

John R. Searle

Why I Am Not a Property Dualist

I have argued in a number of writings¹ that the philosophical part (though not the neurobiological part) of the traditional mind–body problem has a fairly simple and obvious solution: all of our mental phenomena are caused by lower level neuronal processes in the brain and are themselves realized in the brain as higher-level, or system, features. The form of causation is ‘bottom up’ whereby the behaviour of lower-level elements, presumably neurons and synapses, causes the higher-level or system features of consciousness and intentionality. (This form of causation, by the way, is common in nature; for example, the higher-level feature of solidity is causally explained by the behaviour of the lower-level elements, the molecules.) Because this view emphasizes the biological character of the mental, and because it treats mental phenomena as ordinary parts of nature, I have labelled it ‘biological naturalism’.

To many people biological naturalism looks a lot like property dualism. Because I believe property dualism is mistaken, I would like to try to clarify the differences between the two accounts and try to expose the weaknesses in property dualism. This short paper then has the two subjects expressed by the double meanings in its title: why my views are not the same as property dualism, and why I find property dualism unacceptable.

There are, of course, several different ‘mind–body’ problems. The one that most concerns me in this article is the relationship between consciousness and brain processes. I think that the conclusions of the discussion will extend to other features of the mind–body problem, such as, for example, the relationship between intentionality and brain processes, but for the sake of simplicity I will concentrate on consciousness. For the purposes of this discussion, the ‘mind–body problem’ is a problem about how consciousness relates to the brain.

The mind–body problem, so construed persists in philosophy because of two intellectual limitations on our part. First, we really do not understand how brain processes cause consciousness. Second, we continue to accept a traditional

Correspondence:

John R. Searle, Department of Philosophy, University of California, Berkeley, CA 94720, USA

[1] Initially in Searle (1983); subsequently in Searle (1984; 1992), and other writings.

vocabulary that contrasts the mental and the physical, the mind and the body, the soul and the flesh, in a way that I think is confused and obsolete. I cannot overcome our neurobiological ignorance, but I can at least try to overcome our conceptual confusion, and that is one of the things that I will attempt to do in this article.

I think it is because of these two limitations, our ignorance of how the brain works and our acceptance of the traditional vocabulary, that many people find property dualism appealing. Before criticizing it, I want to try to account for its appeal by stating the thesis with as much plausibility as I can. Of course, there are different versions of property dualism, but what I hope to state is the version that is closest to my own views and consequently the one I find most challenging. I will say nothing about ‘neutral monism’, panpsychism, or the various forms of ‘dual aspect’ theories. Notice that in presenting arguments for property dualism I have to use the traditional terminology that later on I will reject.

Here is how the world looks to the property dualist:

There is clearly a difference between consciousness and the material or physical world. We know this from our own experience, but it is also obvious from science. The material world is publicly accessible and is pretty much as described by physics, chemistry, and the other hard sciences; but the conscious, experiential, phenomenological world is not publicly accessible. It has a distinct private existence. We know it with certainty from our inner, private, subjective experiences. We all know that the private world of consciousness exists, we know that it is part of the real world, and our question is to find out how it fits into the public material world, specifically, we need to know how it fits into the brain.

Because neither consciousness nor matter is reducible to the other, they are distinct and different phenomena in the world. Those who believe that consciousness is reducible to matter are called materialists; those who believe that matter is reducible to consciousness are called idealists. Both are mistaken for the same reason. Both try to eliminate something that really exists in its own right and cannot be reduced to something else. Now, because both materialism and idealism are false, the only reasonable alternative is dualism. But substance dualism seems out of the question for a number of reasons. For example it cannot explain how these spiritual substances came into existence in the first place and it cannot explain how they relate to the physical world. So property dualism seems the only reasonable view of the mind–body problem. Consciousness really exists, but it is not a separate substance on its own, rather it is a property of the brain.

We can summarize property dualism in the following four propositions. The first three are statements endorsed by the property dualist, the fourth is an apparent consequence or difficulty implied by the first three:

- (1) There are two mutually exclusive metaphysical categories that constitute all of empirical reality: they are physical phenomena and mental phenomena. Physical phenomena are essentially objective in the sense that they exist apart from any subjective experiences of humans or animals. Mental phenomena are subjective, in the sense that they exist only as experienced by human or animal agents.

