

J.J.C. Smart, “Sensations and brain processes”

Excerpts from J.J.C. Smart, “Sensations and brain processes” (*Philosophical Review* 68: 141-56, 1959)

When I say “I have a round yellowish-orange afterimage”, or “I am in pain”, am I reporting something non-physical, something “irreducibly psychical”? Smart thinks not:

The suggestion I wish if possible to avoid is...that “I am in pain” is a genuine report, and that what it reports is an irreducibly psychical something. And similarly the suggestion I wish to resist is also that to say “I have a yellowish orange after-image” is to report something irreducibly psychical.

Why does Smart wish to resist these suggestions?

Mainly because of Occam’s razor. It seems to me that science is increasingly giving us a viewpoint whereby organisms are able to be seen as physico-chemical mechanisms: it seems that even the behavior of man himself will one day be explicable in mechanistic terms. There does seem to be, so far as science is concerned, nothing in the world but increasingly complex arrangements of physical constituents. All except for one place: in consciousness. That is, for a full description of what is going on in a man you would have to mention not only the physical processes in his tissue, glands, nervous system, and so forth, but also his states of consciousness: his visual, auditory, and tactual sensations, his aches and pains. That these should be *correlated* with brain processes does not help, for to say that they are *correlated* is to say that they are something “over and above.” You cannot correlate something with itself. You correlate footprints with burglars, but not Bill Sikes the burglar with Bill Sikes the burglar. So sensations, states of consciousness, do seem to be the one sort of thing left outside the physicalist picture, and for various reasons I just cannot believe that this can be so. That everything should be explicable in terms of physics (together of course with descriptions of the ways in which the parts are put together—roughly, biology is to physics as radio-engineering is to electro-magnetism) except the occurrence of sensations seems to me to be frankly unbelievable. Such sensations would be “nomological danglers,”... It is not often realized how odd would be the laws whereby these nomological danglers would dangle. It is sometimes asked, “Why can’t there be psycho-physical laws which are of a novel sort, just as the laws of electricity and magnetism were novelties from the stand-point of Newtonian mechanics?” Certainly we are pretty sure in the future to come across new ultimate laws of a novel type, but I expect them to relate simple constituents: for example, whatever

ultimate particles are then in vogue. I cannot believe that ultimate laws of nature could relate simple constituents to configurations consisting of perhaps billions of neurons (and goodness knows how many billion billions of ultimate particles) all put together for all the world as though their main purpose in life was to be a negative feedback mechanism of a complicated sort. Such ultimate laws would be like nothing so far known in science. They have a queer "smell" to them. I am just unable to believe in the nomological danglers themselves, or in the laws whereby they would dangle. If any philosophical arguments seemed to compel us to believe in such things, I would suspect a catch in the argument. In any case it is the object of this paper to show that there are no philosophical arguments which compel us to be dualists....

Why should not sensations just be brain processes of a certain sort? There are, of course, well-known (as well as lesser-known) philosophical objections to the view that reports of sensations are reports of brain processes, but I shall try to argue that these arguments are by no means as cogent as is commonly thought to be the case.

Before doing that, Smart tries to "state more accurately the thesis that sensations are brain processes":

It is not the thesis that, for example; "after-image" or "ache" means the same as "brain process of sort X" (where "X" is replaced by a description of a certain sort of brain process). It is that, in so far as "after-image" or "ache" is a report of a process, it is a report of a process that *happens to be* a brain process. It follows that the thesis does not claim that sensation statements can be *translated* into statements about brain processes. Nor does it claim that the logic of a sensation statement is the same as that of a brain-process statement. All it claims is that in so far as a sensation statement is a report of something, that something is in fact a brain process. Sensations are nothing over and above brain processes. Nations are nothing "over and above" citizens, but this does not prevent the logic of nation statements being very different from the logic of citizen statements, nor does it insure the translatability of nation statements into citizen statements. (I do not, however, wish to assert that the relation of sensation statements to brain-process statements is very like that of nation statements to citizen statements. Nations do not just *happen to be* nothing over and above citizens, for example. I bring in the "nations" example merely to make a negative point:

that the fact that the logic of A-statements is different from that of B-statements does not insure that A's are anything over and above B's.)

What sort of "identity" is relevant?

