COMMENTS ON ADAM PAUTZ "WHY CONSCIOUSNESS CANNOT BE IN THE HEAD"

Sydney Shoemaker – Cornell University

I agree with Adam Pautz in rejecting biological theories of consciousness – theories that hold that every experience property "is necessarily identical with some (perhaps global, distributed) *neural* property." I cannot myself see why such theories are not refuted by the multiple realizablity argument, which Pautz evidently thinks is not up to the job. I will come back to this later. But for now my focus will be on Pautz's claim that such theories are refuted by the fact that no neural state can have, as a matter of necessity, "externally directed" properties that experiences necessarily possess. The first part of my comments will raise an objection to the claim that experience properties do necessarily have the externally directed properties Pautz claims they have, and then suggest a modification of the claim that avoids that objection. The second part will argue that even if the physical realizers of experience properties like Pautz's R have externally directed properties necessarily, this would not give us an identity between these and biological properties. This is because the physical realizers cannot be purely biological, since they involve more than what is inside the head. (It is also because no multiply realizable property can be identical with any of its realizers.) .

Pautz's externally directed properties, which he says necessarily belong to the property R, are the *externality* property of being such that if anyone has it then he has an experience as of a round thing, the *matching* property of being such that if anyone has it then he is in a state that matches the world only if a round thing is present, the *grounding* property of being such that if its possessor has the general capacity to possess beliefs he

will have the additional capacity to have a general belief that is true only if something or other is present that is round, and the *justification* property of being such that if the believer has it in the absence of undercutting defeaters, then he has a justification for believing that a round thing is present.

Let me say first that I think that the intuitions on which the claim that R necessarily has these externality properties rest on a questionable use of the first person perspective. It is certainly true that I can't imagine having an instance of R without these externality properties being instantiated. But that at best shows that it is necessarily true that in creatures of our sort R has these externality properties. What of creatures that are not of our sort? This brings me to an objection that Pautz briefly considers but does not, I think, satisfactorily answer.

The objection is that in a brain-in-a-vat scenario or an inverted spectrum scenario, R can be instantiated without having these properties. I will focus on the brain-in-a vat version of the objection. Pautz's main response to this is that the objection rests on a "controversial, simple pure input-based causal theory of the externally-directed properties on which (for instance) R is an experience of a round thing only if it is normally caused by round things," and can be met if we accept a "more output-based" theory and say that "in these scenarios R counts as an experience of a round thing, on the grounds that having R involves being in a state that would lead to actions appropriate to a round object…." In the case of the brain-in-a-vat, "R would lead to such actions were the brain embodied." .

To bring out why I find this response unsatisfactory let me elaborate the relevant brain-in-a-vat scenario – basically that presented by Putnam. The brain is one that has

never been normally embodied. And it has always been attached to a computer that produces inputs to it that mimic those it would receive if it were normally embodied in a certain normal earthly environment. Moreover, the neural states that in a normally embodied brain would generate bodily behavior affect the computer in ways that modify its input to the brain in ways similar to the ways the actions of a normally embodied person affect his sensory input. In virtue of all of this there are realized in the brain beliefs whose truth conditions have to do with states of the computer and volitional states whose satisfaction conditions involve states of the computer. None of the beliefs and desires have the contents that the possessor of the brain would have if it were normally embodied; but they have contents that are, as it were, counterparts of these. Where the possessor of the normally embodied brain would have the belief "I see something red and round," the BIV would have the belief "I see something that is red* and round*." – where the starred predicates stand for states of the computer. It is plausible that the neural state that realizes R in a normal brain would also realize it in the BIV. But here R would not have the *externality* property, the *grounding* property, and the *justification* property (I am not sure whether it would count as having the *matching* property). Its connections would be with red* and round*, not with red and round

What about Pautz's point that if the brain in the BIV case were normally embodied R would lead to actions appropriate to red and round? No doubt that is true. But it is equally true of the brain of a normal person that if it were hooked up to a computer in the way the BIV is, R in it would lead to actions (or actions*) appropriate to red* and round*. We could formulate counterparts of Pautz's externality, matching, grounding, and justification properties that are defined in terms of red* and round* rather

than in terms of red and round – call these externality*, matching*, grounding* and justification*. Presumably Pautz wouldn't want to say that in the normal person R has these properties, so it hardly seems that the counterfactual true of the BIV would give R in that case the externally directed properties Pautz wants it to have.

