Yet Another Objection to Fading and Dancing Qualia (5th draft)

Nir Aides
nir@winpdb.org

In this paper I present objections to the Fading Qualia and Dancing Qualia thought experiments, which David Chalmers uses to argue that functional organization fully determines conscious experience.

Chalmers describes these thought experiments, and the principle of organizational invariance in the paper: “Absent Qualia, Fading Qualia, Dancing Qualia” (1995; 1996) [1]; he believes that consciousness is irreducible, and the thought experiments are designed to show that irreducibility is not a barrier to the construction of a conscious computer (1996, p. 275), and more generally that the right kind of Turing equivalent computation will possess a mind (1993).

Chalmers claims that while his arguments are only plausibility arguments, they have considerable force (1996, p. 250) and their conclusion is by far the most plausible; my intention is to show this is not true.

Chalmers begins by defining a notion of functional organization, and what it means for two systems to be functionally isomorphic; in particular, according to Chalmers, functionally isomorphic systems have the same functional organization at some level which is “fine enough” to determine behavioral capacities and dispositions, and he suggests that perhaps the neural level is fine enough for his purpose (1996, p. 248).

In the two thought experiments Chalmers supposes a system called Robot, which is functionally isomorphic to his brain, by having a silicon chip for every neuron that he has, and contemplates what it would be like to be an intermediate system, a mixture of neurons and silicon chips.

An empirical objection to fading qualia

In the fading qualia thought experiment Chalmers describes a process of replacing his neurons one by one with functionally isomorphic silicon chips, gradually transforming his brain into Robot, while preserving its functional organization; in his mind the only way Robot may have no qualia is if it suddenly disappears or if it gradually fades away during the process.

Chalmers argues that in each step, replacing a single neuron with a silicon chip cannot result in the sudden disappearance of qualia, since “If this were possible, there would be brute discontinuities in the laws of nature unlike those we find anywhere else” (1996, p. 255), and since “In all fundamental laws known to date, the dependence of one magnitude on another continuous magnitude is continuous in this fashion, and there is no way to compound continuity into discontinuity” (1996, p. 256).

These claims seem false; according to the photoelectric effect described by Einstein, a photon above a threshold frequency has the required energy to detach an electron from metal (1921), and since the electromagnetic spectrum is continuous, and the ejection of an electron is a discrete event, we arguably have an example of “compounding continuity into discontinuity”; to keep things down to earth with a simpler example, one can imagine a delicate chinese vase hanging from a rope, and cutting the rope with a knife, one fibre, or even one molecule, at a time; at some point, the rope will fail and the vase will fall down to earth and crash; since that can happen at an arbitrary time between two successive cuts, we can speak of the vase as disappearing.

As a note, one may also ask why a physical principle may be applied to qualia; according to Chalmers conscious experience is not predictable from physics: “If all we knew about were the facts of physics, and even the facts about dynamics and information processing in complex systems, there would be no compelling reason to postulate the existence of conscious experience” (1996, pp. 4-5); on these grounds it is not clear what the laws of physics can tell us about the emergence, or the disappearance of qualia as neurons are replaced with silicon chips.

Chalmers then argues by *reductio ad absurdum* that fading qualia is impossible as well, but for that purpose he contrasts fading qualia with suddenly disappearing qualia, and writes: “Perhaps in the extreme case, when all is dark inside, it is reasonable to suppose that a system could be so misguided in its claims and judgments … But in the intermediate case, this is much less plausible” (1996, p. 257).

But if suddenly disappearing qualia are not ruled out, and if the *reductio ad absurdum* does not apply to suddenly disappearing qualia, then the thought experiment is unconvincing.

An indeterministic brain objection to dancing qualia

In the dancing qualia thought experiment Chalmers argues by *reductio ad absurdum* that if Robot has inverted qualia or is missing qualia altogether, then there should exist two intermediate systems, between him and Robot, which have significantly different qualia but are different by, say, no more than one tenth of their composition (1996, pp. 266-267); to get from one system to the other all we would have to do is replace a small region of Chalmers' brain with a functionally isomorphic silicon circuit.

Chalmers then describes a process of switching back and forth between that small region of his brain and the functionally isomorphic silicon circuit, by flipping a switch, and he claims that since functional organization is preserved throughout the process, there is no way for the system to notice it (1996, p. 269) [2].

