



UNDER CONSTRUCTION: Please ignore while this notice remains.

Origins of Many Types of Consciousness (DRAFT: Liable to change)

Aaron Sloman

<http://www.cs.bham.ac.uk/~axs>

[School of Computer Science, University of Birmingham](http://www.cs.bham.ac.uk/~axs)

This paper is available as html or pdf

<http://www.cs.bham.ac.uk/research/projects/cogaff/misc/consciousness-origins.html>

<http://www.cs.bham.ac.uk/research/projects/cogaff/misc/consciousness-origins.pdf>

A partial index of discussion notes is in

<http://www.cs.bham.ac.uk/research/projects/cogaff/misc/AREADME.html>

Consciousness as a biological phenomenon

Scientists and philosophers have been thinking about consciousness for centuries (e.g. Plato with his image of prisoners in a cave experiencing shadows on a cave wall, produced by external objects that cannot be seen), although Kathy Wilkes https://en.wikipedia.org/wiki/Kathy_Wilkes argues in [Wilkes\(1984\)](#) that we use the word "consciousness" in different senses, labelling different phenomena and it is not fit to be used as a label for a topic for scientific explanation, because "... *no precise characterization could come close to capturing the thoroughly imprecise and heterogeneous everyday meaning that the term has in the vernacular. The adaptation that the term 'consciousness' would need to undergo before it could be made to cover tidily a systematically related bunch of behaviours would be so great that a study of this 'consciousness' would no more be a study of consciousness as we think of it than the study of the spin of an electron can inform us about the behaviour of whipping-tops.*"

She also (rightly in my view) criticises attempts to characterise consciousness as a topic for systematic investigation by using Nagel's formula [Nagel\(1974\)](#) based on the phrase "What it is like to be ..."

She goes on "*I conclude on a more speculative note. I suggest--no more--that just as science can dispense with the concept of consciousness and lose thereby none of its comprehensiveness and explanatory power, so too could ordinary language. Evidently, this is not a recommendation that we should drop the term; few proposals for linguistic reform have any effect. The suggestion is rather than the presupposition that consciousness is an important, or a real, phenomenon should be*

dropped."

I suggest that we can deal with her objections by treating consciousness by analogy with other topics for biological investigation, including, reproduction, development, disease, food, digestion, locomotion, prey, predator, symbiont, and many more. These do not refer to specific physical, perceptually identifiable entities, states, processes or relations, but to complex collections of relationships relevant to life and the processes that life produces or which serve, impede, or threaten it. Such concepts have the feature that Gilbert Ryle and some of his contemporaries labelled "polymorphism".

On this basis, the fashion among many contemporary researchers to assume that "consciousness" (or some special variant of consciousness often labelled "phenomenal consciousness") can be defined using some simple slogan, e.g. "What it feels like, or what it is like, to be something" (the proposal launched in [Nagel\(1974\)](#)), can be criticised as naively oversimplifying a large and complex collection of phenomena that need to be studied in a far more sophisticated manner, as life itself has increasingly been studied, especially since Darwin.

Fashion and familiarity can provide the illusion of content, even among well educated academics who should know better. Many years ago, I lampooned the "what it is like" approach to specifying a useful concept of consciousness in a semi-serious essay on "What it is like to be a rock": <http://www.cs.bham.ac.uk/research/projects/cogaff/misc/rock/>.) I recently discovered that the philosopher Peter Hacker had presented his own more sober demolition in [Hacker\(2002\)](#).

But criticism of poor terminology and feeble explanations is not enough. My aim is to present a potentially more philosophically and scientifically fruitful approach to research on the many and varied phenomena subsumed under the label "consciousness" (and related words in other languages), emphasising evolutionary origins, biological functions, biological diversity and the need for [explanatory mechanisms](#). (Compare the study of [life](#) -- another slippery subject now including a far wider range of organisms than had been thought of by the time of one of the early pioneering biologists, Aristotle.)

The forms of consciousness that we know of occur mainly in living things, although there are very primitive variants in current AI systems, including forms of introspection in self-monitoring self-debugging systems (one of the topics investigated in [the CogAff project](#)) and perhaps more human-like forms of consciousness in future AI systems that are not yet on the horizon. We can also talk of crowds, or communities or even nations becoming conscious of something, though this paper does not discuss such cases. Population attitudes, emotions, values, goals, etc. are also interesting, but ignored here.

