

COMMENTARY

Can self-awareness be taught? Monkeys pass the mirror test—again

Annamarie W. Huttunen^a, Geoffrey K. Adams^b, and Michael L. Platt^{a,b,1}

“Mirrors,” she said, “are never to be trusted.”
Neil Gaiman, *Coraline*

The ability to recognize oneself in the mirror is often held as evidence of self-awareness. However, children under the age of 2 y and most animals do not behave as if their mirrored reflection represents their own face or body, calling their capacity for self-awareness into doubt. In PNAS, Chang et al. (1) build on prior work (2) to provide further evidence that—with extensive training—rhesus macaques pass the mirror self-recognition test, suggesting this training either uncovers latent self-awareness or teaches the monkeys a new cognitive skill. Have monkeys finally earned a hard-won spot among the few exalted species to display self-awareness? If so, what does mirror self-recognition imply for our understanding of the brain mechanisms that support it?

Upon initial exposure to a mirror, most animals react to their reflection as if seeing another animal. Gallup (3) first observed that, after repeated exposure to a mirror, chimpanzees began to interact with it in a more self-directed and less social way. To test whether the chimps actually recognized themselves, Gallup marked the animals' foreheads with an odorless dye while the chimps were anesthetized. Upon awakening, chimpanzees spontaneously began touching the otherwise imperceptible mark using their reflection in the mirror. Gallup, and many others thereafter, argued that such behavior is a decisive measure of self-awareness. Since Gallup's pioneering study, the mark procedure has become the litmus test for self-awareness. Human toddlers begin to pass a version of the mark test, not requiring anesthesia, at around 16 mo to 24 mo (4), whereas individuals with certain neuropsychiatric disorders, notably schizophrenia, show impairments in mirror self-recognition (5). In addition to chimpanzees, a menagerie of distantly related species, from elephants to magpies, have passed the mark test (6). Other primates, including gorillas and (previously) macaques, typically fail to show signs of self-recognition in a mirror (7).

Failing the mark test, however, is not evidence of the absence of self-awareness. On one hand, children

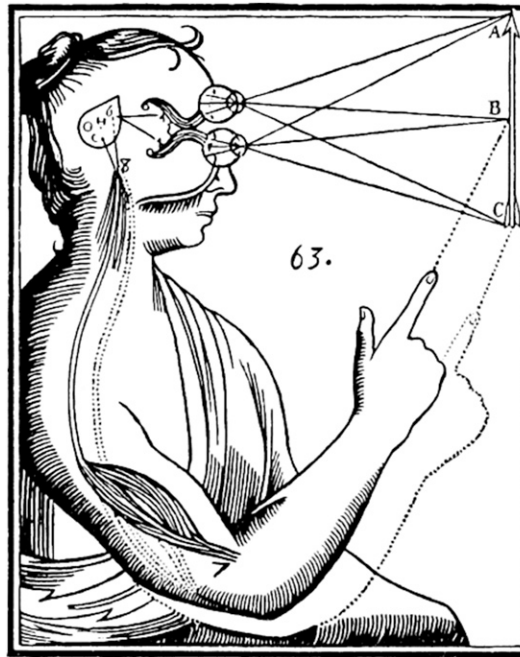


Fig. 1. Reflecting on the self. René Descartes introduced an early version of representationalism, according to which the mind engages with the world through intervening representations constructed by the senses. The “sense of self” expands on this idea to propose that metacognitive abilities like self-recognition rely upon a mental representation of the self. For Descartes, self-awareness was localized to a discrete anatomical substrate—the pineal gland. By contrast, an embodied cognition account of self-awareness regards it as a property of the entire sensorimotor system, rather than a function of a single specialized “self-awareness area.” Figure from René Descartes, *Meditations on First Philosophy*, 1641. Public domain. Retrieved from Wikimedia Commons, https://commons.wikimedia.org/wiki/File%3ADescartes_mind_and_body.gif.

as old as 6 y from rural non-Western societies often do not “pass” the mark test (8), suggesting that differences in culture and experience with mirrors rather than cognitive capacity can influence test performance. On the other hand, there is a strong argument

^aDepartment of Marketing, Wharton School, University of Pennsylvania, Philadelphia, PA 19104; and ^bDepartment of Neuroscience, University of Pennsylvania, Philadelphia, PA 19104

Author contributions: A.W.H., G.K.A., and M.L.P. wrote the paper.

The authors declare no conflict of interest.

See companion article on page 3258 in issue 12 of volume 114.

¹To whom correspondence should be addressed. Email: mplatt@mail.med.upenn.edu.

that species-specific sensory and motor specializations for interacting with the environment, as well as the decidedly unnatural experimental context, may influence behavior displayed in front of a mirror in ways unrelated to self-awareness (9). Nevertheless, the simplicity and intuitiveness of the mark test makes it a useful tool for comparative cognition, if not the final arbiter of self-awareness.

