

HOW VELMANS' CONSCIOUS EXPERIENCES AFFECTED OUR BRAINS

Ron Chrisley and Aaron Sloman

Velmans' paper raises three problems concerning mental causation:

- (1) How can consciousness affect the physical, given that the physical world appears causally closed?¹⁰
- (2) How can one be in conscious control of processes of which one is not consciously aware?
- (3) Conscious experiences appear to come too late to causally affect the processes to which they most obviously relate.

In an appendix Velmans gives his reasons for refusing to resolve these problems through adopting the position (which he labels 'physicalism') that 'consciousness is nothing more than a state of the brain'. The rest of the paper, then, is an attempt to solve these problems without embracing a reductionist physicalism.

Velmans' solution to the first problem is 'ontological monism combined with epistemological dualism': First-person and third-person accounts are two different ways of knowing the same facts. This kind of reply is not new; it is, for example, a twist on the position expressed in Davidson (1970). True, there are substantial differences: For one, Davidson reconciles the tension between descriptions of events in mentalistic and physicalist language, not between first- and third-person descriptions of states; for another, Davidson actually provides an argument for his position, although to do so he assumes that there are no psycho-physical (or indeed, psycho-psycho) laws, something which I suspect Velmans would be reluctant to do. Nevertheless, they have in common the idea that the causal efficacy of the mental is not at odds with the causal closure of physics, since a mind-involving causal story is just another way of talking about the same facts that a purely physical causal story talks about.

This 'dual-aspect' approach is a popular tactic for resolving the mind-body problem, but it has some well-known problems, and it is unfortunate Velmans doesn't reply to these standard objections. For example, a frequently discussed issue in connection with theories of mental causation is the problem of overdetermination (see, e.g., Unger, 1977; Peacocke, 1979). (Although it is usually stated in terms of a contrast between the mental and the physical, I'll translate it into 'consciousness speak'.) It would seem that any account that admits the causal closure of the physical, yet introduces the causal efficacy of conscious states, ends up in the unfortunate situation of having too many causes lying around. If physical event P causes action A, and also the event of having a conscious experience C causes action A, then even if the physical world had been

[10] Actually, that's a big 'given': Although at least one of the present authors favours a Bohmian interpretation of quantum mechanics (Bohm, 1952), which is compatible with determinism, we both acknowledge that physical reality could very well be as orthodoxy says it is: non-deterministic. If so, there are indeed "gaps" in the chain of causation that consciousness might fill'. One might doubt that such gaps could be present or significant in warm, massive brains, but there are several authors who argue otherwise (e.g., Beck, 1996; Beck and Eccles, 1992; Hameroff and Penrose, 1996; Hagan *et al.*, 2002).

otherwise (P' rather than P had occurred, say), it seems that A would still have occurred, since C would still have been there to do the causing. And conversely, if C had not occurred, it looks as if A would have still occurred, since P is still in place to do the causing.

A dual-aspect theory, it is said, allows one to resolve this problem. Since the physical and experiential perspectives are different ways of describing the same underlying reality, one cannot assume that if P does not occur, C still occurs. And, assuming supervenience of the experiential on the physical, one *knows* that if C had not occurred, P definitely would not have occurred. Thus, on an 'ontologically monist but epistemologically dualist' account, there is no problem of causal over-determination of effect.

But many have questioned whether this answer is satisfying (e.g., Sosa, 1984; Block, 1989; LePore and Loewer, 1989; Leiter and Miller, 1994): We'd like to think that our conscious states have causal power by virtue of their being the mental states that they are, not by virtue of being identical with some physical state, which itself has, by virtue of falling under physical laws, the true causal power. Simplistic appeals to a 'neutral' reality, which underwrites both physical and experiential causation talk, will not work here. There is a fundamental asymmetry between the physical and the conscious: Physical laws apply everywhere, both in situations where there is and where there is not consciousness, while the converse does not hold.¹¹ So there seems to be a primacy of the physical, and one must reply to the idea that it is this physical, causal reality which is always doing all the work.