McCarthy

John McCarthy once claimed that a thermostat controlling a heating system has primitive beliefs and goals, and that holding a candle under the thermostat can fool it into having a false belief that causes it to turn off a heater. In contrast, control mechanisms in living organisms are typically far more complex. Instead of achieving or maintaining a single simple state, like being close to a certain temperature, they compete and cooperate in achieving, preserving, or preventing a collection of different states with changing priorities. As organisms become more complex through processes of individual growth and development, or through evolutionary processes, the competing and cooperating control subsystems within each individual become more complex and varied, including control systems for controlling other control systems, e.g. comparing the need to obtain

food and the need to avoid becoming food for something else.

and that there was once no life, or consciousness, on this planet shortly after it first formed, and very likely that there was no life in our galaxy when it first formed, and no life or consciousness in the universe shortly after the Big Bang.

Consciousness is a naturally occurring biological phenomenon taking many different forms (as is common among products of biological evolution), and any serious scientific or philosophical study of consciousness should start from that variety, and place it in the context of evolution, individual development, and the many known functions of consciousness. Moreover, as with evolved forms of life, we need to search for mechanisms underlying consciousness. The mechanisms may be as varied as the mechanisms enabling many forms of life.

Forms of consciousness that we know of occur in products of biological evolution, e.g. in human minds (including minds at various stages of development), and many non-human minds. Some researchers are even studying forms of plant consciousness, though this is controversial. (It isn't clear to me whether that is a substantive or a terminological controversy.)

So anyone wishing to define, or characterise or explain consciousness should examine a wide variety of examples, looking for relevant similarities and differences, and, if possible, generative mechanisms, instead of simply assuming that a fashionable slogan suffices for identifying the topic. Compare trying to define "matter" as "What occupies space" and "Space" as what can be occupied by matter and processes involving matter (and possibly other things).

Wilkes

See [Wilkes\(1984\)](#) for an unusually detailed and broad-minded survey of problems and theories about consciousness, stressing the [heterogeneity](#) of phenomena involving various facets of consciousness or its absence -- e.g. unconsciously adjusting one's posture to reduce (unconscious?) discomfort, including various abnormalities of consciousness connected with drugs, hypnosis, medical conditions, sleep-walking, dissociations produced by brain damage or hypnosis, etc. Wilkes used a range of examples to challenge some of the over-simplifying assumptions usually made by those aiming to produce a scientific or philosophical theory of consciousness, or an explanation of how brains enable it.

However, like many thinkers, she makes unjustified claims about what computer-based minds cannot achieve:

"...The computer, though, like the frog, is severely limited in what it can do with 'seen' objects-hence, of course, the legitimation of the scare-quotes around 'seen'. Computer 'perception' shares too little of the causal/functional role that helps define animal perception to justify removal of the scare-quotes; and, although the difference may be one of degree rather than kind, the degree is very considerable and may be well-nigh insurmountable"(page 35).

I suspect that had she learnt more about the varieties of possible AI mechanisms, and the steadily increasing variety of hardware, software, and networked options for implementing computer-based machinery, she might have retracted that statement (based on [Dreyfus\(1979\)](#)). However, she rightly notes that in future we'll need "careful description of what kinds and levels of information processing we find and do not find in the system (human or artefactual) in question".

Moreover, unlike most who discuss consciousness she rightly includes *propositional attitudes*, such as beliefs, though she could have added hopes, fears, preferences, intentions, values, attitudes, ambitions, skills (mental and physical) and various stored resources for perceiving, thinking, choosing, acting, reflecting, learning, evaluating, composing, creating, ... etc., among the mental contents that can be conscious or unconscious, and which might, in future, be replicated in computer-based agents. However, she also suggests that when flaws in our current terminology and naive theories are removed "the question of consciousness simply drops out; anything that we needed the term 'conscious' to do can be handled more efficiently without it" and she concludes "that consciousness is not the sort of thing that has a 'nature' appropriate for scientific study". But that leaves open whether all the phenomena that we now think of or will in future think of as involving consciousness can be explained in terms of (biological) information-processing mechanisms and can in principle be replicated using non-biological information processing mechanisms.

In a way Wilkes by-passes some of the questions on the basis of her claim

"We are then, I suggest, thrown back on the idea that 'consciousness' and 'conscious' are terms of the vernacular which not only need not but should not figure in the conceptual apparatus of psychology or the neurosciences, for the concepts required in those theories will want to group phenomena along different, more systematic principles. We can support this argument by looking briefly at the way we use the notion in our everyday language; for by so doing it will be seen that consciousness is not the sort of thing that has a 'nature' appropriate for scientific study."

