ARTICLE IN PRESS

Consciousness and Cognition xxx (2012) xxx–xxx

Contents lists available at SciVerse ScienceDirect

Consciousness and Cognition

journal homepage: www.elsevier.com/locate/concog

More questions for mirror neurons

Emma Borg

Philosophy Department, University of Reading, Reading RG6 6AA, UK

ARTICLE INFO

Article history:
Available online xxxx

Keywords:
Mirror neurons
Mindreading
Intentions
Intentionality

ABSTRACT

The mirror neuron system is widely held to provide direct access to the motor goals of others. This paper critically investigates this idea, focusing on the so-called 'intentional worry'. I explore two answers to the intentional worry: first that the worry is premised on too limited an understanding of mirror neuron behaviour (Sections 2 and 3), second that the appeal made to mirror neurons can be refined in such a way as to avoid the worry (Section 4). I argue that the first response requires an account of the mechanism by which small-scale gestures are supposedly mapped to larger chains of actions but that none of the extant accounts of this mechanism are plausible. Section 4 then briefly examines refinements of the mirror neuron-mindreading hypothesis which avoid the intentional worry. I conclude that these refinements may well be plausible but that they undermine many of the claims standardly made for mirror neurons.

1. Introduction

As is by now very well-known, both the monkey brain and (apparently) the human brain contain a special category of neurons which respond in two distinct conditions – neurons which fire both when a subject performs a given motor action and when a subject witnesses the same motor act being performed by a target.[1] So, for instance, one and the same pattern of neuronal stimulation can be observed both when a subject undertakes a precision grip of a cup and when they merely see a target performing that action. These dual-acting neurons – the so-called 'mirror neurons' (MNs) – have given rise to a great deal of interest and a fair amount of speculation since being first reported by Gallese, Fadiga, Fogassi, and Rizzolatti (1996). Much of this interest has stemmed from the idea that mirror neurons might play a fundamental role in social cognition, providing the neurological basis for our understanding of the minds of others.[2] The suggestion has been (e.g. see Gallese & Goldman, 1998) that if self-generated MN activity (i.e. the firing of an agent's MNs during the performance of an action by the agent) forms the neurological basis of the subject's plan to carry out that action, then externally-generated MN activity (i.e. the firing of an agent's MNs during observation of the same action performed by another) might still form the basis of plan formation, but in the latter case the plan is in some way taken 'off-line', not resulting in motor-action by the agent but instead being somehow 'tagged' as belonging to the target. Such externally-generated MN activity would then provide direct access to the action plans (or motor goals) of others, providing at least a first-step on the road to full social cognition. In what follows, I will term this idea the 'MN-mindreading hypothesis'.

As Hickok (2008) points out, one immediate worry for the MN-mindreading hypothesis is that terms like 'goal' and 'action' are ambiguous and can be read in either non-intentional or intentional ways (e.g. 'the goal' of an action might, on the one hand, be simply the object to which it is directed or, on a more intentional reading, it might be what the action

E-mail address: e.g.n.borg@reading.ac.uk

[1] See Chong, Cunnington, Williams, Kanwisher, and Mattingley (2008) and Dinstein (2008) for some questions about the evidence for the existence of mirror neurons in humans (as opposed to monkeys). What is involved in 'witnessing a target perform the same motor act' will be explored below.

[2] For instance, see the *NY Times* article on 10th January 2006 entitled 'Cells that read minds'.

1053-8100/$ - see front matter © 2012 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.concog.2012.11.013

is designed to bring about).[3] Clearly, if MN activity is to provide a route to an intentional explanation of the behaviour of others, descriptions of MN activity must either deploy notions like 'goal' in non-intentional ways or proponents of the hypothesis must explain how intentional readings are made possible. However MN enthusiasts are confident that such an explanation can be given and that MNs can provide the bridge between physical actions and intentional attributions. So, for instance, we find Ramachandran (2000) suggesting that MNs provided "the driving force behind the 'great leap forward' in human evolution" and that their discovery will 'do for mindreading what the discovery of DNA did for biology' (2006). Other authors have been rather more circumspect, but still we find Iacoboni et al. (2005) arguing that MNs should be understood not merely as capturing the *what* of an action but its *why*, i.e. the intention with which a given motor act was performed, and it now seems commonplace to hold (as, e.g., do Buccino et al., 2007: T119) that "the MNS [mirror neuron system] appears [also] to underlie the coding of intentions behind the actions of others". Furthermore, the scope of issues to which MNs may contribute an answer is often held to range widely, as for instance in Oberman et al. (2005): 190–1 "Mirror neurons are primarily thought to be involved in perception and comprehension of motor actions, but they may also play a critical role in higher order cognitive processes such as imitation, theory of mind, language, and empathy".[4]

In Borg (2007) I raised some questions for the MN-mindreading hypothesis and expressed some scepticism about the claim that MNs can provide the key to intentional understanding (see also Csibra (2007) and Jacob (2008a, 2008b) for sceptical stances in this area). The aim of this paper is to revisit the argument against the MN-mindreading hypothesis given in Borg, 2007, exploring further some of the possible solutions to it and assessing whether scepticism about the MN-mindreading hypothesis is still warranted. My conclusion will be that it is.

