New Scientist

The ‘me’ illusion: How your brain conjures up your sense of self

Self-awareness isn't the pinnacle of consciousness - it's just an accidental byproduct of evolution, and a figment of our minds

Reilika Landen/plainpicture

By **Sofia Deleniv**

LOOK into a mirror and you may see pimples, wrinkles or unruly facial hair, but beneath the superficial lies something far more interesting. Every time you lock eyes with your reflection, you know exactly who is looking back at you. The sense of self is unmistakable. It is so much a part of being human that we often fail to notice it. Yet self-awareness is one of the biggest mysteries of the mind. How did it arise and what is it for?

Looking at other animals suggests we are not alone in being able to recognise ourselves in a mirror. Admittedly, it’s a short list of species that seem capable of this feat, but it hints at a possible explanation. Self-awareness may have evolved in only the brightest animals with the biggest brains. If so, it represents the peak of mental complexity – the highest form of

However, some people have started to question this idea. Now, an extraordinary finding lends weight to their scepticism: one monkey species that was previously deemed unable to recognise itself in a mirror can easily learn to do so. This isn't simply another name to add to the echelons of the self-aware. The discovery suggests we need to fundamentally rethink our ideas about mirrors and minds.

The hunt for self-awareness among non-humans has been going on for decades. In the most widely used test – the so-called face-mark test – researchers stealthily apply a spot of odourless dye to an animal's forehead or cheek and then observe its reaction when it is in front of a mirror. The underlying premise is that those with a firm sense of self can acknowledge their reflection and attempt to scrub off the dye.

Most of the animals that have passed this test are considered to be intelligent. They include chimps, bonobos, orangutans, Asian elephants and Eurasian magpies (a member of the notoriously clever corvid family). Killer whales and bottlenose dolphins also seem to recognise themselves in a mirror, although their anatomy means they can't remove a face mark. This apparent correlation with smarts means that self-awareness has become a sort of proxy for mental complexity. But there are some puzzling evolutionary gaps. Gorillas, for instance, usually fail the test – with the notable exception of the recently deceased Koko – yet our more distant primate relatives, the orangutans, pass it. Also, the self-aware elite contains some bizarre anomalies such as pigeons, manta rays, ants and even a robot.

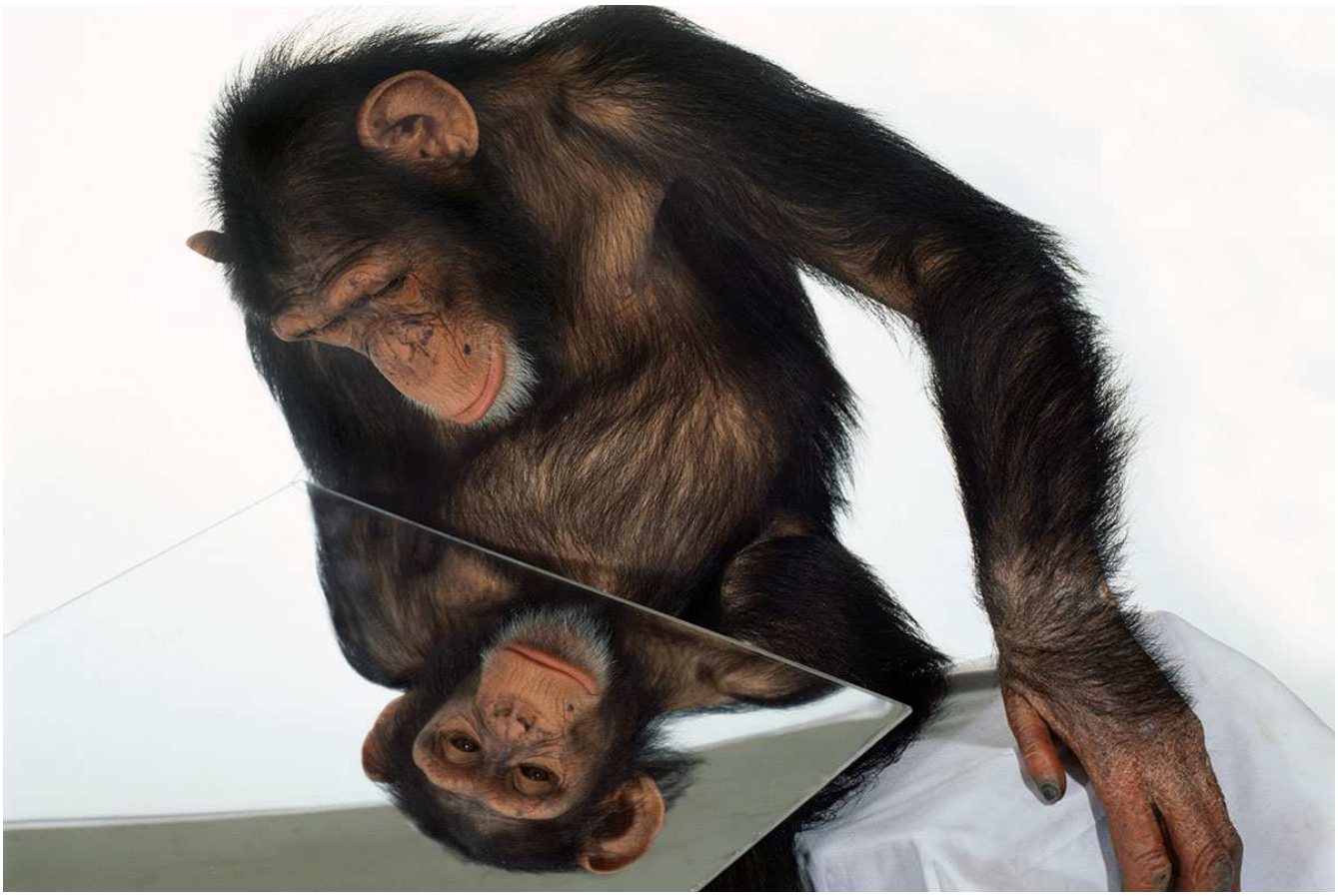
Some of these findings – particularly with ants and pigeons – are contested. Researchers have tried to explain away others, arguing, for example, that gorillas have mentally regressed since their split from the other ape lineages because they face fewer pressures in their environment. But the recent discovery in monkeys is harder to dismiss.

