

Theories of consciousness: history, ai and animals



**ASSIGN
BUSTER**

Consciousness

- Andrew P Allen

History and Philosophy

People can mean various things when they talk about “consciousness”. At a simple level, one can mean awareness of one’s world or one’s internal drives (e. g. thirst). A more complex form of consciousness is awareness of one’s own awareness, the consciousness that allows people to psychologise about themselves. Approaching the concept from a different angle, “consciousness” sometimes means the sense of what it is like to be someone or have a particular experience. Although we may have a sense of what an experience is like, it is very difficult to describe exactly what the experience is like (c. f. Ned Block, 1990, for an interesting discussion).

A key issue within philosophy of mind is the “mind-body problem”: can a physical body produce a subjective, apparently non-physical mind, and if so, how? Materialists take the position that the mind is the product of the brain, while dualists hold that body and mind are not the same thing. The position of dualism is typically associated with René Descartes, who suggested that mind and body are two different types of matter (see <http://plato.stanford.edu/entries/dualism/#HisDua> for a discussion).

In attempting to explain how the brain produces awareness, neuroscientists would tend towards materialism. Regarding brain and consciousness, Place (1956/1990) has drawn an analogy with clouds and the droplets of water that form them. Although a cloud observed at a distance and droplets of water observed close-up *seem* very different, the many droplets of water

nonetheless make up the cloud. So it (perhaps) is with the brain and consciousness; the firing of a neuron may seem very different from a mental image of a new car, but there is no reason to say that this mental image *cannot* consist of nothing more than the action of many neurons.

Daniel Dennett has criticized what he calls the “Cartesian theatre”; a given place in the brain where sensations, memory traces etc. are combined to form consciousness. There is a danger of positing a neural “homunculus” (a “little man” in the head) which observes the various non-conscious parts of the brain and turns them into conscious experience. It does at least seem evasive to propose a single part of the brain is responsible for turning sensations from unconscious information processing to conscious experience without specifying the process whereby such a change occurs. Dennett is setting a high bar for the neural correlate of consciousness (see below); you have to give a full explanation of how the process of consciousness is brought about by the brain without suggesting that some brain area just acts in a conscious way.

If we accept that the brain (and the rest of the body?) produces consciousness, then we have to reject dualism (Edelman, 2003), or at least a strong version thereof. Edelman points out that consciousness has a wide range of interesting properties (e. g. it feels unitary, so it seems it requires the binding of multiple sources of sensory information). He suggests that evolutionary pressure would favour cognitive structures which could integrate information from multiple sources.

Consciousness in the brain

Given that consciousness is stopped when the activity of some regions is stopped, it seems fair to assume that the brain may be responsible for consciousness. *However*, the question remains: how do these brain regions lead to the conscious experience (Churchland, 2012)? Crick and Koch (1998) highlight some of the key issues. At any time, the brain is doing a lot of things, but only some of these things appear in our consciousness. Is there anything special about the neurons involved in consciousness and their type of firing? What about the connections between them? There has been some interest in finding a so-called *neural correlate of consciousness*.

Edelman (2003) takes the approach of looking at connections. He posits “re-entry” as a process which could account for how functionally distinct parts of the brain co-ordinate their activities to produce a combined output. It involves recursive signalling over multiple pathways which are used simultaneously. He suggests that this process allows for the binding of outputs from different brain areas to form an integrated sense of experience. Edelman suggests the thalamocortical system as “a dynamic core” for consciousness. The thalamic intralaminar nuclei (ILN) may play a particularly important role in consciousness; it projects axons widely to all cortical areas, and small lesions to the ILN are associated with significant loss of awareness (Bogen, 1997). Note that the ILN may be necessary but not sufficient for consciousness; it is through its interaction with cortical regions that it could produce something like consciousness. The thalamocortical system contains functionally distinct sub-parts which may act semi-independently, while also being able to integrate information between themselves. By suggesting that consciousness could be brought about by brain processes and their

interaction, Edelman's idea may avoid falling into the trap of the Cartesian theatre.

Attention and consciousness

At first, it might seem like attention and consciousness might be the same thing; when we attend to something, we are conscious of it, and when we are conscious of something, we are attending to it, right? However, it has been argued that you can have either consciousness or attention without having the other (Koch and Tsuchiya, 2007). They cite work which uses interocular suppression (i. e. presenting different images to each eye in order to reduce perception of some/all of these images) to present both a nude image and a meaningless scramble of its pixels, while simultaneously rendering the nude image invisible to consciousness. Nonetheless, heterosexual participants attend to nude images of the opposite sex more than scrambled control images (Jiang et al., 2006). Hence, attention without consciousness! Another example of attention without consciousness is *blindsight*, where patients with damage to the primary visual cortex can report properties of visual stimuli above chance level, but without awareness of having seen anything (Weiskrantz, 1997). Subliminal presentation of stimuli can be processed by brain areas associated with emotional processing, such as the amygdala (Naccache et al., 2005).

