

FACING BACKWARDS ON THE PROBLEM OF CONSCIOUSNESS

*Daniel C. Dennett, Center for Cognitive Studies,
Tufts University, Medford, MA 02155, USA.*

The strategy of divide and conquer is usually an excellent one, but it all depends on how you do the carving. Chalmers's (1995) attempt to sort the 'easy' problems of consciousness from the 'really hard' problem is not, I think, a useful contribution to research, but a major misdirector of attention, an illusion-generator. How could this be? Let me describe two somewhat similar strategic proposals, and compare them to Chalmers' recommendation.

1. The hard question for vitalism

Imagine some vitalist who says to the molecular biologists:

The easy problems of life include those of explaining the following phenomena: reproduction, development, growth, metabolism, self-repair, immunological self-defence . . . These are not all *that* easy, of course, and it may take another century or so to work out the fine points, but they are easy compared to the really hard problem: life itself. We can imagine something that was capable of reproduction, development, growth, metabolism, self-repair and immunological self-defence, but that wasn't, you know, *alive*. The residual mystery of life would be untouched by solutions to all the easy problems. In fact, when I read your accounts of life, I am left feeling like the victim of a bait-and-switch.

This imaginary vitalist just doesn't see how the solution to all the easy problems amounts to a solution to the imagined hard problem. Somehow this vitalist has got under the impression that being alive is something over and above all these subsidiary component phenomena. I don't know what we can do about such a person beyond just patiently saying: your exercise in imagination has misfired; you can't imagine what you say you can, and just saying you can doesn't cut any ice. (Dennett, 1991, p. 281–2.)

2. The hard question for Crick

Francis Crick (1994) gives us an example of what happens when you adopt Chalmers' distinction, when he says, at the close of his book on consciousness. 'I have said almost nothing about qualia — the redness of red — except to brush it to one side and hope for the best.' (p. 256.) But consider what would be wrong with the following claim made by an imaginary neuroscientist (Crick) substituting 'perception' for 'qualia' in the quotation from Crick: 'I have said almost nothing about perception — the actual analysis and comprehension of the visual input — except to brush it to one side and hope for the best.' Today we can all recognize that whatever came before Crick's declaration would be forlorn, because not so many years ago this was a mistake that brain scientists actually made: they succumbed all too often to the temptation to treat vision as if it were television — as if it were simply a matter of getting 'the picture' from the eyes to the screen somewhere in the middle where it could be handsomely reproduced so that the phenomena of appreciation and analysis could then get underway. Today we realize that the analysis — the whatever you want to call it that composes, in the end, all the visual understanding — begins right away, on the retina; if you postpone consideration of it, you misdescribe how vision works. Crick has made a mistake: he has created an artifactual

‘hard’ problem of perception, not noticing that it evaporates when the piecemeal work on the easy problems is completed.

Is it similarly a mistake for Crick, following Chalmers, to think that he can make progress on the easy questions of consciousness *without in the process answering the hard question*? I think so (Dennett, 1991). I make the parallel claim about the purported ‘subjective qualities’ or ‘qualia’ of experience: if you don’t begin breaking them down into their (functional) components from the outset, and distributing them throughout your model, you create a monster — an imaginary dazzle in the eye of a Cartesian homunculus (Dennett, 1995).

Chalmers has not yet fallen in either of these traps — not quite. He understands that he must show how his strategic proposal differs from these, which he recognizes as doomed. He attempts this by claiming that consciousness is strikingly unlike life, and unlike the features of perception misconstrued by Crick: when it comes to consciousness, the hard problem is ‘almost unique’ in that it ‘goes *beyond* problems about the performance of functions.’ *Almost unique*? He gives us no other phenomena with this special feature, but in any case, what he says in support of this claim simply repeats the claim in different words:

To see this, note that when we have explained the performance of all the cognitive and behavioural functions in the vicinity of experience . . . there may still remain a further unanswered question: *Why is the performance of these functions accompanied by experience?* A simple explanation of the functions leaves this question open. (Chalmers, 1995, p. 203.)