I have a lot of sympathy with that position, which is why I've been basing my work not on any definition of "consciousness" or collection of facts relating to consciousness, but on a wide variety of human and non-human capabilities which need to be explained, some of which, but not all, have been replicated, often in simplified forms, in AI systems (whether rule-based, logic-based, symbolic, sub-symbolic, etc.).

I think that's a step too far. We now know a great deal more about the variety of forms of life and mechanisms involved in producing and maintaining life than Aristotle, or even Darwin, did. But that does not mean we should no longer use the concepts they used. However I agree with her statement that "the legitimacy of ascribing consciousness will be a function of what other mental predicates we want to ascribe, and the legitimacy of ascribing these is a function of the range of behaviours open to the agent, the range of stimuli to which he can react discriminatingly", though she adds, in my view mistakenly, "and perhaps above all of our attitude to him or it". This could get in the way of scientific advance, since attitudes change as we become better informed.

Wilkes suggests that whereas the ordinary notion of "intelligence" can be, and has been, transformed into a concept useful for science, that is not possible for "consciousness", since "no precise characterization could come close to capturing the thoroughly imprecise and heterogeneous everyday meaning that the term has in the vernacular." I suspect that that's an exaggerated worry, but accept the need for caution in linking any scientific advance concerning forms of information processing related to familiar varieties of consciousness to a widely used pre-scientific concept. However, all the scientifically informed proposals made so far for giving "consciousness" a precise definition seem to me to be based on unjustified selection of a subset of cases as forming defining cases. For example, no current well known neural theory of consciousness that I have encountered says anything about the forms of mathematical

consciousness I have been working on, inspired by Immanuel Kant as well as my own experience of doing mathematics.

Ned Block

Ned Block is well known for emphasising a contrast between *phenomenal consciousness* and *access consciousness*. Both are thought of as subjects for scientific study, but perceptual or sensory forms of phenomenal consciousness are typically precursors of access consciousness. That latter involves having a functional role in some aspect of cognition, e.g. perception, control of actions, suggesting or testing hypotheses, answering questions, etc. Some of what he writes leads some readers to suppose that access consciousness can be explained by physical mechanisms in brains whereas phenomenal consciousness has a more mysterious status, but he explicitly denies that, claiming (e.g. in personal correspondence) to be a physicalist, i.e. he believes that ultimately all the facts about consciousness can be explained by physical mechanisms in brains, and their relationships to sensors, effectors and objects in the environment. The difference is that [access consciousness](#), is causally involved in cognitive processes of various sorts (e.g. forming beliefs and controlling actions), whereas [phenomenal consciousness](#) is a more primitive form of consciousness that can exist prior to causation or modification of cognitive processes.

In a recent online two-part interview by Daniel Tippens, ([2015a](#)) ([2015b](#)), Block added a third category, a form of consciousness involving "...monitoring, some feedback and maybe some awareness of yourself". (I have not understood why that is not a sub-category of what he had previously called "access consciousness" -- this is just a symptom of the difficulty of pinning down exactly what is being referred to by the phenomenal/access distinction.

To a first approximation Block seems to think that in perceptual (or sensory) processes there is a primary state "phenomenal consciousness" which has sensory or perceptual contents that can only be identified with some difficulty in laboratory experiments. Influenced by [Nagel\(1974\)](#) he explains phenomenal consciousness as "what it's like to see or smell or hear, that internal experience that you get when you have a sensation or images in your mind". He claims that we share that with some other animals (e.g. other mammals) and that "it does not require language or much in the way of cognition -- maybe nothing in the way of cognition". However, it is very difficult (for me, and I think some other readers) to pin down exactly what he means by cognition and why some types of consciousness do not require cognition whereas others do. In part 2 of the 2015 interview he adds a third sub-type of consciousness, in which we are conscious of things. E.g. "We are conscious of our own thoughts. We can be conscious of our pains, of our perceptions. That involves some notion of monitoring, some feedback and maybe some awareness of yourself. So that is another notion. That's called monitoring consciousness or self-consciousness." Another idea is what I call access consciousness. And that's when you have an episode of phenomenal consciousness and it is available to your cognitive systems. So you can think about it. You can reason about it. S

An example discussion

On Wed 6th June 2017 I attended a one-day multidisciplinary workshop on Origins of Consciousness, organised by Jonathan Birch

<https://www.eventbrite.co.uk/e/origins-of-consciousness-tickets-33885705056>

The conference aimed to focus on these questions:

1. What can the neuroscience of consciousness tell us about its evolutionary history?
2. How many times has consciousness evolved? Is it a uniquely human phenomenon, or do we share it with other mammals, birds, reptiles, fish, or even invertebrates?
3. What is the evolutionary function of consciousness, and why did it first evolve?