Last year, Liangtang Chang and colleagues at the Shanghai Institutes for Biological Sciences, China, released video footage of a small group of rhesus macaques interacting with a mirror. It shows the monkeys contorting their bodies, tugging at their facial hair, inspecting their fingertips and making flashy displays of their genitals, all the while keeping their eyes on their reflections. They are captivated, leaving little doubt they recognise themselves. Yet, rhesus macaques have consistently failed the mirror test. And just a few weeks earlier, the ones studied by Chang's team had shown no signs that they understood their reflections. What changed?

In fact, there is anecdotal evidence of macaques in the lab showing a sudden interest in mirrors after being fitted with bulky neural recording devices that protrude from their heads. Chang's team wondered whether the monkeys genuinely lacked self-awareness, or whether they were being held back by a lack of coordination – an inability to link what they saw with internal signals generated by their muscle movements. To test this, they taught the monkeys to link vision and movement by giving them a food reward for touching a projected laser dot. At first, the researchers shone the laser where the monkeys could easily see it, then gradually worked up to shining it in places only visible in the mirror. Fast-forward a few weeks of practice, and they passed the face-mark test with flying colours.

At the least, this indicates that the way we test for self-awareness is flawed (see "Mirror, mirror"). That, in turn, raises the possibility that self-awareness is much more widespread than we think. So, what do we know about the evolution of this prized trait?

Many psychologists and anthropologists hold that there is a hierarchy of consciousness that corresponds with increasing brain complexity. At its base is the minimal consciousness attributed to animals with simple nervous systems. These minds are thought to be permanently adrift in a sea of raw sensory experiences, tossed around between perceptions such as colour, hunger, warmth and fear, with little awareness of their meaning. Few minds are sophisticated enough to experience the world differently – through an introspective lens. Even then, they may have a limited sense of self. Only at the peak of mental complexity do we find minds able to construct a lifelong narrative of experiences centred around an abstract concept of "self" – these are the elite.



Smart animals like chimps and dolphins can recognise themselves in a mirror, but have they led us up the garden path?
James Balog/Aurora Photos

What is the evidence for this hierarchy? After all, mental complexity is a slippery concept and, besides, none of us has insight into even the mind of another human, let alone a bat or a beetle. Well, there's no question that some brains are much bigger and more structurally complicated than others. This disparity is mainly the result of the differing evolutionary demands that animals must meet to survive. For example, the nervous system of a sedentary, filter-feeding oyster consists of just two cell clusters. These allow it to do exactly what an oyster needs to do – control its digestion, and transmit signals from light-sensing tentacles to the muscle that snaps it shut when a predator looms. Meanwhile, at the other end of the spectrum, there is one particular demand that seems to have led to the evolution of complex brains and could also have created the conditions for a sense of self to arise. That challenge is dealing with the minds of others – be they prey, competitors or other members of your social group.

According to the social brain hypothesis, developed by Robin Dunbar at the University of

this, brains needed to evolve from being simply things that experience sensations and thoughts to becoming their observer. To do this, they needed to build a model of a mind, according to neuroscientist Michael Graziano at Princeton University. And once the biological machinery for such model-building evolved, it could be used to represent not only the minds of others, but also one's own mind.

