

ANIMAL SENTIENCE

AN INTERDISCIPLINARY JOURNAL ON ANIMAL FEELING

Shanahan, Murray (2016) [Consciousness as integrated perception, motivation, cognition, and action](#). *Animal Sentience* 9(12)

DOI: 10.51291/2377-7478.1145



This article has appeared in the journal *Animal Sentience*, a peer-reviewed journal on animal cognition and feeling. It has been made open access, free for all, by WellBeing International and deposited in the WBI Studies Repository. For more information, please contact wbisr-info@wellbeingintl.org.



Consciousness as integrated perception, motivation, cognition, and action

Commentary on [Klein & Barron](#) on *Insect Experience*

Murray Shanahan

Department of Computing
Imperial College London

Abstract: This commentary has two aims: first to clarify the behavioural grounds for the ascription of consciousness to non-human animals (including insects), and second to show how Klein & Barron's views can be reconciled with the core claims of global workspace theory.

[Murray Shanahan](#), Professor of Cognitive Robotics, Imperial College London, is interested in the principles that underlie sophisticated cognition, both as it is found in Nature and as it might be realised artificially.

<https://www.doc.ic.ac.uk/~mpsha/>



1. Behavioural Hallmarks of Consciousness in Other Animals

In due course I'm going to discuss the brains of invertebrates and Klein & Barron's (K & B, 2016) claim that insects have the capacity for subjective experience. But first I want to talk about the behaviour of cats. What is it that convinces us, without recourse to neuroscience, that we are in the company of a fellow conscious creature when we are with a domestic animal such as a cat? A child might be taken in by doleful eyes and soft fur (features lacking in insects). But a considered response must look to relevant behavioural cues. What counts is the way an animal reacts to its surroundings, the awareness it manifests of the world and how the world can meet its needs.

A cat's awareness of the world is manifest in a number of ways. It orients towards and attends to objects it perceives through sound or sight or smell. It is especially interested in animate objects such as humans and other animals. It might approach an object it perceives, or retreat from it. It might paw at a perceived object or sniff it or bite it. Much of a cat's behaviour, such as grooming or scratching, is directed towards its own body, which is, in one sense, just another object. Most of the world, though, is empty space, by which I mean space through which an animal can move. A cat is adept at doing this, squeezing through gaps, jumping onto walls, and so on. Significantly, a cat's behaviour can be directed towards an object (or indeed towards a portion of empty space) even when it is no longer directly perceived. When a mouse disappears into a hole, a cat will wait for it to reappear. Without this sensitivity to the persistence of objects, a cat's behaviour might be seen as simply a series of responses to immediate stimuli rather than manifesting an awareness of the world we all share.

As well as displaying an awareness of the world and its contents, the actions of an animal fulfil an underlying purpose. They serve the animal's biological needs, such as attaining food, water, shelter, and the company of others, while avoiding harm. A cat will *approach* a potential food item such as a mouse and (typically) *retreat* from a potential predator such as a dog, not the other way around. But in a complex world, the possibilities for action are numerous. Multiple sources of food, water, or companionship might be available at the same time, along with multiple threats or aversive objects, all occupying different regions of the space accessible to the animal. Since it only has one body, and since it can only move in one direction at a time, it must weigh up the options in relation to its bodily needs, integrating all the relevant information at its disposal, and choose. To choose is to commit the whole animal to a single course of action. In short, the constraints imposed by having a single body confer a degree of unity on the animal's perception, motivation, cognition, and action.

Without knowing anything about their brains, we are inclined to see animals who conform to the above description as fellow conscious creatures, as capable of subjective experience. Although their emphasis is neuroanatomy rather than behaviour, this observation gels with K & B's case for insect consciousness. K & B argue that (a) the evolutionarily ancient core control system of the vertebrate brain has the necessary functional architecture to support subjective experience (Merker, 2007), and (b) the same functional architecture is present in the insect brain (Barron & Klein, 2016). The elements of this functional architecture, common to the basal brains of both vertebrates and insects, map closely onto the behavioural traits enumerated above. As K & B say, "both systems have specialized regions for processing the position of the moving animal in space. In both systems action selection is resolved by combining information on position with information on the environment, the relevance of stimuli in the environment to the animal, and the state of the animal."

