QUALIA AIN'T OUT HERE EITHER

Christopher Searle

What follows is a short thought experiment that aims to show, reductio ad absurdum*, that narrow qualia internalism is probably inconsistent with a physicalist/functionalist theory of mind. Those wishing to rebut the argument presented here will need to demonstrate why spatial proximity and the right sort of causal connection of functionally isolated components are necessary to the instantiation of qualia.*

Let us imagine a typical brain-in-a-vat scenario where an evil scientist has wired up a super-scientific computer to a fully functioning human brain.

As per usual, the sensory inputs are provided by the computer so that the experiencing subject does not believe it is actually a brain-in-a-vat, rather it believes it is a normal human being engaged in some physical activity or other.

For the sake of argument, let us imagine that the evil scientist has engineered the situation in such a way that our brain-in-a-vat experiences exactly the same five-minute sequence over and over again indefinitely.

Using super-advanced nanotechnology, the evil scientist is able automatically to return the brain to its previous physical state at the end of each five-minute sequence. Every single neural connection is returned to its previous configuration. The inputs, the outputs, and the overall function of the system are precisely the same for each five-minute cycle, so phenomenologically speaking, according to identity theory, the perception of the brain-in-a-vat is the same for each cycle.

doi:10.1017/S1477175619000368 © The Royal Institute of Philosophy, 2019
Think 54, Vol. 19 (Spring 2020)

https://doi.org/10.1017/S1477175619000368 Published online by Cambridge University Press

So, *ceteris paribus*, each consecutive five-minute cycle is an intrinsic duplicate of its predecessor with precisely the same internal qualia.

Let us consider a simplistic view of the operation of a single neuron that is intrinsically involved in the instantiation of these qualia.

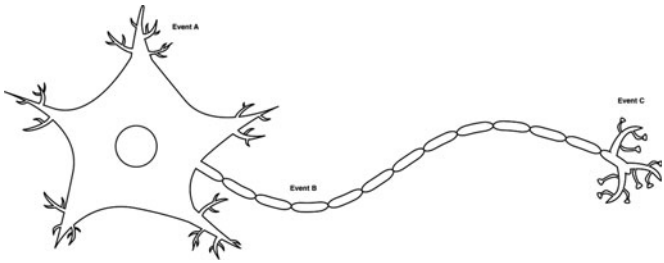


Figure 1. Neurological event A–B–C.

So, with the original brain-in-a-vat, our single neuron undergoes an action potential. A synaptic signal is received as neurotransmitters stimulate the dendrites; this is neurological event A. The sodium channels open, the neuron fires, and an electrical signal is sent down the length of the axon; this is neurological event B. The signal reaches synaptic knobs of the axon terminals and a neurotransmitter is released; this is neurological event C. Neurological event A leads to neurological event B which in turn leads to neurological event C. See Figure 1.

Let us imagine that our brain-in-a-vat is being fed a five-minute loop of the experience of playing croquet on the Quad lawn. It becomes conscious, but it is unaware that it has only just become conscious because its consciousness, of course, contains all of the associated memories of its immediate and distant past. It simply experiences playing the same five-minutes of croquet over and over again, blissfully unaware of its predicament.

Now, using super advanced Star-Trek-style transporter technology, our evil scientist is able to duplicate the entire

https://doi.org/10.1017/S1477175619000368 Published online by Cambridge University Press

brain-in-a-vat system. The duplicate that he creates has precisely the same structure as the original; every neural connection is the same, and therefore it would seem to have qualitatively identical qualia.

The evil scientist then uses his advanced nanotechnology to divide the duplicate brain into large removable sections in the following way: every single neuron that crosses the border of the section to be divided is severed across its axon. A super-scientific nano-computer is then carefully installed at the termination of each side of the divided axon. These tiny computers can communicate via a super-advanced signal device. Of course, billions of these devices are required, but here we shall focus on just one.

Splitting the neurons of the duplicate brain into sections in this way allows our evil scientist to move the various sections of the brain to different locations around his laboratory. The occipital lobe sits in a jar next to his High-Voltage Jacob's Ladder, while the temporal lobe sits in a jar by his crackling Tesla Coil, and all the while the brain sections are communicating happily together and the perceiving subject perceives no change in his perception.

