International Journal of Philosophical Studies Vol. 19(4), 593–602

Routledge
Taylor & Francis Group

Discussion

Polger on the Illusion of Contingent Identity

Don Merrell

Abstract

Thomas Polger has argued in favour of the mind–brain type-identity theory, the view that mental states or processes are type-identical to states of the central nervous system. Acknowledging that the type-materialist must respond to Kripke's modal anti-materialist argument, Polger insists that Kripke's argument rests on dubious assumptions concerning the identity conditions of brain states. In brief, Polger claims that one knows that *x* and *y* are non-identical when one knows the *identity conditions* for both *x* and *y*. Replace *x* and *y* with 'brain states' and 'sensations' and it follows that one can know that brain states and sensations are *non*-identical only if one knows the identity conditions for brain states. But according to Polger, we do *not* know the identity conditions for brain states. Hence, we should not be so confident that brain states and sensations are non-identical after all. But Polger's account is terribly flawed. Ironically, if Polger's scepticism is warranted, then Polger himself has no good reasons to be a type-materialist. But more importantly, Polger's scepticism regarding the identity conditions of brain states is deeply defective. We do, I submit, understand the identity conditions of brain states. In the end, I submit, Kripke is safe from Polger.

1 Introduction

Thomas Polger has argued in favour of the mind–brain type-identity theory, the view that mental states or processes are type-identical to states of the central nervous system.[1] Acknowledging that the type-materialist must respond to Kripke's modal anti-materialist argument, Polger insists that Kripke's argument rests on dubious assumptions concerning the identity conditions of brain states. In brief, Polger claims that one knows that *x* and *y* are non-identical when one knows the *identity conditions* for both *x* and *y*. Replace *x* and *y* with 'brain states' and 'sensations' and it follows that one can know that brain states and sensations are *non*-identical only if one knows the identity conditions for brain states.[2] But according to Polger, we do *not* know the identity conditions for brain states. Hence, we should not be so confident that brain states and sensations are non-identical after all.

International Journal of Philosophical Studies
ISSN 0967–2559 print 1466–4542 online
http://www.tandfonline.com
http://dx.doi.org/10.1080/09672559.2011.539370

But Polger's account is terribly flawed. Ironically, if Polger's scepticism is warranted, then Polger himself has no good reasons to be a type-materialist. But more importantly, Polger's scepticism regarding the identity conditions of brain states is deeply defective. We do, I submit, understand the identity conditions of brain states. In the end, I submit, Kripke is safe from Polger.

2 Kripke versus the Identity Theory

Recall that Kripke's argument proceeds from his theory of rigid designation. A 'rigid designator', according to Kripke, is a term that refers to the same thing in any possible world in which it happens to refer. If 'pain' and 'c-fibre firing ('CFF') are rigid designators – and Kripke maintains that they are – then if in fact pain is CFF, then *necessarily* pain is CFF. But the identity clearly does not *seem* necessary, for we can imagine, for example, a pure Cartesian soul world where pain and other mental types exist in the absence of brains or nervous systems; and we can imagine worlds where CFF exists but designates, say, euphoria instead of pain. In short, it appears that 'pain' and 'CFF' could have designated different types of events. So, for the identity theorist to triumph, she must explain the apparent contingency of the mind–brain identity statement. And this can be done with other apparent contingencies like 'water = H_2O' and 'heat = molecular motion' – identities that do not seem necessarily true. But these apparently contingent identity statements may be explained away since one can conceive of 'epistemic counterparts' of water and heat – respectively, some substance which has all the phenomenal properties of water but which is, say, XYZ instead of H_2O; and some epistemic or qualitative counterpart of heat that is caused by something other than molecular motion. Imagining these possibilities in no way shows that water is not H_2O or that heat is not molecular motion.

So, Kripke offers the type-materialist a way out: come up with an epistemic counterpart of this-worldly pain and insist that what we are imagining when we imagine that pain is not identical to CFF is, really, an imagining of something else, something with the familiar phenomenal profile of pain but which is not genuine pain. This epistemic imposter of pain may indeed not be identical with CFF in the same way that the epistemic imposter of water may not be identical to H_2O. Imagining this in no way shows that pain is not CFF. The identity theorist, it seems, is back on solid ground.

