. Perrett DI, Harries MH, Bevan R *et al. J Exp Biol* **146**, 87–113 (1989).
13. Wachsmuth E, Oram MW, and Perrett DI. *Cerebr Cortex* **4**, 509–522 (1994).
14. Buccino G, Binkofski F, Fink GR *et al. Eur J Neurosci* **13**, 400–404 (2001).
15. Rizzolatti G and Arbib MA. *Trends Neurosci* **21**, 188–194 (1998).
16. Wolpert DM and Ghahramani Z. *Nature Neurosci* **3** (Suppl), 1212–1217 (2000).
17. Miall RC and Wolpert DM. *Neural Netw* **9**, 1265–1279 (1996).
18. Haruno M, Wolpert DM and Kawato M. *Neural Comp* **13**, 2201–2220 (2001).
19. Iacoboni M. Understanding others: Imitation, language, empathy. In: Hurley S and Chater N (eds). *Perspectives on Imitation: From Cognitive Neuroscience to Social Science, Volume 1: Mechanisms of Imitation and Imitation in Animals*. Cambridge, MA: MIT Press; 2003: in press.
20. Iacoboni M, Koski LM, Brass M *et al. Proc Natl Acad Sci USA* **98**, 13995–13999 (2001).
21. Flanagan JR and Johansson RS. *Nature* **424**, 769–771 (2003).
22. Demiris J and Hayes G. Imitation as a dual-route process featuring predictive and learning components: a biologically plausible computational model. In: Dautenhahn K and Nehaniv CL (eds). *Imitation in Animals and Artifacts*. Cambridge, MA: MIT Press; 2002, pp. 327–361.
23. Schaal S, Ijspeert A and Billard A. *Phil Trans R Soc Lond B Biol Sci* **358**, 537–547 (2003).
24. Miall RC, Weir DJ, Wolpert DM *et al. J Motor Behav* **25**, 203–216 (1993).
25. Wolpert DM, Miall RC and Kawato M. *Trends Cogn Sci* **2**, 338–347 (1998).
26. Schmahmann JD and Pandya D. *J Comp Neurol* **289**, 53–73 (1989).
27. Stein JF and Glickstein M. *Physiol Rev* **72**, 967–1017 (1992).
28. Dum RP and Strick PL. *J Neurophysiol* **89**, 634–639 (2003).
29. Amino Y, Kyuhou S, Matsuzaki R *et al. Neurosci Lett* **309**, 29–32 (2001).
30. Clower DM, West RA, Lynch JC *et al. J Neurosci* **21**, 6283–6291 (2001).
31. Andersen RA, Snyder LH, Bradley DC *et al. Annu Rev Neurosci* **20**, 303–30 (1997).
32. Pouget A, Deneve S and Duhamel JR. *Nature Rev Neurosci* **3**, 741–747 (2002).
33. Duhamel JR, Colby CL and Goldberg ME. *Science* **255**, 90–92 (1992).
34. Wolpert DM, Goodbody SJ and Husain M. *Nature Neurosci* **1**, 529–533 (1998).
35. Wohlschlager A, Gattis M and Bekkering H. *Phil Trans R Soc Lond B Biol Sci* **358**, 501–515 (2003).
36. Wolpert DM, Doya K and Kawato M. *Phil Trans R Soc Lond B Biol Sci* **358**, 593–602 (2003).

Acknowledgements: This work was supported by the Wellcome Trust and the McDonnell Foundation. I am very grateful to M. Iacoboni for his comments.

DOI: 10.1097/01.wnr.0000098751.87269.77