I'm less convinced by Koch and Tsuchiya's argument that one can have consciousness without attention. Their argument seems to be based on limiting their point to top-down attention processes. For example, they suggest that one can make out the gist of an image after a very brief

presentation. Of course, there may be little top-down processing going on here, and 30 ms may be too short a time to talk about “sustained attention”, but after all, one has to orient to the image in order to perceive it. Perhaps you may see it otherwise...

Are non-humans conscious?

Trying to define consciousness at a brain level may be even more difficult when it comes to non-human animals. This question is also important for the ethical consideration of neuroscientists who work with animals. If one is to work with a particular species, one should at least try to be aware of its capacity for suffering.

Panksepp (2005) argues that affect is largely produced by processes concentrated in subcortical, limbic regions in the mammalian brain. He defines consciousness as brain states which are associated with feeling or experience. He distinguishes raw, primary-process consciousness from secondary consciousness, which can relate to how external events relate to internal states, and tertiary consciousness, which is basically meta-cognition.

Panksepp attacks what he seems to perceive as a wilful ignorance of the affective experience by neuroscientists working with animals, and criticises those who suppose that all consciousness is dependent upon the advanced linguistic and reasoning skills possessed by humans. However, the fact is acknowledged that outward behaviour may give a misleading impression of internal affective states. Nonetheless, he defends an internal affective life in animals, citing evidence of differing vocalisations of rats in response to environments associated with pleasurable/unpleasant drugs (Burgdorf,

Knutson, Panksepp, & Ikemoto, 2001a, 2001b), as well as neural mechanisms underlying desire for certain drugs which are similar to those in humans. Given similar subcortical machinery in other mammalian life, such research may give insight into the affective life of humans. However, studying consciousness in animals can be tricky; although anaesthesia is often used in certain techniques, if one wishes to study *consciousness* then any form of anaesthesia or sedation may bias results (Crick & Koch, 1998).

The work of Gallup (1970) used a simple behavioural test to examine self-awareness in chimpanzees. A mirror was inserted in their environment. Although the animals initially responded socially to it, they began to groom in response to it. When they were marked with a red dot in their sleep they used the mirror to try to clean the dot off. However, this level of performance was not evident in other primates.

However, the so-called “hard problem of consciousness” (what is it really *like?*) may be insoluble. Thomas Nagel (1974) famously used animal life as an illustration of how difficult it is to grasp *qualia* (i. e. the subjective feeling of what something is like) by asking the question “what is like to be a bat?” Aside from bringing up again the issue of knowing others’ minds, the comparison here is stronger because it shifts from trying to second-guess the thoughts of fellow humans to trying to imagine the thoughts and feelings of a strange species. The implication is that, even if we were to understand all the neural processes tied up with the bat’s nervous system which bring about consciousness, we would still not be able to fully imagine what it is like to be a bat.

Artificial intelligence and models of altered consciousness

Although a large proportion of neuroscience involves backwards engineering of the brain (i. e. taking something which has already been engineered by evolution and trying to tease apart its structure and function), artificial intelligence, by engineering intelligent systems, can also be used to observe if a particular account of how the brain works actually produces a comparable output when you run it through a computer program (if the program doesn't produce the same output as the "natural" brain, this may pose a problem for your theory, or vice versa). (Note this process of back-propagation is somewhat reminiscent of Edelman's idea of re-entry).

Takeno has found that the robot can distinguish between its own image in a mirror from either a second robot or another robot which follows the test robot's instructions. The robot is equipped with LED lights allowing it to demonstrate distinct responses to its own mirrored behavior compared to that of another robot, including another robot engaging in the same behaviour (Takiguchi, Mizunaga, & Takeno, 2013). See the following brief video: <https://www.youtube.com/watch?v=TK0M02aKXLE>

A neural network was used to model how the excessive loss of synapses during adolescence could lead to auditory hallucinations reported in schizophrenia (Hoffman & McGlashen, 1997). Pruning was carried out in a "Darwinian" fashion by removing neural units which were less well-connected to other units, in addition to modelling cell death which could be associated with excessive loss of neurons. Excessive loss of neural units produced a model of hallucination whereby words were coming up as perceived at the

output layer of the network even when words were not being entered at the input layer. Although the authors admit that such models are vastly simplified models of the real thing, by reproducing (modeled) phenomena visible in the world (in this case, auditory hallucinations), they allow one to study such phenomena by testing if the mechanisms one hypothesizes explain such phenomena (in this case, excessive loss of neurons involved in working memory) actually produce the phenomenon under investigation. Interestingly, the neurons pruned were modeled on corticocortical connections rather than thalamocortical connections (the type suggested by Edelman to play a key role in producing conscious experience itself).