K & B's neuroscientific conclusion that "processing of this kind supports the capacity for a subjective experience of the environment" complements the behavioural perspective just outlined. If insects were closer to human size, if their actions were closer to human speeds, and if their appearance were more appealing to the average non-entomologist, perhaps a closer acquaintance with their behaviour would lead naturally to the same conclusion as K & B's, even without an appeal to neuroscience.

2. The Relation to Global Workspace Theory

Now, suppose we accept these minimal behavioural and/or neural requirements for the capacity for subjective experience. What are the implications, with respect to consciousness, of the possession of larger, more complex brains, such as those of humans? Merker (2007), whose opposition to corticocentric theories underpins K & B's target article, remarks that "the telencephalon serves as a medium for the increasingly sophisticated elaboration of conscious contents." Beside our awareness of the outside world as revealed to us through our senses, humans enjoy an inner parade of thoughts, memories, associations, and imaginings, all coloured by a palette of emotions and open to reflective, higher-order conscious scrutiny. This is the domain of the human cortex.

But where does all this leave other prominent candidates for a scientific theory of consciousness, such as integrated information theory (Tononi, 2008; Tononi & Koch, 2015) and global workspace theory (GWT; Baars, 1988; Shanahan, 2010; Dehaene et al., 2014). To what extent are they compatible with K & B's theory? As K & B affirm, the strong claims of Tononi's integrated information theory do not sit well with their standpoint, according to which "it is not integration *per se* that matters ... but the dynamic and ongoing connection between perception, interoception, associative memory, and motor feedback." But what about GWT? GWT is not discussed in the target article, but in the authors' earlier *PNAS* paper (Barron & Klein, 2016) they acknowledge that both theories "emphasize the role of consciousness in bringing together disparate brain processes into a common arena." On the other hand, as Barron & Klein note, GWT has a strong cortical bias, "focusing on the contribution of human fronto-parietal regions to reflective self-awareness of our mental states."

The affinity between the two theories might run deeper than Barron & Klein suspect, as I will now attempt to show. But there is an obstacle to reconciling them. Specifically, GWT purports to account for the distinction between conscious and non-conscious processing in the brain, where reportability is the gold-standard behavioural indicator of conscious processing in humans. But what looks like non-conscious processing to GWT looks like subjective experience according to K & B. To make the point more clearly, a brief introduction to GWT is in order.

In its original guise, GWT was couched in terms of a computational architecture comprising multiple parallel processes that communicate through a global workspace. Conscious processing was hypothesised to involve broadcast of information, via the global workspace, to the full cohort of parallel processes, while non-conscious processing was localised to a small subset of processes (Baars, 1988). In its more neural contemporary guise, the same essential distinction is preserved, but the concept of broadcast is refined into that of an integrated brain state wherein multiple, functionally distinct brain regions exercise mutual influence on each other (Shanahan, 2010; Dehaene et al., 2014). This integrated brain state is realised by the connective core of the brain's white matter pathways, which constitute a modular small-world network with hub nodes (Shanahan, 2012). Consciously mediated behaviour correlates with globally integrated brain states, while automatic or habitual behaviour (such as driving a familiar route or brushing teeth) does not require this level of integration.

The difficulty here, in relation to the proposal of Barron & Klein, is that many aspects of automatic or habitual human behaviour seemingly meet the criteria for subjective awareness of the environment, implicating the relevant basal brain structures, despite the fact that the subject cannot report them. For example, in order to brush my teeth, I have to pick up the toothbrush and toothpaste and execute a complex motor sequence involving the two objects. At some level, my brain has to register the objects that are within reach, it has to select two specific objects to pick up, and it has to be sensitive to their location and orientation in order to do so. Yet immediately after performing this action, I might be unable to recall what other objects were in the vicinity, let alone their locations or orientations. How are we to reconcile the minimal behavioural and neural requirements just proposed,

according to which (it seems) this action would involve subjective awareness, with GWT, according to which it does not?

