

## ESSAY



D. PARKINS

## Behind the looking-glass

To understand how mirror neurons help to interpret actions, we must delve into the networks in which these cells sit, say **Antonio Damasio** and **Kasper Meyer**.

A remarkable discovery was made more than ten years ago: some neurons in the brains of macaques are active both when a monkey moves and when it sees a person move in a comparable way<sup>1</sup>. The lead researchers, Giacomo Rizzolatti and Vittorio Gallese, called the cells involved 'mirror neurons'. This evocative name, and the significant implications of the finding, led to a surge of scientific and public interest in these cells. But perhaps the name was too evocative for the finding's own good. It seems to have tempted people into thinking of these neurons as tiny, miraculous mirrors that allow us to understand each other, diverting attention from the search for how they work.

After mirror neurons were found in two main regions of macaque brains<sup>1,2</sup>, corresponding results were found in human brain scans<sup>3</sup>. Researchers began to invoke mirror neurons to explain the fundamental task of

how humans and other primates recognize the actions and, by extension, the intentions of others. Mirror-neuron activity is thought to generate a more-or-less explicit simulation of others' actions in the observer's brain. If this is the case, then simulation, rather than inference, would provide a deep understanding of those movements. Some have proposed that mirror neurons are a vital part of how

infants first learn to imitate other people's actions; of why seeing a grimace can trigger empathetic feeling in a watcher; and even of how language is learned. Although simulation is unlikely to account entirely

for how people understand what they see other people do, the identification of mirror neurons was an important step on the path to unpacking that process.

Unfortunately, the lack of a satisfactory explanation for how the simulation process occurs has allowed the notion to take hold

— among the public and scientists alike — that mirror neurons accomplish mirroring on their own in some mysterious way.

Surely, if 'action understanding' is to occur by internal simulation, the process must enlist both motor and sensory systems in the brain. Mirror neurons cannot accomplish this task alone. Rather, they must be acted on by other structures; and, just as importantly but a fact generally neglected, mirror neurons must act on other structures.

### Mutual support

Unravelling how mirror neurons work requires knowledge of the complex architecture in which these cells are embedded. A model of neural architecture proposed by one of us (A.D.)<sup>4</sup> nearly 20 years ago may help with that. Impressive findings from mirror-neuron research at the single-cell level have added persuasive empirical support for this early model.

The model, originally called 'time-locked multimodal retroactivation', was derived

**"Hearing a peanut being broken without seeing or feeling it triggers a whole neural network."**

during the 1980s from observations in patients with brain lesions. Some patients with damage in the anterior (higher-order) sectors of the temporal cortex, for example, could not recall complex memories that combine separate components of a specific event, such as a face, a name, an action or a place. These patients could not recognize close friends or relatives, nor could they picture unique events in time and place, such as their own wedding or the birth of a child. Yet they could easily imagine or recognize a picture of a non-specific wedding or a baby.

In other words, anterior damage did not preclude their retrieval of mental representations of objects, places or actions; but it did stop them recalling certain combinations of representations that signified a particular person or event. Only damage to posterior sectors of the cerebral cortex, near sensory and motor cortices, impaired access to separable memory components. Given what was already known about memory systems, this suggested that anterior sites held the key to some process needed to reconstruct, elsewhere, the parts that made up a complex memory.

### Mind mapping

An idea for a brain architecture that could explain these findings was inspired by experimental neuroanatomy studies showing that, surprisingly, signals are conveyed within the brain in both forwards and backwards directions. For example, signals are obviously sent from the eye to the visual cortex and on to areas of higher-level processing in the brain. But these high-level areas also send signals back to the visual cortex, and even to the visual thalamus, below the level of the cortex<sup>5</sup>. The same forward-backward signalling arrangement was found for the hippocampus — a brain region involved in making factual memories<sup>6</sup>.

All this led to a model that posited the existence of 'convergence-divergence zones' (CDZs). These neural ensembles collect signals from separate sites, and signal back to those sites. When several signals converge on a CDZ, the ensemble creates an abstract record of the coincident activations — a memory trace, in other words. The model contained two broad types of CDZ. 'Local CDZs' were proposed to coordinate information within regions close to a sensory cortex, such as the visual cortex. These local hubs were proposed to converge on 'non-local CDZs' in higher-order sectors of the brain.