Consciousness: a clinical case

A vegetative state is where a patient shows no overt signs of awareness, even though they are visibly awake. However, the idea that people in a persistent vegetative state lack consciousness has been challenged by recent research. Patients in a minimally conscious state or persistent vegetative state have been instructed to perform mental imagery tasks while undergoing fMRI (Monti et al., 2010). The tasks used are associated with activity in the parahippocampal gyrus and the supplementary motor area; areas which are associated with actually carrying out the activity. A minority of the participants showed activity in response to the tasks similar to healthy controls. However, bearing in mind that information can be processed without conscious awareness (as alluded to in the discussion of attention and consciousness), is it possible that this brain activity may have emerged automatically, without the patients having any conscious awareness of the scene described to them by the researchers? Such an interpretation is

challenged by the following finding: a number of healthy controls and 1 patient were asked questions, and instructed to think of one mental image if the answer was “ yes” and a different mental image if the answer was “ no”. The patient showed signs of being able to complete this task. The fact that the participants could control what was imagined suggests that they may have been aware of their own awareness.

Adrian Owen talks about these issues at the following link: <http://tedxtalks.ted.com/video/TEDxUWO-Adrian-Owen-The-Quest-f>

Embodiment

The idea that the brain is, or at least is very much like, a computer is quite popular. Indeed, computers themselves have increasingly sophisticated artificial intelligence. Of course, a lot of the information we process is not purely symbolic for us; it is viscerally linked to our bodily states and physiological drives, and thus *embodied* . Returning to the question of what it is like to be a bat, we can consider the brain of this animal and how it works, but even if we could understand all brain functions of the bat, there would still be other differences between our species. For example, bats have wings which they can use to fly. What is it really *like* at a subjective level to do this? If we were to both given the chance to experience this kind of flight, your answer to this question could be completely different from mine, and yet perhaps we would both be right about our own experience.

References

Block, N. (1990). Inverted Earth. *Philosophical Perspectives* , 4 , 53-79.

Bogen, J. E. (1997). Some neurophysiologic aspects of consciousness. *Seminars in Neurology* , 17 (2), 95-103.

Burgdorf, J., Knutson, B., Panksepp, J., & Ikemoto, S. (2001a). Nucleus accumbens amphetamine microinjections unconditionally elicit 50 kHz ultrasonic vocalizations in rats . *Behavioral Neuroscience* , 115, 940-944.

Burgdorf, J., Knutson, B., Panksepp, J., & Shippenberg, T. (2001b). Evaluation of rat ultrasonic vocalizations as predictors of the conditioned aversive effects of drugs. *Psychopharmacology* , 155, 35-42.

Churchland, P. M., (2012). Consciousness, in: Gregory, R. L. (Ed.), The Oxford companion to the mind. Oxford University Press, Oxford, UK.

Crich, F., Koch, C. (1998). Consciousness and neuroscience. *Cerebral Cortex*, 8 , 97-107.

Edelman, G. (2003). Naturalizing consciousness: A theoretical framework. *PNAS* , 100 (9), 5520-5524.

Gallup, G. (1970). Chimpanzees: Self-recognition. *Science* , 167 (3914), 85-87.

Hoffman, R. E., & McGlashen, T. H. (1997). Synaptic elimination, neurodevelopment, and the mechanism of hallucinated “ voices” in schizophrenia. *American Journal of Psychiatry* , 154 , 1683-1689.

Jiang, Y., Costello, P., Fang, F., Huang, M., He, S., (2006). A gender- and sexual orientation-dependent spatial attentional effect of invisible images. *Proceedings of the National Academy of Sciences* 103, 17048-17052.

<https://assignbuster.com/theories-of-consciousness-history-ai-and-animals/>

Koch, C., Tsuchiya, N., (2007). Attention and consciousness: two distinct brain processes. *Trends in cognitive sciences* 11, 16-22.

Monti, M. M., Vanhaudenhuyse, A., Coleman, M. R., Boly, M., Pickard, J. D., Tshibanda, L., Owen, A. M., Laureys, S., (2010). Willful modulation of brain activity in disorders of consciousness. *New England Journal of Medicine* 362, 579-589.

Naccache, L., Gaillard, R., Adam, C., Hasboun, D., Clémenceau, S., Baulac, M., Dehaene, S., Cohen, L., (2005). A direct intracranial record of emotions evoked by subliminal words. *Proceedings of the National Academy of Sciences of the United States of America* 102, 7713-7717.

Nagel, T. (1974). What is it like to be a bat? *Philosophical Review* , 83 , 435-450.

Panksepp, J. (2005). Affective consciousness: core emotional feelings in animals and humans. *Consciousness and cognition*.

Place, U. T., (1956/1990). Is consciousness a brain process?, in: Lycan, W. G. (Ed.), *Mind and cognition: An anthology*. Blackwell, Malden, Massachusetts, pp. 14-19.

Takiguchi, T., Mizunaga, A., & Takeno, J. (2013). A study of self-awareness in robots. *International Journal of Machine Consciousness*, 5 , 142.

Weiskrantz, L. (1997). *Consciousness lost and found* . Oxford: Oxford University Press.