But though this explanation may succeed in explaining the contingency of water being H_2O, and of heat being molecular motion, it is unsuccessful in the case of pain being CFF. And the reason for this is that pains (presumably *all* sensation types) do not admit of an appearance–reality distinction, or what amounts to the same thing, being painful is not a contingent property of pain. Here's Kripke:

> It is impossible to explain the apparent possibility of C-fiber stimulation not having been pain in the same way. Here too, we would have to suppose that we could have been in the same epistemological situation, and identify something in the same way we identify pain, without its corresponding to C-fiber stimulation. But the way we identify pain is by feeling it, and if a C-fiber stimulation could have occurred without our feeling any pain, then the C-fiber stimulation would have occurred without there being any pain, contrary to the necessity of the identity. The trouble is that although 'heat' is a rigid designator, heat is picked out by the contingent property of its being felt in a certain way; pain, on the other hand, is picked out by an essential (and indeed necessary and sufficient) property. For a sensation to be felt as pain is for it to be pain.[3]

In the end, it seems that the type-materialist will not be able to employ Kripke's epistemic model for explaining away the contingency of brain states and mental states. This means that the appearance of contingency has *not* been explained and, hence, type-materialism is contingently true – true in *some* worlds but not all. But there's the rub: since there is no contingent identity between rigid designators, type-materialism is false.

3 Polger on Candidacy Criteria for Identity Conditions

There are several ways one might respond to Kripke. For example, one could deny that 'pain' is a rigid designator, one could deny that pain is essentially painful, one could deny that imagination entails metaphysical possibility, etc. – any of which would challenge Kripke's project by attacking his analytical tools. But Polger accepts all of Kripke's tools and instead takes seriously the idea that the way to refute him is to accept the challenge of coming up with an epistemic model distinct from Kripke's which will allow us to explain our Cartesian intuitions and which is consistent with type-materialism. As Kripke says, '[The materialist] has to hold that we are under some illusion in thinking that we can imagine that there could have been pains without brain states. ... So the materialist is up against a very stiff challenge. He has to show that these things we think we can see to be possible are in fact not possible. He has to show that these things we think we can imagine are not in fact things we can imagine.'[4]

Like Kripke's, Polger's model is epistemic. The idea is to show that it is possible that something has gone wrong, epistemically speaking, when we imagine pain and CFF coming apart. And since the identities in question – the necessary a posteriori – appear contingent when in fact they are not, Polger owes us an epistemic model that will do what Kripke's will not, namely, explain away the apparent contingency of mind–brain identity claims.

The first step in Polger's model building is to ask how we discover necessary a posteriori scientific identities. Drawing on the work of Alan Sidelle, Polger notes that an obvious question we must ask when entertaining the possibility that two things are identical is why we entertain *some* candidates for identification over others. For example, though water turned out to be H_2O, it could have turned out to be H_2SO_4;[5] and naturally we would have been just as willing to accept it as H_2O as a candidate for identity with water. But notice that we would not consider 'the liquid enjoyed by George Washington' as a candidate for identity with water. Sidelle writes, 'What this suggests is that there are "candidacy" criteria which anything must meet if it is to be a candidate for identification with our subject of inquiry. And I suggest that in order to be such a candidate, something must have identity conditions which are compatible with those for our subject.'[6] Now, it is probably the case that guiding our recognition of identity conditions for types of things is some intuitive, folkish knowledge of the way the world fits together. For example, it was discovered that gold is the element with atomic number 14; 'but', Polger says, 'we do not even consider that gold is a Bordeaux wine. Common material stuffs have identity conditions ... that involve their contemporaneous microstructure, whereas French wines have identity conditions ... that involve their appellation and vintage. We do not even consider the identification of gold with Bordeaux, for the identity conditions ... for gold are incompatible with those for Bordeaux. Something is gold whether or not it comes from a certain place, at a certain time; not so for Chateau Haut-Brion.'[7]