- (2) Because mental states are not reducible to neurobiological states, they are something *distinct from* and *over and above* neurobiological states. The irreducibility of the mental to the physical, of consciousness to neurobiology, is by itself sufficient proof of the distinctness of the mental, and proof that the mental is something over and above the neurobiological.
- (3) Mental phenomena do not constitute separate objects or substances, but rather are features or properties of the composite entity, which is a human being or an animal. So any conscious animal, such as a human being, will have two sorts of properties, mental properties and physical properties.
- (4) The chief problem for the property dualists, given these assumptions, is how can consciousness ever function causally? There are two possibilities, neither of which seems attractive. First, let us assume, as seems reasonable, that the physical universe is causally closed. It is closed in the sense that nothing outside it, nothing non-physical, could ever have causal effects inside the physical universe. If that is so, and consciousness is not a part of the physical universe, then it seems that it must be epiphenomenal. All of our conscious life plays no role whatever in any of our behaviour.

On the other hand, we may assume that the physical universe is not causally closed, that consciousness can function causally in the production of physical behaviour. But this seems to lead us out of the frying pan and into the fire, because we know, for example, that when I raise my arm, there is a story to be told at the level of neuron firings, neurotransmitters and muscle contractions that is entirely sufficient to account for the movement of my arm. So if we are to suppose that consciousness also functions in the movement of my arm, then it looks like we have two distinct causal stories, neither reducible to the other; and to put the matter very briefly, my bodily movements have too many causes. We have causal overdetermination.

The property dualist has a conception of consciousness and its relation to the rest of reality that I believe is profoundly mistaken. I can best make my differences with property dualism explicit by stating how I would deal with these same issues.

(1) There are not two (or five or seven) fundamental ontological categories, rather the act of categorization itself is always interest relative. For that reason the attempt to answer such questions as, 'How many fundamental metaphysical categories are there?', as it stands, is meaningless. We live in exactly one world and there are as many different ways of dividing it as you like. In addition to electromagnetism, consciousness, and gravitational attraction, there are declines in interest rates, points scored in football games, reasons for being suspicious of quantified modal logic, and election results in Florida. Now, quick, were the election results mental or physical? And how about the points scored in a football game? Do they exist only in the mind of the scorekeeper or are they rather ultimately electronic phenomena on the scoreboard? I think these are not interesting, or even meaningful, questions. We live in one world, and it has many different

types of features. My view is not 'pluralism' if that term suggests that there is a nonarbitrary, noninterest-relative principle of distinguishing the elements of the plurality. A useful distinction, for certain purposes, is to be made between the biological and the non-biological. At the most fundamental level, consciousness is a biological phenomenon in the sense that it is caused by biological processes, is itself a biological process, and interacts with other biological processes. Consciousness is a biological process like digestion, photosynthesis, or the secretion of bile. Of course, our conscious lives are shaped by our culture, but culture is itself an expression of our underlying biological capacities.

(2) Then what about irreducibility? This is the crucial distinction between my view and property dualism. Consciousness is causally reducible to brain processes, because all the features of consciousness are accounted for causally by neurobiological processes going on in the brain, and consciousness has no causal powers of its own in addition to the causal powers of the underlying neurobiology. But in the case of consciousness, causal reducibility does not lead to ontological reducibility. From the fact that consciousness is entirely accounted for causally by neuron firings, for example, it does not follow that consciousness is nothing but neuron firings. Why not? What is the difference between consciousness and other phenomena that undergo an ontological reduction on the basis of a causal reduction, phenomena such as colour and solidity? The difference is that consciousness has a first-person ontology; that is, it only exists *as experienced* by some human or animal, and therefore, it cannot be reduced to something that has a third-person ontology, something that exists independently of experiences. It is as simple as that.

The property dualist and I are in agreement that consciousness is ontologically irreducible. The key points of disagreement are that I insist that from everything we know about the brain, consciousness is causally reducible to brain processes; and for that reason I deny that the ontological irreducibility of consciousness implies that consciousness is something 'over and above', something distinct from, its neurobiological base. No, causally speaking, there is nothing there, except the neurobiology, which has a higher level feature of consciousness. In a similar way there is nothing in the car engine except molecules, which have such higher level features as the solidity of the cylinder block, the shape of the piston, the firing of the spark plug, etc. 'Consciousness' does not name a distinct, separate phenomenon, something over and above its neurobiological base, rather it names a state that the neurobiological system can be in. Just as the shape of the piston and the solidity of the cylinder block are not something over and above the molecular phenomena, but are rather states of the system of molecules, so the consciousness of the brain is not something over and above the neuronal phenomena, but rather a state that the neuronal system is in.