The talks were all excellent and between them covered a lot of ground, but I felt there were important omissions, both conceptual and factual, including a lack of discussion of evolutionary stages relevant to understanding origins of known types of consciousness. That's partly because researchers tend to think of consciousness in terms of a small subset of examples, without investigating precursors of all those examples, and trying to characterise the branching space of possible alternative forms of consciousness. Compare a study of mechanics that considers only levers, pulleys, and gears.

A more challenging comparative investigation of evolved forms of information processing, including forms of consciousness, is the goal of the multi-branched, Turing-inspired, Meta-Morphogenesis (M-M) project, begun in 2011, since when it has inspired a collection of branching sub-projects, especially the study of evolved construction kits, mentioned below. The current top-level overview (revised from time to time) is here:

<http://www.cs.bham.ac.uk/research/projects/cogaff/misc/meta-morphogenesis.html>

<http://www.cs.bham.ac.uk/research/projects/cogaff/misc/meta-morphogenesis.pdf>

The Meta-Morphogenesis project

The project was triggered by an invitation from Barry Cooper to comment on Turing's 1952 paper "The chemical basis of morphogenesis" (among others) for Elsevier's (prize-winning) centenary volume on Turing. More information about that volume is here:

<http://www.cs.bham.ac.uk/~axs/amtbook>

<http://www.mathcomp.leeds.ac.uk/turing2012/give-page.php?300>

I conjectured that the answer to my question: *Why did Turing write that paper, and what might he have done if he had lived several more decades?* was that he had begun to investigate alternative forms of information processing that were used or could have been used in living organisms between the very simplest organisms, or proto-organisms, and those on our planet now. Insofar as those chemical mechanisms combine both continuous and discrete processes they go beyond forms of computation that are equivalent to what a Turing machine (among several others) can do. In particular, a synapse containing billions of

It soon became clear that that project included a vast collection of different sub-projects, including a project to investigate the variety of construction kits developed and used by evolution and its products, some of which are for creating various kinds of information-processing mechanism, increasingly required both during biological evolution of new designs and during development of an individual whose needs, capabilities, and opportunities become more varied and more complex over time.

New construction kits for creating new construction kits (meta-construction-kits??) are also products of evolution, individual development, and in some cases social/cultural processes in a group. Cross-species construction kits occur in some symbiotic relationships.

A partial survey of varieties of fundamental and evolved construction kits of many kinds is in this strand of the M-M project:

<http://www.cs.bham.ac.uk/research/projects/cogaff/misc/construction-kits.html>

<http://www.cs.bham.ac.uk/research/projects/cogaff/misc/construction-kits.pdf>

Since 2011, the project has branched out in various ways, including the study of forms of compositionality produced and used by biological evolution,

<http://www.cs.bham.ac.uk/research/projects/cogaff/misc/compositionality.html>

<http://www.cs.bham.ac.uk/research/projects/cogaff/misc/compositionality.pdf>

and the study of increasingly sophisticated forms of perception of, and reasoning about, spatial structures and processes, in many different species, including abilities to detect, reason about and make use of spatial impossibilities, illustrated (chaotically) here:

<http://www.cs.bham.ac.uk/research/projects/cogaff/misc/impossible.html>

<http://www.cs.bham.ac.uk/research/projects/cogaff/misc/impossible.pdf>

Some species develop meta-cognitive mechanisms and capabilities allowing their spatial reasoning processes to be attended to, compared, taught, debugged and creatively combined in new multi-component competences, including obtaining and consuming food, making clothes, building shelters, fighting conspecifics and many more (e.g. designing and building temples and pyramids). Massive current engineering projects require intricately organised cooperative use of vast numbers of competences, in vast numbers of cooperating individuals, for example the London Crossrail project (look for videos).

Relevance to consciousness

I hope to return later to discussion of how this relates to origins of consciousness.