On the other hand, it would have been entirely reasonable to entertain the idea that, as far as candidacy goes, gold was not identical to the element with atomic number 13 but with whatever is the atomic structure of lead. In fact, though, the specific identity conditions of gold are incompatible with those of lead. 'But', as Polger notes, 'that empirical discovery is prefaced on knowing that the general identity conditions for each were microstructural, and then discovering that gold and lead differ in precise microstructure.'[8] At the early stages of scientific inquiry into identity conditions, the best we can hope for is *compatibility* of conditions – not identity. That is, with gold, for example, we know that most if not all atomic structures are candidates for identification with gold. What is relevant to candidacy is compatibility, not sameness, since we are trying to discover identities – that is, we cannot know the identity conditions until we know which properties or characteristics are compatible.

What happens when one does not know the identity conditions for a thing? Polger mentions the grade-school pottery project in illustration. Imagine that your child arrives with a painted lump of clay: 'Look what I made!' Polger asks, 'Is it an urn, you wonder? A bowl? The notorious *ashtray*? If the identity conditions involve shape it might be one thing; if they involve function, another; and so on. It seems that one must know the

identity conditions (generally specified) for this kind of thing in order to begin to guess what it is.'[9]

And now we get to the crux of Polger's argument. Let's say we are considering the type identity of two things – widgets and thingamajigs. Polger writes,

> If we know that their identity conditions are incompatible, then we know that they cannot be identical. ... And if their identity conditions are compatible as far as we know, then – as far as we know – they are candidates for identity. ... Likewise, if we do not know how to individuate either widgets or thingamajigs, or both ... then as far as we know they are candidates for identification – for as far as we know their identity conditions are compatible. ... If we are uninformed about the identity conditions for some widgets and thingamajigs, then even if they are identical it might seem that they could fail to be, that it could be otherwise. Thus arises the appearance of contingency.[10]

This is the formation of Polger's epistemic model for explaining the contingency of a posteriori identities. More specifically, with respect to the scientific identities which concern Kripke and the identity theorists, imagine (what was obviously the case) that before the dawn of molecular chemistry, the identity conditions for water were not thought to be incompatible with its being either an element or a molecule, and let's assume that water could have been an element. It follows, then, that water could fail to be the molecule H_2O. But, after the identity conditions for water were more precisely specified, it turns out that water could not be other than H_2O. 'The identity conditions for water are identical to the identity conditions for H_2O, and they are incompatible with any other stuff. Having discovered that the identity conditions for water and H_2O are identical, the apparent contingency of "water is H_2O" is eliminated.'[11]

As we can see, Polger's epistemic model for explaining the contingency of a posteriori identities utilizes this idea that some identity claims may appear contingent – when they are not – because we are ignorant of the relevant criteria of identity to apply. And, Polger says, 'This will often be the case if two things are candidates for identity, but are not known to be identical or non-identical. Candidacy – compatibility of identity conditions (generally specified) – thus provides a second model for explaining away the apparent contingency of putative identities; call it the *candidacy model*.'[12]

4 The Candidacy Model and Kripke's Challenge

Can the candidacy model save the type-materialist from Kripke? Polger thinks so, and here's how. If true, the identity 'pain = CFF' is necessarily true, though this relationship appears contingent. According to Polger's

candidacy model, this is due to the fact that we do not know that pains and CFFs have the same principles of individuation and, therefore, we might wrongly believe that they have different identity conditions. According to Polger, 'This would be the case if we do not know the precise identity conditions for brain states, or if we did not know the precise identity conditions for sensations. I am suggesting that we do not know the precise identity conditions for either – neither for sensations nor for brain processes.'[13] Here I will concentrate on Polger's contention that we do not understand the identity conditions of brain states. I will have nothing to say about the contention that we are ignorant of the identity conditions of sensations, and my argument, I submit, will go through just as easily without saying anything about such identity conditions.