One question these observations raise is whether incentivized experience with mirrors can uncover latent self-awareness or even teach it. An initial study by Chang et al. (2) provided rhesus macaques with visual-somatosensory training in the form of an irritant laser pointer directed at the forehead, visible only in a mirror. This explicit training regimen not only resulted in macaques passing the classic mark test with dyes of various colors but also subsequently evoked spontaneous exploration of otherwise unseen body parts. These findings suggest that previous mark test studies may have confused motivation to touch the mark, or lack of experience with the necessary sensory-motor remapping of mirrored information, with self-recognition.

The Chang et al. (2) study met strong criticism from Anderson and Gallup (10), who argued that using an irritant mark and extensive training effectively “engineered” behavior that only appeared to be self-aware. They contend that, for the mark response to reflect true self-awareness and not simply a “simulation” of self-awareness, the behavior must arise spontaneously, without training, in response to an otherwise undetectable stimulus. This argument is similar to Searle’s (11) so-called “Chinese room” critique of artificial intelligence, which holds that a computer program simulating Chinese does not truly understand the language.

The current study by Chang et al. (1) is a response to these critiques. Rather than training monkeys to detect an irritant spot, monkeys were trained to touch a spot projected onto a surface in the monkey’s close personal space. In an initial training period, monkeys could see the spot directly, as well as seeing its reflection in the mirror, and so could complete the task without using the mirror. Following this training, monkeys failed the classic mark test. In a second training period, monkeys’ direct view of the spot was obscured, forcing them to use the mirror to complete the task. When given the mark test a second time, monkeys touched their own faces near the spot, thus passing the mark test. The authors reasonably conclude that performance on the mark test depends upon an individual’s ability to accurately guide movements using mirrored visual information, and thus failure on the test may not indicate an absence of self-awareness. An intriguing corollary finding of this study is that, after training, monkeys spontaneously initiated self-directed behaviors when a mirror was placed in the home cage. Untrained monkeys failed to use the mirror in this way, suggesting that extensive training can unveil a latent capacity for self-recognition.

For most of us, mirror self-recognition is so automatic and effortless that it is difficult to appreciate its complexity. Even the simplest account requires that reflected visual information is integrated with information from proprioceptive and motor control systems to identify the reflection, an external stimulus not normally associated with one’s own perceptions and actions, as an image of one’s own body (typically conflated with the “self”). In fact, our capacity for this type of multimodal self-recognition is surprisingly flexible, as illustrated by the “rubber hand illusion.” If one views a rubber hand touched while one’s own hand is simultaneously touched, one typically experiences a powerful sense of “owning” the rubber hand (12). Intriguingly, macaques (or at least

neurons in their brains) appear to experience something like the rubber hand illusion as well. While recording activity of neurons in parietal cortex, Graziano et al. (13) presented monkeys with an artificial arm. They found that firing rates of neurons clearly discriminated between the sight of the monkey’s real arm and the artificial arm, but that synchronous touching of the real and artificial arms caused neuronal responses to the artificial arm to more closely resemble responses to the real arm. Although it is not clear that monkeys psychologically experienced the rubber hand illusion in the same way that humans do, this study reveals that plasticity in neural coding of the type required for mirror self-recognition is evolutionarily conserved among primates and is

Clever studies like the one by Chang et al. help expose our preconceptions about ourselves and point the way toward deeper understanding of the way our brains, and the brains of other animals, construct reality and our place within it.

present in a brain area important for both sensorimotor integration and awareness of one’s body and actions (14).

Attempts to identify a neurobiological substrate for a sense of self in humans, independent of the kind of visuomotor tasks described above, have relied on functional imaging of the brain during presentation of self- and other-associated words, images, or even heartbeats. Such studies have identified a conceptual self-awareness network, a set of brain areas, including medial frontal pole, temporal pole, hippocampus, and retrosplenial cortex, specifically involved in processing declarative information about the self (15). This network partially overlaps the “default mode network” (DMN), a poorly understood set of brain regions that is active at rest but suppressed during performance of attention-demanding tasks. Indeed, the DMN has been implicated in several aspects of self-related processes, including projection of oneself into the future and “mentalizing” the thoughts and feelings of others through simulation (16, 17). Notably, the DMN has been identified not only in rhesus macaques (18) but also in rodents (19), implying this network may be an evolutionarily ancestral mammalian adaptation.

Philosophically, Chang et al.’s (1) findings encourage us to reflect upon whether passing the mark test requires a representational explanation, that is, understanding that the mirrored image is oneself (Fig. 1). Radical interpretations of embodied cognition (20) proffer that a sense of self is not explicitly represented in the brain but rather emerges in real time from the dynamic interaction of our bodies with the environment. The mark test requires the ability to integrate mirrored visual information with proprioception and efference copy of movements to guide subsequent behavior. If the brain detects incongruence between visual information and proprioception (i.e., a mark that can be seen but not felt), that might naturally recruit attention and subsequent exploration of the incongruent stimulus (i.e., touching the mark). Brain areas, including anterior cingulate cortex (21) and insular cortex (22), that identify such incongruencies could recruit the so-called “salience network” (23) to guide attention to the relevant stimulus (24), especially when made motivationally significant (25).