This is only to mention one issue which must be addressed by any proposal such as Velmans'; there are others. For example, Honderich (1993) would seem to argue that Velmans' position cannot give a proper account of mental causation, while Kim (1993) explicitly argues that non-reductive approaches that try to do justice to mental causation end up violating the causal closure of the physical. By mentioning these examples, we are not saying that dual-aspect theories cannot be defended against them, nor indeed that such theories are not good contenders for a proper account of the relation between consciousness and the physical; but Velmans has not given such a defence, and his proposal would benefit from his locating it within the discussion that has already occurred in this area.¹²

One particular benefit of such contextualisation would be the clarification of Velmans' position itself. Although the above discussion takes the 'ontological monism combined with epistemological dualism' slogan at its word, there are several passages in Velmans' text which are in tension with that position. For example, the 'epistemological dualism' part of the slogan is supposed to rule out the reductionist physicalism rejected in the appendix. But Velmans seems to be assuming a reductionist position himself when he discusses the neural correlates

[11] This asymmetry implies that there is more work to be done than Velmans, even in footnote 14, acknowledges. Not only will there have to be laws relating the first- and third-person characterisations of a psychophysical state, but there will also have to be laws which tell us the conditions under which a state that has a physical/third-person characterisation also has an experiential/first-person characterisation in the first place.

[12] Some overviews of the positions in the mental causation debate can be found in Crane (1995) and Jackson (1996).

of consciousness in the light of his theory. When Velmans moves, without argument, from the representational nature of experiences to the existence of neural correlates of these experiences which have the same representational content as these experiences, he is either making a rather strong reductionist assumption, or (worse) postulating a dubious causal connection (between the structure of experience and the structure of neural states) that needs to be explained.

A general point can be made here. When people consider the proposal that computation or artificial intelligence (AI) can help us understand the mind, they often assume that this would only be true if the mind were in some sense computational. But this is to ignore a different way in which work in AI can be of assistance: by being a test-bed for our metaphysical theories. If one interacts with virtual machines implemented in computational hardware, one can come to realize that the mirroring of structure that Velmans is assuming need not hold. That is, computational examples make it clear that although level of description (aspect) Y is implemented or realized in (is an aspect of the same thing as) level of description (aspect) X, it does not follow that for every entity, structure or property referred to in Y there is something referred to in X to which it corresponds. For example, a computer can be understood as computing with sparse arrays, even though for any particular cell of the array there will likely be nothing localizable in in computer hardware to which that cell corresponds (Sloman, 2001). It is through designing, implementing and/or interacting with computational systems realizing multiple levels of virtual machines that one's comprehension of the metaphysical possibilities is expanded. So, no, it does not follow that there must be neural correlates of consciousness (in Velmans' strong sense of the term), just as there are not, in general, silicon correlates of computation. (See, e.g., Hurley, 1998, for a quite different reason for believing that the structure of consciousness need not be mirrored in the structure of the vehicles of consciousness.) In assuming that there must be such correlates, it is hard to see how Velmans is less reductionist than (some of) the positions that he argues against in the appendix.

Perhaps sensing that he is falling into reductionism, Velmans uses three analogies to attempt to convince us that all is well, but they in fact make matters worse. Sameness of information structure does not mean that experiences are nothing more than physical states, Velmans points out. A video recording of a TV broadcast of Hamlet, Velmans says, has the same 'sequential informational structure' as the screen of the TV receiving that programme, and yet the videotape and screen states are 'not ontologically identical'.¹³ Even if we assume that Velmans is correct in this claim, his attempt to make a point with it misfires, for two reasons. First, the problem was not that we had as data sameness of information structure, and this seemed to force us into monism. Rather, it seemed that Velmans could only move from the existence of informationally structured

[13] Surely Velmans does not need to claim that the experience and its correlate have the *same* informational structure; doubtlessly (and in the light of his footnote 9) the physical aspect of the system will typically encode much more information than is experientially represented. Thus Velmans may say (and *should* say, in order to reflect the asymmetry between the experiential and the physical) that the physical aspect must contain at least as much as informational structure as the experiential aspect. However, this proviso will not on its own answer the other objections we are making.

experiences to identically informationally structured physical states by assuming some kind of systematic, law-like relations between experience and physical reality. Perhaps this is not an ontological reduction, but it is an epistemological one; yet epistemological dualism is the only thing separating Velmans from the physicalist positions he rejects. It is this need to distance himself from physicalism which raises the second problem with the analogy: he admits that the videotape and the screen are ontologically distinct, yet he was supposedly defending an ‘ontologically monist’ position! It seems Velmans ends up with the converse of the position for which he was aiming: ontological dualism, but epistemological monism (in the sense that strong assumptions are made about ‘informational mirroring’).