But this suggests a way of modifying Pautz's claim so as to finesse the BIV objection. Perhaps what necessarily belong to R are not the externality, matching, grounding, and justification properties as he characterizes them, but more abstract properties that stand to these as determinables to determinates – ones that are realized in us by Pautz's externality, etc. properties and are realized in the BIV by the externality*, etc., properties. E.g., what is necessarily true of R is that if someone has it then he has an experience of something having properties that are in certain ways functionally like the properties red and round.

This view goes with a functionalist view about how properties like R have the representational properties they have, and Pautz would probably reject it for that reason. For my own part, I think that it is only on such a functionalist view that it is plausible that properties like R necessarily have the externally directed properties.

I turn now to Pautz's claim that the biological realizers of properties like R do not have, as a matter of necessity, the externally directed properties they have. I take it to be analytic that the instantiation of a total realizer of a property is necessarily sufficient for the instantiation of that property. And if the realized property has certain externally directed properties, so also must any of its total realizers. If one is not a necessitarian about causal laws, then no biological property can by itself be a total realizer of property S; at best, it could be only a core realizer that is part of a total realizer of it, the rest of that

total realizer being the holding of a certain set of causal laws. And in that case the biological property could not have, necessarily, the externally directed properties S has. It might seem that if one is a necessitarian (as I am), then one could hold that a biological property could be by itself a total realizer of R, and so have its externally directed properties. But the very fact that R necessarily has externally directed properties rules out its having any biological property as a total realizer. I think its having these externally directed properties makes it an intentional property. I take it that an intentional property like the belief that there is water in the glass is realized in part by things outside the head of the believer – in this case, by the existence of water in the believer's environment.. The realizer will be physical, and in part biological, but it won't be purely biological. And the same will be true of R, given that it necessarily has externally directed properties. This will be so if it necessarily has the ones Pautz's claims it has. But it will also be so if has the more abstract ones I suggested he should hold it has. Here too the realizers will be in part biological. But they will also consist in part of facts about the subject's environment; on the account I suggested, these will be very messy, highly disjunctive physical facts.[1]

But, assuming physicalism, the extra-cranial part of the total realizer will still be physical. So why can't R be identical with something physical, namely such a realizer? Given necessitarianism about laws, it will have the externally directed properties R has. But I don't think one has to reject necessitarianism in order to rule out such an identity

[1] Pautz says that "where N is a total realizer the view that R is identical with N is not a version of the biological theory." His reason for saying this seems to be that "big state theory is very much like functionalism," which I take it means that big states, e.g., total realizers, are too much like functional properties to be biological properties. But which total realizers may realize functional properties, they are not themselves functional properties or like functional properties. If R were completely internal to its possessors, its total realizer could be biological. What prevents the total realizer from being biological is the fact that R fails to be completely internal due to its possession of the externally directed properties.

claim. A much better reason for rejecting it is the multiple realization argument. Properties like R are multiply realizable, or so I think. And as a matter of logic, no multiply realizable property can be identical with one of its realizers. (Pautz mentions a number of writers who have called into question the multiple realizability objection to psychophysical identity theories; but I cannot imagine that any of these writers would hold that a multiply realizable property could be identical with one of its realizers – that is ruled out by Leibniz's Law – and I don't see how the physicalists among them could deny that mental properties are multiply realizable.) I might add that I think it is also clear that no instance of a multiply realizable property can be identical with an instance of one of its realizers – the instances will differ because they differ in their causal powers, which they will do because the powers conferred by the realized property will be a proper subset of those conferred by its realizer.

I offered Pautz a modified version of his view that properties like R necessarily have externally directed properties, but I am myself uncertain whether even that modified version is true. I am uncertain of this because it conflicts with the view, which I find attractive, that the phenomenal character of conscious states is an entirely internal matter. If I accept the latter I can hold that properties like R can have purely biological total realizers. These won't have externally directed properties because the properties they realize won't have such properties. Even so, these biological properties will not be identical with the properties they realize, just because the latter are multiply realizable. At the moment I cannot decide between the purely internalist view about phenomenal character and the view I have offered to Pautz. But I lean towards the latter, because I think there may be a version of it that gives me what I want from the former; one could

hold that whether a creature has phenomenal states at all depends on facts "outside the head," but that where creatures do have phenomenal states the phenomenal character of their states supervenes on what is inside the head.