REFERENCES

(To be expanded)

Heidi Appel and Rex Cocroft (2014)

<https://academicminute.org/2014/10/appel-cocroft/>

To see whether predator noises would affect plants, two University of Missouri researchers exposed one set of plants to a recording of caterpillars eating leaves, and kept another set of plants in silence. Later, when caterpillars fed on the plants, the set that had been exposed to the eating noises produced more of a caterpillar-repelling chemical.

<https://link.springer.com/article/10.1007/s00442-014-2995-6>

Plants respond to leaf vibrations caused by insect herbivore chewing

<http://www.dailymail.co.uk/sciencetech/article-2677858/Bad-news-vegetarians-Plants-hear-eaten.html>

Ned Block, (1995), On a confusion about the function of consciousness, *Behavioral and Brain Sciences*, 18, pp. 227--47,

Ned Block,(2015a) interviewed on phenomenal consciousness, Part I May 18, 2015

Daniel Tippens, asks Professor Ned Block, of New York University, about his work on the relationship between phenomenal consciousness and access consciousness. This is part I of that interview,

Ned Block (2015b), interviewed on phenomenal consciousness, Part II May 20, 2015 Daniel Tippens, asks Professor Ned Block, of New York University, about his work on the relationship between phenomenal consciousness and access consciousness. This is part II of that interview,

Paco Calvo, What Is It Like to Be a Plant?, *Journal of Consciousness Studies*, 2017, vol 24, 9-10, pp. 205-227,

<http://www.ingentaconnect.com/content/imp/jcs/2017/00000024/F0020009/art00012>

Hubert L. Dreyfus, (1979), *What Computers Can't Do*, Harper and Row, (revised ed.), New York,

P.M.S. Hacker, 2002, Is there anything it is like to be a bat?, *Philosophy*, Vol. 77, pp. 157--74,

<http://info.sjc.ox.ac.uk/scr/hacker/docs/To%20be%20a%20bat.pdf>

Thomas Nagel (1974), What is it like to be a bat, *The Philosophical Review* Vol. 83, No. 4 (Oct., 1974), pp. 435-450, <https://www.jstor.org/stable/2183914>; also in

The mind's I: Fantasies and Reflections on Self and Soul Eds D.R. Hofstadter and D.C.Dennett Penguin Books 1981, pp391--403 (Followed by commentary by D.R.Hofstadter, pp403--414.)

Arthur S. Reber, (2018) *The First Minds: Caterpillars, Karyotes, and Consciousness*, ISBN-13:9780190854157, Oxford Scholarship Online,

<http://dx.doi.org/10.1093/oso/9780190854157.001.0001>

Clarence A. Ryan and Andre Jagendorf, May 1995, Self defense by plants

Proc. Natl. Acad. Sci. USA Vol. 92, p. 4075, Colloquium entitled "Self-Defense by Plants: Induction and Signalling Pathways"

September 15-17, 1994, National Academy of Sciences, in Irvine, CA.

<http://www.pnas.org/content/92/10/4075.full.pdf>

Aaron Sloman, students and colleagues (1978,ff) The Birmingham Cognition and Affect project and its precursors at Sussex University.

<http://www.cs.bham.ac.uk/research/projects/cogaff/misc/emotions-affect.html>

<http://www.cs.bham.ac.uk/research/projects/cogaff/>

Aaron Sloman, (1996 ...). What is it like to be a rock? Online, semi-serious, discussion note.

<http://www.cs.bham.ac.uk/research/projects/cogaff/misc/rock/>

Kathleen V. Wilkes, (1984), Is Consciousness Important?, *British Journal for the Philosophy of Science*, Vol 35, No 3, Sep, 1984 pp. 223--243, <https://doi.org/10.1093/bjps/35.3.223>

Document History

The first draft of this document was installed on 8th June 2017. In 2019, after attending a conference on consciousness in Oxford at the Mathematical Institute (Sept 9-12,2019), I decided to re-write this to expand a point I made about the importance of consideration of evolution and changing functions of biological consciousness during the discussion. The conference web site is:

<https://models-of-consciousness.org/>

The conference booklet, including talk abstracts (3.1MB)is:

https://models-of-consciousness.org/booklet/MoC19_Conference_Booklet.pdf

Oxford Mathematics of Consciousness and Applications Network:

<https://omcan.web.ox.ac.uk/> **Last updated:** 19 Sep 2019

Maintained by [Aaron Sloman](#)
[School of Computer Science](#)
[The University of Birmingham](#)