Our vitalist can surely ask the same dreary question: Why is the performance of these functions accompanied by life? Chalmers says that this would be a conceptual mistake on the part of the vitalist, and I agree, but he needs to defend his claim that his counterpart is not a conceptual mistake as well.

When he confronts the vitalist parallel head-on, he simply declares that whereas vitalist scepticism was driven by doubts about whether physical mechanisms could ‘perform the many remarkable functions associated with life’, it is otherwise with his scepticism:

With experience, on the other hand, physical explanation of the functions is not in question. The key is instead the *conceptual* point that the explanation of functions does not suffice for the explanation of experience. (p. 209.)

I submit that he is flatly mistaken in this claim. Whether people realize it or not, it is precisely the ‘remarkable functions associated with’ consciousness that drive them to wonder about how consciousness could possibly reside in a brain. In fact, if you carefully *dissociate* all these remarkable functions from consciousness — *in your own, first-person case* — there is nothing left for you to wonder about.

What impresses *me* about my own consciousness, as I know it so intimately, is my delight in some features and dismay over others, my distraction and concentration, my unnamable sinking feelings of foreboding and my blithe disregard of some perceptual details, my obsessions and oversights, my ability to conjure up fantasies, my inability to hold more than a few items in consciousness at a time, my ability to be moved to tears by a vivid recollection of the death of a loved one, my inability to catch myself in the act of framing the words I sometimes say to myself, and so forth. These are *all* ‘merely’ the ‘performance of functions’ or the manifestation of various complex dispositions to perform functions. In the course of making an introspective catalogue of evidence, I

wouldn't know what I was thinking about if I couldn't identify them for myself by these functional differentia. Subtract them away, and nothing is left beyond a weird conviction (in some people) that there is some ineffable residue of 'qualitative content' bereft of all powers to move us, delight us, annoy us, remind us of anything.

Chalmers recommends a parallel with physics, but it backfires. He suggests that a theory of consciousness should 'take experience itself as a fundamental feature of the world, alongside mass, charge, and space-time.' As he correctly notes, 'No attempt is made [by physicists] to explain these features in terms of anything simpler,' but they do cite the independent evidence that has driven them to *introduce* these fundamental categories. Chalmers needs a similar argument in support of his proposal, but

when we ask what data are driving him to introduce this concept, the answer is disappointing: It is a belief in a fundamental phenomenon of 'experience'. The introduction of the concept does not do any explanatory work. The evidential argument is circular. (Roberts, 1995, fn 8.)

We can see this by comparing Chalmers' proposal with yet one more imaginary non-starter: *cutism*, the proposal that since some things are just plain cute, and other things aren't cute at all — you can just see it, however hard it is to describe or explain — we had better postulate *cuteness* as a fundamental property of physics alongside mass, charge and space-time. (Cuteness is *not* a functional property, of course; I can imagine somebody who wasn't actually cute at all but who nevertheless functioned exactly as if cute — trust me.) Cutism is in even worse shape than vitalism. Nobody would have taken vitalism seriously for a minute if the vitalists hadn't had a set of independently describable phenomena — of reproduction, metabolism, self-repair and the like — that their postulated fundamental life-element was hoped to account for. Once these phenomena were otherwise accounted for, vitalism fell flat, but at least it had a project. Until Chalmers gives us an *independent* ground for contemplating the drastic move of adding 'experience' to mass, charge, and space-time, his proposal is one that can be put on the back burner.

References

- Chalmers, David (1995), 'Facing up to the problem of consciousness', *Journal of Consciousness Studies*, 2 (3), pp. 200–19.
- Crick, Francis (1994), *The Astonishing Hypothesis: The Scientific Search for the Soul* (New York: Scribners).
- Dennett, Daniel (1991), *Consciousness Explained* (Boston, MA: Little, Brown & Co.).
- Dennett, Daniel (1995), 'Our vegetative soul' — review of Damasio, *Descartes' Error*, in *Times Literary Supplement*, August 25, 1995, pp. 3–4.
- Roberts, John (1995), 'Qualia and animal consciousness', Center for Cognitive Studies working paper, 1995-#1, Tufts University (unpublished).