This is not a problem which can be disposed of by simply deleting the Hamlet TV programme/video analogy, but is rather a deep tension in Velmans’ position which surfaces at several points. Consider another analogy which Velmans offers to illustrate his view of the ontological relation between experience and the physical: that of electricity and magnetism. Velmans observes, in footnote 14: ‘it does not make sense to suggest that the current in the wire is nothing more than the surrounding magnetic field, or vice versa (reductionism)’. But then Velmans wants to also have it that the duality implied by this observation is one of aspects, not of ontological character. This is meant to be analogous to the relation between the experiential and the physical. However, the analogy doesn’t work: electricity and magnetism are not simply two ways of thinking about the same phenomenon, but two different physical phenomena that can be related to each other mathematically. In contrast, and crucially, Velmans claims that the difference between first- and third-person ways of thinking of psychophysical stuff is merely that of differently formatted ways of representing the same information. This is not what is happening in the case of electromagnetism: the electrical phenomenon is not just an aspect, a way of formatting the same information as that represented by the magnetic way of looking at the situation. There are situations where only the electrical description applies, and other situations where only the magnetic description applies. *Prima facie*, this suggests that there are two distinct phenomena involved; to argue that there is actually only one, root phenomenon will require further work from Velmans. The analogy is also spoiled by the symmetry of the electrical/magnetic relation, and the asymmetry (discussed above) of the consciousness/physical relation. It may be disputed which of the following is the case:

- Electrical phenomena can exist in the absence of magnetic phenomena, and vice versa; or
- Whenever an electrical phenomenon exists, there also exists a corresponding magnetic phenomenon, and vice versa.

However, in both cases, the ‘vice versa’ implies an ontological symmetry which is not shared by the experiential/physical relation. The only way to impose symmetry would be to assume (as others have been forced to do, e.g. Chalmers, 1996) that whenever there is a physical phenomenon, there is some experiential phenomenon, however slight or imperceptible or implausible, accompanying it. Panpsychism threatens.

The third analogy, that of wave/particle complementarity, is even worse. More and more physicists and philosophers take the appeal to complementarity as a *reductio ad absurdum* of particular ontological positions in quantum mechanics. They do not deny the veracity of the data that have led some to conclude that quanta have both wave and particle aspects; but they do deny that the paradox of complementarity is a satisfying way of accounting for that data. There are other, less paradoxical and thus more satisfying metaphysical pictures on offer (e.g. Bohm, 1952; Hiley and Pylykänen, 2001). To say that your metaphysics of mind is akin to the wave/particle complementarity metaphysics of quanta is just another way of saying that you don't have a satisfying metaphysics, and choose instead to 'live with' the paradoxes.

Velmans' proposed solutions to the other two problems are relatively independent of his proposed solution to the first; we now consider them in reverse order. Concerning the third problem, Velmans concedes that in many cases (e.g., those documented in Libet, 1985), conscious experiences do come too late to causally affect the processes to which they most obviously relate (although they may have longer-term causal effects). Some might think that this would have the unpalatable consequence that many of our actions are involuntary, but Velmans' solution to the second problem shows why this is not so: non-conscious processes can nevertheless produce voluntary action. That seems very sensible, but there are other difficulties with Velmans' proposed solution to the timing problem. Specifically, it concedes too much: if we can be wrong about conscious experience playing a causal role in our decision to press a button, why shouldn't we be sceptical about its role in all action? Epiphenomenalism threatens.

A better response would be to resist the conclusion that conscious experience is not playing a causal role in the Libet cases. One could do this in two different ways. One could deny that subjects are infallibly accurate about the timing of their experiences. The fact that the subject takes the experience to be happening when the revolving dot is at a particular location on the screen does not imply that the experience is in fact occurring at that the time when the dot is actually at that location — it might be happening about 350 milliseconds earlier than that. Alternatively, one could make the obvious point that conscious experience can play a causal role, even if preceded by a predicting readiness potential. For example, it might be that the readiness potential causes the experience, which itself causes the action. Any dual-aspect theory worth its salt will not be troubled by the fact that there is a neural causal account which explains our action; that is entirely consistent with there also being an experience-involving account.

