

On the zombie within

Christof Koch and Francis Crick

To what extent are we conscious of everything going on in our brains? Nietzsche and Freud popularized the notion of the unconscious as a realm of the mind that controls human behaviour but is not itself accessible to conscious introspection or knowledge. By 'unconscious' we mean any neuronal activity that does not give rise to conscious sensation, thought or memory. Although many of Freud's ideas, involving penis envy, Oedipus complex, the Id and other fanciful creations, are mere myths that lack objective standing, science has provided credible evidence for the existence of sensory-motor systems in the primate brain that function in the absence of consciousness.

Many mammalian brain systems perform complex yet routine tasks without direct conscious input. Melvyn Goodale and David Milner described the brain parts responsible for this as 'online systems', by analogy with software that processes information in real time. Such systems can deal with certain commonly encountered situations automatically, which is why we call them 'zombie' agents. One can be conscious of the actions of one's own zombie, but usually only in retrospect.

The best evidence comes from studying dissociation of 'vision for perception' and 'vision for action' in both healthy humans and

patients. For instance, an experimental subject, sitting in the dark, stares at a single light source and points a finger at it. Suddenly a new light appears in the periphery and the observer has to move his/her eyes and finger rapidly towards this target. While the eyes are in transit, the light is moved a little bit to the left or right. Neither eyes nor finger has a problem in correcting for this and both end up right on target. Yet the subject does not see the extra target motion, even though the oculomotor system automatically corrects for it. All that is perceived is a flashing target and the eyes and hands making a movement towards it.

Other instances of online systems are those that adjust for body posture on the basis of visual or vestibular cues, or that control the way we shape our hand and fingers when we reach out to pick something up. Rigorous experiments involving forced choices between two alternatives show that we can respond to scary pictures of snakes or spiders even when not consciously aware of them.

Within the clinical domain, some neurological patients show very selective deficits. A key element of the Milner-Goodale argument is patient D.F., who cannot visually recognize objects or shapes since suffering from carbon-monoxide-induced anoxia. Ask her whether an elongated slot is vertical or horizontal and she has to guess. Yet when she 'posts' a letter into this slot, she effortlessly rotates her hand into the appropriate position. However, when she is shown the slot, the light is turned off and she has to wait a few seconds before executing the hand movement, she fails. This might imply that her online system does not have access to explicit memory. When she reaches out to pick up a pencil or a glass, she scales her handgrip accordingly, as do normal subjects. Yet she denies seeing the object.

Even more spectacular cases of zombie behaviour can occur in patients with complex partial seizures and in sleepwalkers. Both involve complex yet relatively stereotypical motor patterns: wandering around, moving furniture and even driving cars. This automatic behaviour follows an internal programme that can be influenced by the environment (for example, when sidestepping an obstacle). In general, neither the epileptic patient nor the sleepwalker responds to commands or remembers anything later. The simplest interpretation is that, although consciousness is shut down by the partial seizure or by deep sleep, enough of the forebrain remains active to subservise online systems. Both syndromes accentuate the difficulty of assessing the degree of consciousness in the absence of either explicit recall or language.

Online systems are fast, outpacing consciousness. Anecdotal evidence and psy-

Unconsciousness

"There is credible evidence for the existence of sensory-motor systems in the primate brain that function in the absence of consciousness. So why bother with consciousness?"

chophysical research emphasize rapid and effortless behaviour that predates consciousness. This is particularly true of the highly practised and ritualized sensory-motor activities that humans love, such as rock-climbing, fencing and dancing. Mastery of these requires a surrendering of the conscious mind, allowing the body to take over.

The hallmarks of a zombie system are stereotypical, limited sensory-motor behaviour, and immediate, rapid action. Its existence raises two questions. First, why aren't we just big bundles of unconscious zombie agents? Why bother with consciousness, which takes hundreds of milliseconds to set in? It may be because consciousness allows the system to plan future actions, opening up a potentially infinite behavioural repertoire and making explicit memory possible.

Second, what is the difference between the neuronal pathways that subservise zombie agents and the neural networks that give rise to specific, conscious perception? Both probably involve the cerebral cortex and the thalamus. Are they based on activity in different subsets of neurons, segregated according to brain areas? Or could the neuronal correlates of consciousness correspond to a cell type, intermixed with other types that are responsible for unconscious behaviour? Might the difference be the type of neural activity involved? For instance, we hypothesized that consciousness involves synchronized firing of neurons at the millisecond level, whereas uncorrelated firing can influence behaviour without generating that special buzz in the head.

Could mutation of a single gene turn a conscious animal into a zombie? If so, what test would show that they are unconscious? Tracking down the neuronal correlates of consciousness in humans, monkeys and mice should illuminate the central mystery of how neural activity in specific feedback circuits gives rise to subjective states. ■

Christof Koch is in the Divisions of Biology and of Engineering and Applied Science, Caltech, Pasadena, California 91125, USA. Francis Crick is at the Salk Institute, La Jolla, California 92037, USA.

FURTHER READING

<http://www.klab.caltech.edu/cns120>

Crick, F. *The Astonishing Hypothesis* (Charles Scribner's Sons, New York, 1994).

Milner, A. D. & Goodale, M. A. *The Visual Brain in Action* (Oxford Univ. Press, Oxford, 1995).

CLIVE BARDA/PAL



Caption to come.