When I say that a sensation is a brain process or that lightning is an electric discharge, I am using "is" in the sense of strict identity. (Just as in the—in this case necessary—proposition "7 is identical with the smallest prime number greater than 5.") When I say that a sensation is a brain process or that lightning is an electric discharge I do not mean just that the sensation is somehow spatially or temporally continuous with the brain process or that the lightning is just spatially or temporally continuous with the discharge...I wish to make it clear that the brain-process doctrine asserts identity in the *strict* sense.

Smart discusses eight objections to "the view that the processes reported in sensation statements are in fact processes in the brain", two of which are omitted from this excerpt.

Objection 1. Any illiterate peasant can talk perfectly well about his after-images, or how things look or feel to him, or about his aches and pains, and yet he may know nothing whatever about neurophysiology. A man may, like Aristotle, believe that the brain is an organ for cooling the body without any impairment of his ability to make true statements about his sensations. Hence the things we are talking about when we describe our sensations cannot be processes in the brain.

Reply. You might as well say that a nation of slug-abeds, who never saw the morning star or knew of its existence, or who had never thought of the expression "the Morning Star," but who used the expression "the Evening Star" perfectly well, could not use this expression to refer to the same entity as we refer to (and describe as) "the Morning Star."

(The Morning/Evening "Star" is the planet Venus.)

You may object that the Morning Star is in a sense not the very same thing as the Evening Star, but only something spatiotemporally continuous with it. That is, you may say that the Morning Star is not the Evening Star in the strict sense of "identity" that I distinguished earlier. I can perhaps forestall this objection by considering the slug-abeds to be New Zealanders and the early risers to be Englishmen. Then the thing the New Zealanders describe as "the Morning Star" could be the very same thing (in the strict sense) as the Englishmen describe as "the Evening Star." And yet they could be ignorant of this fact.

There is, however, a more plausible example. Consider lightning. Modern physical science tells us that lightning is a certain kind of electrical discharge due to ionization of clouds of water-vapor in the atmosphere. This, it is now believed, is what the true nature of lightning is. Note that there are not two things: a flash of lightning and an electrical discharge. There is one thing, a flash of lightning, which is described scientifically as an electrical discharge to the earth from a cloud of ionized water-molecules. The case is not at all like that of explaining a footprint by reference to a burglar. We say that what lightning really is, what its true nature as revealed by science is, is an electric discharge. (It is not the true nature of a footprint to be a burglar.)

To forestall irrelevant objections, I should like to make it clear that by "lightning" I mean the publicly observable physical object, lightning, not a visual sense-datum of lightning. I say that the publicly observable physical object lightning is in fact the electric discharge, not just a correlate of it. The sense-datum, or at least the having of the sense-datum, the "look" of lightning, may well in my view be a correlate of the electric discharge. For in my view it is a brain state *caused* by the lightning. But we should no more confuse sensations of lightning with lightning than we confuse sensations of a table with the table.

In short, the reply to Objection 1 is that there can be contingent statements of the form "A is identical with B," and a person may well know that something is an A without knowing that it is a B. An illiterate peasant might well be able to talk about his sensations without knowing about his brain processes, just as he can talk about lightning though he knows nothing of electricity.

Objection 2. It is only a contingent fact (if it is a fact) that when we have a certain kind of sensation there is a certain kind of process in our brain. Indeed it is possible, though perhaps in the highest degree unlikely, that our present physiological theories will be as out of date as the ancient theory connecting mental processes with goings on in the heart. It follows that when we report a sensation we are not reporting a brain-process.

Reply. The objection certainly proves that when we say "I have an after-image" we cannot *mean* something of the form "I have such and such a brain process." But this does not show that what we report (having an after-image) is not *in fact* a brain process. "I see lightning" does not *mean* "I see an electric discharge." Indeed, it is logically possible (though highly unlikely) that the electrical discharge account of lightning might one day be given up.

Again, "I see the Evening Star" does not *mean* the same as "I see the Morning Star," and yet "the Evening Star and the Morning Star are one and the same thing" is a contingent proposition...

Objection 3. Even if Objections 1 and 2 do not prove that sensations are something over and above brain processes, they do prove that the qualities of sensations are something over and above the qualities of brain processes. That is, it may be possible to get out of asserting the existence of irreducibly psychic processes, but not out of asserting the existence of irreducibly psychic *properties*. For suppose we identify the Morning Star with the Evening Star. Then there must be some properties which logically imply that of being the Morning Star, and quite distinct properties which entail that of being the Evening Star. Again, there must be some properties (for example, that of being a yellow flash) which are logically distinct from those in the physicalist story...