We agree with Velmans that there are philosophical problems concerning the causal efficacy of the experiential which need to be addressed by any proper theory of consciousness. We also agree that some sort of monist metaphysics, such as is required to explain the relation between virtual machines (in computers, say) and the physical machines in which they are implemented, is required. Despite Velmans' efforts, however, these needs remain unsatisfied. We believe that the clinical, psychological and philosophical methodologies Velmans musters should be supplemented with and informed by experimental, synthetic AI

work, in order to facilitate the acquisition of new concepts and refinement of old concepts that are required for advances in our understanding of the place experience occupies in the natural world.

References

- Beck, F. (1996), 'Can quantum processes control synaptic emission?', *International Journal of Neural Systems*, 7, pp. 343–53.
- Beck, F. and Eccles, J. (1992), 'Quantum aspects of brain activity and the role of consciousness', *Proceedings of the National Academy of Sciences USA*, 89, pp. 11357–61.
- Block, N. (1989), 'Can the mind change the world?', in *Meaning and Method: Essays in Honor of Hilary Putnam*, ed. G. Boolos (Cambridge: Cambridge University Press).
- Bohm, D. (1952), 'A suggested interpretation of the quantum theory in terms of hidden variables', *Physical Review*, 85, pp. 166–93.
- Chalmers, D.J. (1996), *The Conscious Mind* (New York: Oxford University Press).
- Crane, T. (1995), 'The mental causation debate (mental causation i)', *Aristotelian Society Supplement*, 69, pp. 211–36.
- Davidson, D. (1970), 'Mental events', in *Essays on Action and Events* (Oxford: OUP).
- Hagan, S., Hameroff, S. and Tuszynski, J. (2002), 'Quantum computation in brain microtubules: Decoherence and biological feasibility', *Physical Review E* (To appear).
- Hameroff, S. and Penrose, R. (1996), 'Conscious events as orchestrated space-time selections', *Journal of Consciousness Studies*, 3 (1), pp. 36–53.
- Hiley, B. and Pylykänen, P. (2001), 'Naturalizing the mind in a quantum framework', in *Dimensions of Conscious Experience*, ed. P. Pylykänen and T. Vaden (Amsterdam: John Benjamins).
- Honderich, T. (1993), 'The union theory and anti-individualism', in *Mental Causation*, ed. J. Heil, J. and A. Mele (Oxford: Oxford University Press).
- Hurley, S. (1998), 'Vehicles, contents, conceptual structure, and externalism', *Analysis*, 58, pp. 1–6.
- Jackson, F. (1996), 'Mental causation', *Mind*, 105, pp. 377–413.
- Kim, J. (1993), 'The nonreductivist's trouble with mental causation', in *Mental Causation*, ed. J. Heil, J. and A. Mele (Oxford: Oxford University Press).
- Leiter, B. and Miller, A. (1994), 'Mind doesn't matter yet', *Australasian Journal of Philosophy*, 72.
- LePore, E. and Loewer, B. (1989), 'More on making mind matter', *Philosophical Topics*, 17.
- Libet, B. (1985), 'Unconscious cerebral initiative and the role of conscious will in voluntary action', *Behavioral and Brain Sciences*, 8, pp. 529–66.
- Peacocke, C. (1979), *Holistic Explanation: Action, Space, Interpretation* (Oxford: Clarendon Press).
- Sloman, A. (2001), 'Diagrams in the mind', in *Diagrammatic Representation and Reasoning*, ed. M. Anderson, B. Meyer and P. Olivier (Berlin: Springer-Verlag).
- Sosa, E. (1984), 'Mind-body interaction and supervenient causation', *Midwest Studies in Philosophy*, 9.
- Unger, P. (1977), 'The uniqueness in causation', *American Philosophical Quarterly*, 14, pp. 177–88.

BRIDGING EASTERN AND WESTERN PERSPECTIVES ON CONSCIOUSNESS

K. Ramakrishna Rao

The target article by Max Velmans is yet another step in his continuing struggle to understand the complementarity of first-person and third-person aspects of consciousness. This struggle began about a decade ago with his first major effort in a target article (Velmans, 1991) in *Behavioral and Brain Sciences* and culminated in his *Understanding Consciousness* (Velmans, 2000). The present article, which covers the middle ground, is an instructive summary exposition of his main thesis that relates to the reconciliation of the irreducibility of conscious experience to brain states, on the one hand, and to its inalienable connection to the physical processes in the brain, on the other. In order to understand the full import of Velmans' reasoning and arguments, it is necessary to read the target article along with his earlier publications. Therefore, in the following comments I draw freely from Velmans' other writings.