A model – be it for mind reading, weather forecasting or whatever – usually starts with some assumptions about the factors that contribute to the system in question and their relative importance. It then runs a simulation and, depending on how much the result diverges from physical observations, modifies the assumptions. The model thus acquires an accurate representation of the forces at work, allowing it to make reasonable predictions about the future. "The brain is a model-builder," says Graziano. "You can't move your arm properly if your motor system doesn't know where it is, can't predict where it will be in the next few seconds, and can't run simulations about what will happen if it sends out this or that command to the muscles." And, he argues, the brain uses exactly the same strategy to model minds so that it can interact socially. If he is correct, then what you consciously experience is the simulation.

"Perhaps self-awareness isn't even a simulation but just a hall of mirrors"

By extension, self-awareness is the conscious state of running that simulation on your own mind. Graziano believes we have no reason to put it on a pedestal. "Self-awareness is not higher-order, or intrinsically more complicated, than consciousness," he says. "It is another example of consciousness." A mind is just an object that some brains can model, and so become aware of. Moreover, it is hard to establish whether this ability is associated with uniquely complex biological machinery. After all, we are still struggling to pin down what consciousness looks like in the brain.

Most researchers agree that the brain operates at least partly by generating simulations. However, many disagree that consciousness is a functional piece of the modelling machinery. Instead, a widely held view sees it as the unintended by-product of information rushing through the closed loop of connections that is the brain. Consciousness can't help existing despite serving no particular purpose, just like the noise emitted by a running engine, which has no bearing on the workings of the engine itself. By this way of thinking, self-awareness isn't even a simulation; it is just a hall of mirrors.

Such emergent phenomena are common in nature. They give the mesmerising impression of complexity and intentionality, despite stemming from a system whose components operate with no regard for the phenomenon itself. One notable example is the collective behaviour of flocks of birds, which can be modelled using individuals driven by just two opposing forces – an instinct to follow their nearest few neighbours, and to back off if they get too close. Apparent complexity emerges even in Petri-dish-bound bacterial colonies, where individual bacteria automatically respond to chemical signals secreted by their neighbours to regulate their proximity. The structure that emerges has no agency or purpose – it is purely an indicator of the forces at work in each individual.

Similarly, self awareness may be an apparently complex phenomenon that emerges from the brain. However, unlike with birds or bacteria, a mind cannot observe its individual components.

moment. But some paths are more well trodden than others. In humans, the predominant connections seem to be those used to contemplate the minds of others – the same connections used to contemplate ourselves. What emerges from this is a pattern that seems constant. To you, that is your sense of self, confined inside the Petri dish of your brain.

In other animals, the well-trodden paths in the brain will be different. In bats, for example, it might be those transmitting information from the echolocation clicks used to construct a 3D model of the world. There will be a huge diversity of emergent mental patterns that serve the various survival needs of different species. Looked at this way, there is no clear hierarchy of consciousness corresponding to mental complexity.

Consider the octopus

In fact, some of nature's most sophisticated minds probably lack a sense of self as we know it. In mammals, those with bigger social groups generally have bigger brains, implying that a sense of self goes hand in hand with intelligence. But some other animals seem to have evolved to be highly intelligent without having had to understand the minds of others.

Take cephalopods – a group of marine animals that includes cuttlefish and octopuses. Having spent years collaborating with marine biologists, philosopher of science Peter Godfrey-Smith at the University of Sydney believes that the particularly large brain of the common octopus is shaped mainly by the unique demands on a soft-bodied animal inhabiting an environment dominated by vertebrates. This challenge might have triggered the evolution of a bodily self-awareness akin to that of primates, but Godfrey-Smith sees a clear distinction between the two. "When one watches an octopus squeeze through a tiny space, it certainly looks [different]," he says. Either way, we can rest assured that if an octopus has a sense of self, it will have very little in common with the "self" that inhabits our brains. It is even less likely to be something we can measure with a mirror.

Indeed, all this makes clear that the best we can hope for with mirrors is an imperfect glimpse into minds like our own. What's more, if we proceed under the assumption that such minds are the true pinnacles of complexity, then we will miss out on the most beautiful thing about minds – that they are biological machines for adaptation, with contents that can be sophisticated in so many ways.

Mirror, mirror

Joe Raedle/Newsmakers

The ability to recognise oneself in a mirror is generally taken to be an indicator of self-awareness, but that idea is being challenged. For a start, developmental psychologists argue that it doesn't necessarily reveal an awareness of self that extends beyond the here and now. Experiments show that children can acknowledge themselves in a mirror at the age of 3, yet cannot recognise themselves in videos taken a few months earlier. They will struggle with the idea of existing in the past for another year or two.

It is even less clear what it means for a non-human animal to recognise itself in a mirror. Only a handful of species seem capable of the task. The majority are either our primate relatives or animals with complex social lives, like us. So, rather than reflecting mental complexity, it could simply indicate that their minds have evolved to face similar challenges to our own. Besides, the discovery that animals can learn to pass the mirror self-recognition test hints that there could be many species with undetected self-awareness.

This article appeared in print under the headline "The why of me"

Leader: "We're not unique – lots of species can recognise themselves"

Magazine issue 3194, published 8 September 2018

Magazine issue 3194, published 8 September 2018