So there is a sense in which consciousness is reducible: the mark of empirical reality is the possession of cause and effect relations, and consciousness (like other system features) has no cause and effect relations beyond those of its microstructural base. There is nothing in your brain except neurons (together with glial cells, blood flow and all the rest of it) and sometimes a big chunk of the

thalamocortical system is conscious. The sense in which, though causally reducible, it is ontologically irreducible, is that a complete description of the third-person objective features of the brain would not be a description of its first-person subjective features.

(3) I say consciousness is a feature of the brain. The property dualist says consciousness is a feature of the brain. This creates the illusion that we are saying the same thing. But we are not, as I hope my response to points 1 and 2 makes clear. The property dualist means that *in addition to* all the neurobiological features of the brain, there is an extra, distinct, nonphysical feature of the brain; whereas I mean that consciousness is a state the brain can be in, in the way that liquidity and solidity are states that water can be in.

Here is where the inadequacy of the traditional terminology comes out most obviously. The property dualist wants to say that consciousness is a mental and therefore not physical feature of the brain. I want to say consciousness is a mental and therefore biological and therefore physical feature of the brain. But because the traditional vocabulary was designed to contrast the mental and the physical, I cannot say what I want to say in the traditional vocabulary without sounding like I am saying something inconsistent. Similarly when the identity theorists said that consciousness is nothing but a neurobiological process, they meant that consciousness as qualitative, subjective, irreducibly phenomenological (airy fairy, touchy feely, etc.) does not even exist, that only third-person neurobiological processes exist. I want also to say that consciousness is nothing but a neurobiological process, and by that I mean that precisely because consciousness is qualitative, subjective, irreducibly phenomenological (airy fairy, touchy feely, etc.) it has to be a neurobiological process; because, so far, we have not found any system that can cause and realize conscious states except brain systems. Maybe someday we will be able to create conscious artifacts, in which case subjective states of consciousness will be 'physical' features of those artifacts.

(4) Because irreducible consciousness is not something over and above its neural base, the problems about epiphenomenalism and the causal closure of the physical simply do not arise for me. Of course, the universe is causally closed, and we can call it 'physical' if we like; but that cannot mean 'physical' as opposed to 'mental'; because, equally obviously, the mental is part of the causal structure of the universe in the same way that the solidity of pistons is part of the causal structure of the universe; even though the solidity is entirely accounted for by molecular behaviour, and consciousness is entirely accounted for by neuronal behaviour. The problems about epiphenomenalism and the causal closure of the physical can only arise if one uses the traditional terminology and take its implications seriously. I am trying to get us to abandon that terminology.

But if consciousness has no causal powers in addition to its neurobiological base, then does that not imply epiphenomenalism? No. Compare: the solidity of the piston has no causal powers in addition to its molecular base, but this does not show that solidity is epiphenomenal (try making a piston out of butter or water). The question rather is: why would anyone suppose that causal reducibility implies epiphenomenalism, since the real world is full of causally efficacious

higher-level features entirely caused by lower-level micro phenomena? In this case the answer is: because they think that consciousness is something distinct from, something 'over and above' its neuronal base. The typical property dualist thinks that the brain 'gives rise to' consciousness, and this gives us a picture of consciousness as given off from the brain as a pot of boiling water gives off steam. In the epiphenomenalist version of property dualism, the consciousness given off has no causal powers of its own, though it is caused by the brain. In the full-blooded version consciousness has a kind of life of its own, capable of interfering with the material world.