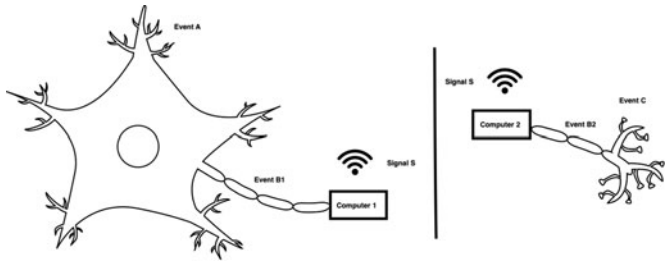


Figure 2. Neurological event $A–B_1–S–B_2–C$.

With the duplicate brain-in-a-vat, neurological event A leads to neurological event B_1 which causes nano-computer 1 to initiate signal S. On detection of signal S, nano-computer 2 initiates neurological event B_2, which leads to neurological event C. See Figure 2.

https://doi.org/10.1017/S1477175619000368 Published online by Cambridge University Press

This signal is of course instantaneous, so dividing the neurons in this way has no effect at all on the global functionality of the brain.

Neurological event A–B–C has precisely the same function as neurological event A–B_1–S–B_2–C, and precisely the same parts of the neuron are active in each case.

So, according to internalist intentionalism, unless the proximity of neurons is required, the duplicated and dissected brain will have the same qualia as the functionally identical original. The mental states of each brain have the same content and the same qualia. Both perceiving subjects experience playing a game of croquet on the Quad lawn. Yet one is a brain-in-a-vat, the other – a brain in several.

Intuitively it would seem that the spatial proximity between various parts of the brain has no bearing on our willingness to attribute qualia to neurological processes.

So far so good, for narrow qualia internalism.

Now let us imagine that our evil scientist goes one step further. He is able to programme the super-scientific nanocomputers so that the communication between them is no longer required.

Suppose that for the original brain-in-a-vat, neurological event A occurs at time t_1, event B occurs at time t_2, and event C occurs at time t_3.

In the duplicate brain-in-a-vat, neurological event A occurs at time t_1, event B_1 occurs at time t_2, and event B_2

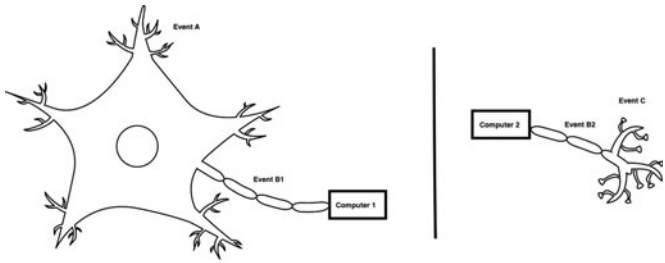


Figure 3. Neurological events A–B1 and B2–C.

https://doi.org/10.1017/S1477175619000368 Published online by Cambridge University Press

also occurs at time t_2 so that event C occurs at time t_3. See Figure 3.

There is no communication between Computer 1 and Computer 2, it is just that Computer 2 has an internal clock and is carefully programmed to initiate neurological event B_2 immediately after the time at which neurological event B_1 occurs in every five-minute cycle.

There is a pre-established harmony between the inputs and outputs, and the overall function of each neuron is the same between the duplicate and the original, even though in the former the neuron is physically divided, spatially separated, and the two halves do not communicate and are causally isolated.

So now our mad scientist has various different brain sections stored in different jars at different locations around his laboratory, none of which are communicating in any way, there is no causal connection between them, but together they are supposed somehow to instantiate the qualia associated with the experience of playing croquet on the Quad lawn. Do the qualia exist in the lab as a whole instead of in the head, or in the vat? If we are unwilling to attribute qualia to this system, why?

Qualia clearly ain't in the head, but they ain't out here either.

Post Scriptum

There are clear parallels between the argument presented here and the central thesis of Dr Zuboff's 'The Story of Brain' (1981) in D. Hofstadter, and D. Dennett (eds.), The Mind's I *(New York: Basic Books), 202–12. These parallels are purely coincidental. Up until the time of submission, the work of Dr Arnold Zuboff was unknown to the author, but the reader is now advised to turn to the work of Dr Zuboff for a more thorough exploration of this issue.*

Christopher Searle is a part-time student reading Philosophy at the University of Oxford. christopher.searle@conted.ox.ac.uk

https://doi.org/10.1017/S1477175619000368 Published online by Cambridge University Press