Now how do I get over [this] objection that a sensation can be identified with a brain process only if it has some [irreducibly psychic] property...whereby one-half of the identification may be, so to speak, pinned down?

My suggestion is as follows. When a person says, "I see a yellowish-orange after-image," he is saying something like this: "*There is something going on which is like what is going on when I have my eyes open, am awake, and there is an orange illuminated in good light in front of me, that is, when I really see an orange.*" ...Notice that the italicized words, namely "there is something going on which is like what is going on when," are all quasi-logical or topic-neutral words. This explains why the ancient Greek peasant's reports about his sensations can be neutral between dualistic metaphysics or my materialistic metaphysics. It explains how sensations can be brain processes and yet how those who report them need know nothing about brain processes. For he reports them only very abstractly as "something going on which is like what is going on when..." Similarly, a person may say "someone is in the room," thus reporting truly that the doctor is in the room, even though he has never heard of doctors. (There are not two people in the room: "someone" *and* the doctor.) This account of sensation statements also explains the singular elusiveness of "raw feels"—why no one seems to be able to pin any properties on them. Raw feels, in my view, are colorless for the very same reason that *something* is colorless. This does not mean that sensations do not have properties, for if they are brain processes they

certainly have properties. It only means that in speaking of them as being like or unlike one another we need not know or mention these properties.

This, then, is how I would reply to Objection 3. The strength of my reply depends on the possibility of our being able to report that one thing is like another without being able to state the respect in which it is like. I am not sure whether this is so or not, and that is why I regard Objection 3 as the strongest with which I have to deal.

Objection 4. The after-image is not in physical space. The brain process is. So the after-image is not a brain process.

Reply. This is an *ignoratio elenchi*. I am not arguing that the after-image is a brain process, but that the experience of having an after-image is a brain process. It is the *experience* which is reported in the introspective report. Similarly, if it is objected that the after-image is yellowy-orange but that a surgeon looking into your brain would see nothing yellowy-orange, my reply is that it is the experience of seeing yellowy-orange that is being described, and this experience is not a yellow-orange something. So to say that a brain-process cannot be yellowy-orange is not to say that a brain-process cannot in fact be the experience of having a yellowy-orange after-image. There is, in a sense, no such thing as an after-image or a sense-datum, though there is such a thing as the experience of having an image, and this experience is described indirectly in material object language, not in phenomenal language, for there is no such thing. We describe the experience by saying, in effect, that it is like the experience we have when, for example, we really see a yellowy-orange patch on the wall. Trees and wallpaper can be green, but not the experience of seeing or imagining a tree or wallpaper. (Or if they are described as green or yellow this can only be in a derived sense.)

Objection 5. It would make sense to say of a molecular movement in the brain that it is swift or slow, straight or circular, but it makes no sense to say this of the experience of seeing something yellow.

Reply. So far we have not given sense to talk of experiences as swift or slow, straight or circular. But I am not claiming that "experience" and "brain process" mean the same or even that they have the same logic. "Somebody" and "the doctor" do not have the same logic, but this does not lead us to suppose that talking about somebody telephoning is talking about someone over and above, say, the doctor. The ordinary man when he reports an experience is reporting that something is going on, but he leaves it open as

to what sort of thing is going on, whether in a material solid medium, or perhaps in some sort of gaseous medium, or even perhaps in some sort of nonspatial medium (if this makes sense). All that I am saying is that "experience" and "brain process" may in fact refer to the same thing, and if so we may easily adopt a convention (which is not a change in our present rules for the use of experience words but an addition to them) whereby it would make sense to talk of an experience in terms appropriate to physical processes...

Objection 7. I can imagine myself turned to stone and yet having images, aches, pains, and so on.

Reply. I can imagine that the electrical theory of lightning is false, that lightning is some sort of purely optical phenomenon. I can imagine that lightning is not an electrical discharge. I can imagine that the Evening Star is not the Morning Star. But it is. All the objection shows is that "experience" and "brain process" do not have the same meaning. It does not show that an experience is not in fact a brain process.