The structure of the paper is as follows: in Section 2 I will re-introduce the challenge I raised in Borg (2007) (here labelled 'the intentional worry') and explore the responses that can be made to it. As we will see, there are, I think, two main lines of response to the initial worry: the first (introduced in Section 2 and explored in Section 3) is to claim that it rests on a misunderstanding of MNs (a route pursued, I take it, by Sinigaglia (2008)). The second, more concessive response, is to claim that it rests on a misunderstanding of the MN-mindreading hypothesis. This kind of refinement of the MN-mindreading hypothesis can be found in recent work by both Gallese and Goldman and the move will be examined further in Section 4.[5] The conclusion of Section 3 will be that the problem does remain for strong accounts which take MNs as the key to at least some attributions of intentional mental states. Furthermore, while the weaker reading of the role of MNs to be examined in Section 4 may yet be vindicated, I will argue that it does alter the MN-mindreading hypothesis in certain crucial ways. Thus, I will contend, at this stage of the debate, it remains completely unclear whether or not MNs do provide a critical first-step on the path to social comprehension.

2. Mirror neurons and the intentional worry

The intentional worry for the claim that MNs provide the key to attributing (at least some) mental states to others is, at its simplest, that MN activity is just too closely tied to behaviour to make it (at least unaided) a route to the attribution of intentional mental states. If MNs respond to gestures and simple motor acts (e.g. firing differentially in response to a precision grip versus a whole-hand pick up) it would seem that they operate at the wrong level to underpin intentional attribution, since, on the one hand, a single gesture can map to a multitude of intentions, while, on the other, multiple gestures can map to a single intention. So, for instance, I can grasp a cup because I want to drink from it or because I want to examine it. Yet if MN activity responds merely to the grasping act itself it will underdetermine the intention. Or again, a range of different specific motor acts can all be realisations of a single intention (e.g. an intention to have a drink being realised through grasping a cup with a precision grip or grasping a cup with the whole hand, etc.). So, again, if MNs are sensitive just to brute kinematics they will underdetermine the intentions which accompany our gestures.

The most robust response to this worry (see, e.g., Sinigaglia's thought-provoking 2008 paper) is to reject it entirely, arguing that it is premised on far too simplistic an understanding of MN activity – given a proper understanding of how MNs behave and their sensitivity to the motor goals of the target, we will see that the intentional worry simply dissolves away. To get a proper understanding of the behaviour of MNs we need to appreciate that, while *some* MNs do respond simply to brute kinematics, others seem sensitive not to mere bodily movements but rather to the sequence of actions within which a given motor act is embedded and thus to the goal which a sequence of actions is designed to realise. So, for instance, the MN sequence for the grasp of an object can be triggered by seeing a conspecific reach behind a screen, so long as the observing subject knows that there is a graspable object behind the screen (so grasping MNs fire even when the brute kinematics of the grasp are unobserved). Furthermore, as Fogassi et al. (2005) demonstrated, one and the same precision grip of a piece of food results in different patterns of MN stimulation in macaque monkeys depending on whether the movement is embedded in a larger act of grasping food to eat or grasping food to place in a container (notably even if the container is located near the monkey's mouth, thus making the brute kinematics of the two action sequences extremely similar). On the other hand, very

[3] See also Goldman (2009): 239: "The Parma team often stress that motor MNs code for the goals of observed actions. This may be uncontroversial if 'goal' means 'goal object', which might refer either to a physical object (e.g. a cup) or a physical event or outcome (e.g. a cup being moved to one's mouth). But the same term 'goal' can be used to refer to a mental state, a state of desire with a certain intentional object or relation of 'aboutness'".

[4] See also Arbib, Oztop, and Zuckow-Goldring (2005) and elsewhere for claims relating MNs and language understanding.

[5] Though we should also note that in their original Gallese and Goldman (1998), the authors were careful (and perhaps more careful than some of those who came after them) to stress both the modesty and the potentially tendentious nature of their claims about MNs.

Please cite this article in press as: Borg, E. More questions for mirror neurons. *Consciousness and Cognition* (2012), http://dx.doi.org/10.1016/j.concog.2012.11.013