Such a model could account for the spontaneous emergence of monkeys’ self-directed behaviors toward body parts that were previously felt but not seen (e.g., inspecting the inside of the mouth or the hindquarters in the mirror). Here, no a priori conceptualization of a self is required to detect the mark as out of place or

to account for subsequent exploration in the mirror. Insofar as self-awareness is defined as representational knowledge of oneself, behavior directed toward one's reflection in a mirror does not necessarily indicate self-awareness. Clever studies like the one

by Chang et al. (1) help expose our preconceptions about ourselves and point the way toward deeper understanding of the way our brains, and the brains of other animals, construct reality and our place within it.

- 1 Chang L, Zhang S, Poo M-m, Gong N (2017) Spontaneous expression of mirror self-recognition in monkeys after learning precise visual-proprioceptive association for mirror images. *Proc Natl Acad Sci USA* 114(12):3258–3263.
- 2 Chang L, Fang Q, Zhang S, Poo M-M, Gong N (2015) Mirror-induced self-directed behaviors in rhesus monkeys after visual-somatosensory training. *Curr Biol* 25(2): 212–217.
- 3 Gallup GG, Jr (1970) Chimpanzees: Self-recognition. *Science* 167(3914):86–87.
- 4 Amsterdam B (1972) Mirror self-image reactions before age two. *Dev Psychobiol* 5(4):297–305.
- 5 Gallup GG, Anderson JR, Platek SM (2003) Self-awareness, social intelligence and schizophrenia. *The Self in Neuroscience and Psychiatry*, eds Kircher T, David AS (Cambridge Univ Press, New York), pp 147–165.
- 6 de Waal FBM (2008) The thief in the mirror. *PLoS Biol* 6(8):e201.
- 7 Suarez SD, Gallup GG, Jr (1981) Self-recognition in chimpanzees and orangutans, but not gorillas. *J Hum Evol* 10(2):175–188.
- 8 Broesch T, Callaghan T, Henrich J (2011) Cultural variations in children's mirror self-recognition. *J Cross-Cult Psychol* 42:1018–1029.
- 9 De Veer MW, van den Bos R (1999) A critical review of methodology and interpretation of mirror self-recognition research in nonhuman primates. *Anim Behav* 58(3):459–468.
- 10 Anderson JR, Gallup GG, Jr (2015) Mirror self-recognition: A review and critique of attempts to promote and engineer self-recognition in primates. *Primates* 56(4): 317–326.
- 11 Searle JR (1980) Minds, brains, and programs. *Behav Brain Sci* 3(03):417–424.
- 12 Botvinick M, Cohen J (1998) Rubber hands 'feel' touch that eyes see. *Nature* 391(6669):756.
- 13 Graziano MS, Cooke DF, Taylor CS (2000) Coding the location of the arm by sight. *Science* 290(5497):1782–1786.
- 14 Blakemore SJ, Frith C (2003) Self-awareness and action. *Curr Opin Neurobiol* 13(2):219–224.
- 15 Tacikowski P, Berger CC, Ehrsson HH (January 24, 2017) Dissociating the neural basis of conceptual self-awareness from perceptual awareness and unaware self-processing. *Cereb Cortex*, 10.1093/cercor/bhx004.
- 16 Qin P, Northoff G (2011) How is our self related to midline regions and the default-mode network? *Neuroimage* 57(3):1221–1233.
- 17 Mars RB, et al. (2012) On the relationship between the "default mode network" and the "social brain." *Front Hum Neurosci* 6:189 10.3389/fnhum.2012.00189.
- 18 Hayden BY, Smith DV, Platt ML (2009) Electrophysiological correlates of default-mode processing in macaque posterior cingulate cortex. *Proc Natl Acad Sci USA* 106(14):5948–5953.
- 19 Lu H, et al. (2012) Rat brains also have a default mode network. *Proc Natl Acad Sci USA* 109(10):3979–3984.
- 20 Clark A (1998) *Being There: Putting Brain, Body, and World Together Again* (MIT Press, Cambridge, MA).
- 21 Kolling N, et al. (2016) Value, search, persistence and model updating in anterior cingulate cortex. *Nat Neurosci* 19(10):1280–1285.
- 22 Greene JD, Nystrom LE, Engell AD, Darley JM, Cohen JD (2004) The neural bases of cognitive conflict and control in moral judgment. *Neuron* 44(2):389–400.
- 23 Seeley VWW, et al. (2007) Dissociable intrinsic connectivity networks for salience processing and executive control. *J Neurosci* 27(9):2349–2356.
- 24 Gao W, Gilmore JH, Alcauter S, Lin W (2013) The dynamic reorganization of the default-mode network during a visual classification task. *Front Syst Neurosci* 7:34: 10.3389/fnsys.2013.00034.
- 25 Platt ML, Glimcher PW (1999) Neural correlates of decision variables in parietal cortex. *Nature* 400(6741):233–238.