In this view, when a monkey breaks open a peanut<sup>7</sup>, local CDZs collect information about various sensory inputs, and feed these to a non-local CDZ that records the coincident

information about the sound, sight and feel of this action. The CDZ does not hold all the details of this information; rather, it contains the potential to retroactivate the separate auditory, visual, tactile and motor sites, and thus reconstitute the original distributed set of memories and information. Imagine the CDZ as a repository of instructions for a book that must be printed by several different printers. Having the instructions alone will not get you the book, and the printers alone will not help either. You need both to get the final product.

In future, hearing a peanut being broken without seeing or feeling it triggers a series of events. First, it activates the auditory cortices and local auditory CDZ; second, it activates the non-local CDZ that previously collected the memory trace associated with this sound; third, it precipitates simultaneous signalling outwards from this non-local CDZ to all the local CDZs involved in the original event (motor, visual, auditory); fourth, it reactivates all or some of these sites. This leads to a more-or-less successful replay of the coincident set of separate brain activities that accompanied the monkey breaking open a peanut.

Looked at in this way, mirror neurons correspond to non-local CDZs. Their connections to other CDZs, and their ability to collect and distribute signals based on learned experience, allow the brain to reconstruct an action from only part of the story. A whole neural network underlies the understanding of action, rather than a single anatomical site or even a single cell. The monkey's comprehension of the sound of a cracking nut is not created just by mirror-neuron sites, but also by the nearly simultaneous triggering of widespread memories throughout the brain.

The neurons at the heart of this process, and at the heart of non-local CDZs, are not so much like mirrors, after all. They are more like puppet masters, pulling the strings of various memories.

Recent findings are in line with this view. Studies in humans and monkeys show that the neural network stimulated by watching an action goes beyond the original mirror-neuron sites; it encompasses more widespread sensorimotor cortices<sup>3,8</sup>. Conversely, carrying out an action recruits sensory cortical areas even when subjects can neither see nor hear the actions they perform<sup>9</sup>. This lends further support to the notion that the neural description of an action goes far beyond its

motor components. At least one other study has invoked the necessity of convergent signals into mirror-neuron areas to explain such observations<sup>10</sup>. The CDZ model, and our interpretation of mirror neurons, adds the aspect of divergence.

### Neural networks

The CDZ framework allows us to see the role of mirror neurons more clearly. Cells in mirror-neuron areas do not themselves hold meaning, and they alone cannot carry out the internal simulation of an action. This runs counter to the understandable public misconception that mirror neurons alone 'mirror' action. Rather, mirror neurons induce widespread neural activity based on learned patterns of connectivity; these patterns generate internal simulation and establish the meaning of actions. Mirror neurons pull the strings, but the puppet itself is made of a large brain network.

The CDZ model was well received, but the lack of single-cell experiments providing direct evidence for this functional architecture has limited the application of the idea. The mirror-neuron evidence sits well alongside the model, and each seems to make more sense in light of the other. But this is not proof. The ultimate test of the convergence-divergence model, and its explanation of how mirror neurons do what they do, depends on the ability to record brain activity simultaneously from neurons in separate sites, and on probing the underlying connectivity between neural areas. Such goals are within reach, albeit technically difficult to achieve. In the meantime, bringing together mirror-neuron data and the CDZ model could guide future efforts to explain the relationship between what we see and what we do. ■

**Antonio Damasio and Kaspar Meyer are neuroscientists at the Brain and Creativity Institute of the University of Southern California, Los Angeles, USA.**

1. Gallese, V., Fadiga, L., Fogassi, L. & Rizzolatti, G. *Brain* **119**, 593-609 (1996).
2. Fogassi, L., Gallese, V., Fadiga, L. & Rizzolatti, G. *Soc. Neurosci.* **24**, abstr. 257.5 (1998).
3. Buccino, G. et al. *Eur. J. Neurosci.* **13**, 400-404 (2001).
4. Damasio, A. R. *Cognition* **33**, 25-62 (1989).
5. Rockland, K. S. & Pandya, D. N. *Brain Res.* **212**, 249-270 (1981).
6. Van Hoesen, G. W. *Trends Neurosci.* **5**, 345-350 (1982).
7. Kohler, E. et al. *Science* **297**, 846-848 (2002).
8. Raos, V., Evangelidou, M. N. & Savaki, H. E. *J. Neurosci.* **27**, 12675-12683 (2007).
9. Gazzola, V., Aziz-Zadeh, L. & Keysers, C. *Curr. Biol.* **16**, 1824-1829 (2006).
10. Keysers, C. & Perrett, D. I. *Trends Cogn. Sci.* **8**, 501-507 (2004).

**"The neurons are not so much like mirrors, after all. They are more like puppet masters, pulling the strings of various memories."**