But suppose the brain is indeterministic [3], and that random evolution of brain states may take place on the same time scale of perceived noisy phenomena such as closed eyes noise, tinnitus, or split second decision making.

If so, then the small region of Chalmers' brain and its isomorphic silicon circuit may, in principle, quickly evolve into different states and generate different outputs.

I believe that even slightly different outputs may be enough for the switching between them to open the door for "noticing", and that therefore Chalmers' argument loses its force.

To see why, suppose a set of neurons *A*, and its isomorphic duplicate *B*; the rest of the brain "sees" a switching between *A* and *B*; that switching can be viewed as a new function which we can call *ABAB*; if *A* and *B* evolve into different states and generate different outputs, then it may be arguably possible for *ABAB* to generate outputs which *A* cannot generate, and therefore *ABAB* is not necessarily isomorphic to *A*.

In analogy, suppose a nondeterministic algorithm *F* that takes an input state and outputs a video of algorithmic art, and suppose that *F* is run twice on the same input state *S*, and generates two slightly different output videos *V1* and *V2*.

These videos may appear indistinguishable to an observer (the rest of the brain); however, if we create a third video *V1212* by alternating between frames of *V1* and *V2*, an observer may easily notice there is something "wrong" with *V1212*. Both *V1* and *V2* are valid outputs of *F(S)* but *V1212* is generally not.

Interestingly, on introspection, it seems that closed eyes noise and tinnitus can produce noise which seems like *ABAB* or *V1212* kind of output, and that therefore it may be argued that such output will not necessarily get noticed as "wrong"; nevertheless, it doesn't seem reasonable to rule out such noticing in principle as Chalmers does.

If noticing is not ruled out, then the thought experiment is unconvincing.

Notes

[1]: There are some differences between the online paper (1995) and the book chapter (1996).

[2]: "Even as we flip the switch a number of times and my qualia dance back and forth, I will simply go about my business, noticing nothing unusual. By hypothesis, my functional organization remains normal throughout. In particular, my functional organization after flipping the switch evolves just as it would have if the switch had not been flipped. There is no special difference in my behavioral dispositions. I am not suddenly disposed to say 'Hmm! Something strange is going on!' There is no room for a sudden start, for an exclamation, or even for a distraction of attention. Any unusual reaction would imply a functional difference between the two circuits, contrary to their stipulated isomorphism. By design, my cognitive organization is just as it usually is, and in particular is precisely as it would have been had the switch not been flipped" (1996, pp. 268-269)

[3]: Whether the brain is deterministic or not is in an open question; according to Glimcher (2005) behavioral science has been traditionally anchored in a deterministic world view, but there are scientists who are now questioning that assumption; one example is growing interest in stochastic resonance in the brain

(2009); while SR can be reproduced with deterministic chaos, it is typically modeled and analyzed in terms of random noise; Brascamp (2006) reports that neural noise has a role in the timing of binocular rivalry.

References

Arrhénius, S. A. (1921). Nobel Prize in Physics - Presentation Speech. Retrieved May 27, 2015, from http://www.nobelprize.org/nobel_prizes/physics/laureates/1921/press.html

Chalmers, D. (1993). A Computational Foundation for the Study of Cognition. Retrieved May 15, 2015, from http://cogprints.org/319/

Chalmers, D. (1995). Absent Qualia, Fading Qualia, Dancing Qualia. Retrieved May 15, 2015, from http://cogprints.org/318/

Chalmers, D. (1996). Absent Qualia, Fading Qualia, Dancing Qualia. In The conscious mind: In search of a fundamental theory (pp. 247-275). New York: Oxford University Press.

Glimcher, P. W. (2005). Indeterminacy in brain and behavior. Psychology, 56(1), 25.

Brascamp, J. W., Van Ee, R., Noest, A. J., Jacobs, R. H., & van den Berg, A. V. (2006). The time course of binocular rivalry reveals a fundamental role of noise. Journal of vision, 6(11), 8.

McDonnell, M. D., & Abbott, D. (2009). What is stochastic resonance? Definitions, misconceptions, debates, and its relevance to biology. PLoS computational biology, 5(5), e1000348.