One way to attempt a resolution would be to appeal to the distinction between primary and higher-order consciousness. According to this way of thinking, when I pick up my toothbrush there is primary consciousness of the relevant sensory stimuli, but higher-order consciousness is lacking. I am not aware of my awareness, hence my inability to report the toothbrush's orientation. Similarly, when a bee selects a particular flower to alight on, it has primary consciousness of the flower's appearance. But the bee is not aware of its awareness. Indeed it lacks the necessary (cortical) apparatus for higher-order consciousness altogether. However, this approach strikes me as unsatisfactory. From a first-personal standpoint, the toothbrush's location simply didn't impact on my conscious experience. As far as I'm concerned, it made no contribution at all to what it was like to be me in the relevant few minutes of my life. It would be imprudent to quarrel with such a first-personal claim from a third-personal point of view, and there is little support in our ordinary language for the possibility of primary consciousness without higher-order consciousness in humans.

So the desired reconciliation must be achieved some other way. As I see it, the key is to recognise that it is possible for the very same type of behaviour to qualify as an indicator of subjective awareness in some animals but not in others. Whether it qualifies or not depends on the neurological context. To see an animal as a fellow conscious creature is to perceive a certain unity. It is to see in the animal the potential for a unified, integrated response to the ongoing situation, a response that draws on the resources of the system as a whole. In a brain with a large cortex, a properly unified response must integrate much more than in a smaller brain. The point is that in the conscious condition the available resources must be integrated *whatever they happen to be* for the animal in question. In a large brain with a highly elaborated cortex there are a lot of resources to integrate. In the brain of a small invertebrate there are fewer.

The upshot is that, for humans, effective behaviour is possible without full integration, implicating just a small set of sensorimotor processes, and without recourse to language, working memory, executive control, and so on. This sort of behaviour is not deemed conscious, either by the subjects themselves or by external observers. On the other hand, when a cat exhibits analogous behaviour, we see this as evidence of subjective awareness of the world. The central claim of GWT, shorn of its cortical emphasis, is that integration of the brain's neural resources is the hallmark of the conscious condition in both cases. Seen in this light, GWT applies not only to mammals, but also to birds (Shanahan et al., 2013), and potentially to invertebrates including insects. And if insects are admitted into the consciousness club, then notwithstanding K & B's passing remark that "all currently existing man-made artifacts fall below the line," perhaps it won't be long before some form of embodied artificial intelligence makes the grade too.

References

- Baars, B.J. (1988). *A Cognitive Theory of Consciousness*. Cambridge University Press.
- Barron, A.B. & Klein, C. (2016). What Insects Can Tell Us About the Origins of Consciousness. *Proceedings of the National Academy of Science*, 113(18), 4900–4908.
- Dehaene, S., Charles, L., King, J.-R. & Marti, S. (2014). Toward a Computational Theory of Conscious Processing. *Current Opinion in Neurobiology*, 25, 76–84.
- Klein, C., & Barron, A. B. (2016). [Insects have the capacity for subjective experience](#). *Animal Sentience* 2016.100.
- Merker, B. (2007). Consciousness Without a Cerebral Cortex: A Challenge for Neuroscience and Medicine. *Behavioral and Brain Sciences*, 30, 63–134.
- Shanahan, M. (2010). *Embodiment and the Inner Life: Cognition and Consciousness in the Space of Possible Minds*. Oxford University Press.
- Shanahan, M. (2012). The Brain's Connective Core and its Role in Animal Cognition. *Philosophical Transactions of the Royal Society B*, 67, 2704–2714.
- Shanahan, M., Bingman, V.P., Shimizu, T., Wild, M. & Güntürkün, O. (2013). Large-scale Network Organization in the Avian Forebrain: A Connectivity Matrix and Theoretical Analysis. *Frontiers in Computational Neuroscience*, 7, 89.
- Tononi, G. (2008). Consciousness as Integrated Information: A Provisional Manifesto. *Biological Bulletin*, 215, 216–242.
- Tononi, G. & Koch, C. (2015). Consciousness: Here, There and Everywhere? *Philosophical Transactions of the Royal Society B*, 370, 20140167.