I think this whole way of thinking of the matter is profoundly mistaken and I want to explain this point in a little more detail. The fact that the dilemma of either epiphenomenalism or causal overdetermination can even seem to be a problem for property dualism is a symptom that something is radically wrong with the theory. Nobody thinks that we are forced to postulate that solidity is epiphenomenal on the grounds that it has no causal powers in addition to the causal powers of the molecular structures, nor do they think that if we recognize the causal powers of solidity we are forced to postulate causal overdetermination, because now the same effect can be explained either in terms of the behaviour of the molecules or the solidity of the whole structure. And what goes for solidity goes for photosynthesis, digestion, electricity, earthquakes, hurricanes in Kansas, and pretty much everything else that we normally cite in causal explanations. In every case the higher-level phenomenon is causally reducible to its microstructural basis, in exactly the same way that consciousness is causally reducible to its microstructural basis. Why are we inclined to make this mistake for consciousness when we would not think of making it for other causal phenomena? I think the answer is obvious. Because the traditional vocabulary tells us that the mental and the physical are two distinct ontological categories and because consciousness is not ontologically reducible to its neuronal base, we suppose that is not a part of the physical world, in the way that these other phenomena are. That is the deeper mistake of property dualism. And that is precisely where I part company with the property dualist. The problem is not only that we have an obsolete seventeenth-century vocabulary that contrasts the mental and the physical, but that we also have a misconception of the nature of reduction. Causal reduction does not necessarily imply ontological reduction, though typically where we have a causal reduction as in the case of the liquidity, solidity and colour we have tended to make an ontological reduction. But the impossibility of an ontological reduction in the case of consciousness does not give it any mysterious metaphysical status. Consciousness does not exist in a separate realm and it does not have any causal powers in addition to those of its neuronal base any more than solidity has any extra causal powers in addition to its molecular base.

Both materialism and dualism are trying to say something true, but they both wind up saying something false. The materialist is trying to say, truly, that the universe consists entirely of material phenomena such as physical particles in fields of force. But he ends up saying, falsely, that irreducible states of consciousness do not exist. The dualist is trying to say, truly, that ontologically

irreducible states of consciousness do exist, but he ends up saying, falsely, that these are not ordinary parts of the physical world. The trick is to state the truth in each view without saying the falsehood. To do that we have to challenge the assumptions behind the traditional vocabulary. The traditional vocabulary is based on the assumption that if something is a state of consciousness in the strict sense — it is inner, qualitative, subjective, etc. — then it cannot in those very respects be physical or material. And conversely if something is physical or material then it cannot in its physical or material respects be a state of consciousness. Once you abandon the assumptions behind the traditional vocabulary it is not hard to state the truth. The universe does consist entirely in physical particles in fields of force (or whatever the ultimately true physics discovers), these are typically organized into systems, some of the systems are biological, and some of the biological systems are conscious. Consciousness is thus an ordinary feature of certain biological systems, in the same way that photosynthesis, digestion, and lactation are ordinary features of biological systems.

Addendum

There is an important issue that I have not pursued in this article, but I want at least to raise as a further problem for property dualism. It is not at all easy to see how the property dualist can maintain simultaneously that consciousness is a property or feature of the brain and that there is a metaphysical dualism of the mental and the physical. How, in short, can the property dualist avoid lapsing into substance dualism? The difficulty comes out in the metaphors that the property dualist uses to express the thesis. Typical metaphors are that consciousness is something ‘over and above’ brain processes, that brains ‘give rise to’ consciousness and, of course, that consciousness is an ‘emergent’ property of the brain. But all of these metaphors suggest that the picture the dualist has is that consciousness is something separate from the brain. I said the property dualist thinks of consciousness as like steam rising from a pot of boiling water, but here is another picture suggested by these metaphors: we are to think of consciousness as like the frosting on the cake of the brain (and in its panpsychist version, the frosting on the whole universe). The frosting is something distinct from the cake and it is on top of (over and above it). I have argued that these are the wrong pictures. The right picture, if we are going to persist in the metaphor of the cake, is that consciousness is the state that the cake (brain) is in. Officially, the property dualist says that consciousness is a property of the brain; but if you consider uncontroversial properties of the brain, like weight, shape, colour, solidity, etc., nobody says that these ‘arise from’ or are ‘over and above’ the brain; and only in a special sense can some of them be described as ‘emergent’ (cf. Searle, 1992, pp. 111–12), and certainly not as ‘emergent from’ the brain. The official claim is that consciousness is a property, not a thing, object or substance. But that claim is inconsistent with the conception of consciousness as something that is ‘over and above’, that the brain ‘gives rise to’, etc.; this conception requires that consciousness be a separate thing, object, or non-property type of entity. The dualism in

property dualism forces them to postulate a separate entity. Ironically, the very dualism of the property dualist picture makes it impossible to state the theory without implying a version of substance dualism.

References

- Searle, J.R. (1983), *Intentionality: An Essay in the Philosophy of Mind* (Cambridge: Cambridge University Press).
Searle, J.R. (1984), *Minds, Brains and Science* (Cambridge, MA: Harvard University Press).
Searle, J.R. (1992), *The Rediscovery of the Mind* (Cambridge, MA: MIT Press).

Paper received October 2002