This objection is perhaps much the same as one which can be summed up by the slogan: "What can be composed of nothing cannot be composed of anything." The argument goes as follows: on the brain process thesis the identity between the brain process and the experience is a contingent one. So it is logically possible that there should be no brain process, and no process of any other sort, either (no heart process, no kidney process, no liver process). There would be the experience but no "corresponding" physiological process with which we might be able to identify it empirically.

I suspect that the objector is thinking of the experience as a ghostly entity. So it is composed of something, not of nothing, after all. On his view it is composed of ghost stuff, and on mine it is composed of brain stuff. Perhaps the counter-reply will be that the experience is simple and uncompounded, and so it is not composed of anything after all. This seems to be a quibble, for, if it were taken seriously, the remark "What can be composed of nothing cannot be composed of anything" could be recast as an a priori argument against Democritus and atomism and for Descartes and infinite divisibility. And it seems odd that a question of this sort could be settled a priori. We must therefore construe the word "composed" in a very weak sense, which would allow us to say that even an indivisible atom is composed of something (namely, itself). The dualist cannot really say that an experience can be composed of nothing. For he holds that experiences are something

over and above material processes, that is, that they are a sort of ghost stuff. (Or perhaps ripples in an underlying ghost stuff.) I say that the dualist's hypothesis is a perfectly intelligible one. But I say that experiences are not to be identified with ghost stuff but with brain stuff. This is another hypothesis, and in my view a very plausible one. The present argument cannot knock it down a priori...

Having discussed eight objections, Smart concludes with "some remarks on the logical status of the thesis itself":

U. T. Place¹ seems to hold that it is a straight-out scientific hypothesis. If so, he is partly right and partly wrong. If the issue is between (say) a brain-process thesis and a heart thesis, or a liver thesis, or a kidney thesis, then the issue is a purely empirical one, and the verdict is overwhelmingly in favor of the brain. The right sorts of things don't go on in the heart, liver, or kidney, nor do these organs possess the right sort of complexity of structure. On the other hand, if the issue is between a brain-or-heart-or-liver-or-kidney thesis (that is, some form of materialism) on the one hand and epiphenomenalism on the other hand, then the issue is not an empirical one. For there is no conceivable experiment which could decide between materialism and epiphenomenalism. This latter issue is not like the average straight-out empirical issue in science, but like the issue between the nineteenth-century English naturalist Philip Gosse and the orthodox geologists and paleontologists of his day. According to Gosse, the earth was created about 4000 B.C. exactly as described in *Genesis*, with twisted rock strata, "evidence" of erosion, and so forth, and all sorts of fossils, all in their appropriate strata, just as if the usual evolutionist story had been true. Clearly this theory is in a sense irrefutable: no evidence can possibly tell against it. Let us ignore the theological setting in which Philip Gosse's hypothesis had been placed, thus ruling out objections of a theological kind, such as "what a queer God who would go to such elaborate lengths to deceive us." Let us suppose that it is held that the universe just began in 4004 B.C. with the initial conditions just everywhere as they were in 4004 B.C., and in particular that our own planet began with sediment in the rivers, eroded cliffs, fossils in the rocks, and so on. No scientist would ever entertain this as a serious hypothesis, consistent though it is with all possible evidence. The hypothesis offends against the principles of parsimony and simplicity. There would be far too many brute and inexplicable facts. Why are pterodactyl bones just as they are? No

¹ See the next reading, Place's "Is consciousness a brain process?"

explanation in terms of the evolution of pterodactyls from earlier forms of life would any longer be possible. We would have millions of facts about the world as it was in 4004 B.C. that just have to be *accepted*.

The issue between the brain-process theory and epiphenomenalism seems to be of the above sort. (Assuming that a behavioristic reduction of introspective reports is not possible.) If it be agreed that there are no cogent philosophical arguments which force us into accepting dualism, and if the brain-process theory and dualism are equally consistent with the facts, then the principles of parsimony and simplicity seem to me to decide overwhelmingly in favor of the brain-process theory. As I pointed out earlier, dualism involves a large number of irreducible psychophysical laws (whereby the "nomological danglers" dangle) of a queer sort, that just have to be taken on trust, and are just as difficult to swallow as the irreducible facts about the paleontology of the earth with which we are faced on Philip Gosse's theory.