4a Our Ignorance of the Identity Conditions of Brain States

Polger boldly states: 'We do not know how to individuate brain states, properties, processes, events, and so forth. Not only do we not know how to individuate these things, we don't really even have a clue what such things are.'[14] Polger contends that the term 'brain state' is a place holder for something we hope someday to understand. There is, according to Polger, a 'thin' notion of a brain state that we have no problem understanding. That is, that there are brains – physical objects – which have properties at times. But according to Polger, this notion is of little use in the neurosciences: 'The philosophical notion of a brain state is that of a mechanism that will play an explanatory role in a science of brains. When philosophers talk about brain states they are gesturing toward a thicker notion of a brain state – one that calls for a substantial neuroscientific theory. And neuroscientists don't – yet – have much to say.'[15] The neuroscientists are lost, according to Polger. They just do not know what they are talking about. And,

> I do not know how to thoroughly defend the negative existential claim I am making: that as a matter of empirical fact at this time, we do not know the identity conditions for brain states, properties, processes, events, and such. How does one defend the claim that some putative bit of knowledge is not now had? The best I can do is to consider the only candidate I know of for our knowledge about brain states: brain-imaging studies.[16]

But it turns out that brain-imaging studies don't deliver us identity conditions of brain states. Here's the idea. It is commonly thought that we can pick out brain states indirectly through the imaging studies produced by high-tech devices such as fMRIs or PET scans. Polger says, 'The images are clear and distinct, the stuff of science fiction. And the spatial and temporal resolution of imaging techniques is improving almost daily. Aren't those

functional magnetic resonance imaging (fMRI) pictures, for example, images of brain states? No. We must be careful about what we take those wonderful images to show.'[17] Polger is not saying that we make the mistake of pointing to certain coloured portions of MRI films and saying 'That's me thinking about *X* at 2.43 today.' No, strictly speaking that thing to which we point is a plastic film. (We're not *that* silly.) But we *do* think that those coloured portions of the MRI film *represent* my thinking about *X* at 2.43 today – the brain state that I am in during the scan. But Polger disagrees. MRI scans are, he says, 'big-picture' wide-view techniques, and he thinks that this picture is simply 'too big' because it is already highly interpreted. How is this?

> The genius of fMRI is that, through careful experimental design and sophisticated signal-processing algorithms, we can get a picture indicating which regions of the brain are, statistically, differentially active during specific kinds of mental tasks. The resulting images show those regions for which the activation is statistically correlated with the task. Analyzing the fMRI data into those pretty pictures we see takes some serious mathematics. But then we face a tough question: What is the 'object' that an fMRI image depicts?[18]

Obviously, the implication is that the 'object' depicted is not the state itself but a small amount of localized brain activity. Only a small part of the total activity of the brain is represented by the clear and distinct image on the film.

And it won't help to go 'small picture' by concentrating on a single or a few neurons by putting electrodes directly into brain tissue. This is because they do not tell us what is going on in other relevantly causally related parts of the brain. As Polger says, 'One or a few isolated neurons or ensembles do not a brain state make.'[19] In short, Polger concludes that brain scans are unable to supply the needed information about brain-state individuation. And since brain-scan imagining studies provide us with the best candidates for brain-state individuators, Polger's sceptical message is clear: we do not know how to individuate brain states and, hence, our apparent ability to imagine neural and phenomenal states existing apart does not, in fact, provide us with any evidence of their genuine contingency.

5 Objections

5a Polger's Irrational Commitment to Type-Materialism

Essentially, Polger has argued that if it is denied that we have knowledge of the identity conditions of brain states, then we make room for doubt that pain and CFF are non-identical. *But we also make room for doubt that pain and CFF are identical.* How can Polger be a committed type-materialist – one who believes that pain is identical to CFF – if he claims

to be *ignorant* of the identity conditions of brain states? Once you've adopted this scepticism, on what possible grounds could you maintain your type-materialism? Let's consider an analogy. Imagine a racist whose prejudice is based on the idea that, say, whites are superior to non-whites, and that different races should not inter-marry, or be educated together, etc. Why? 'Because', this racist might say, 'race-mixing is unnatural. Mother Nature did not intend these racially different groups to intermingle. It isn't natural.' Notice that this racist assumes that *race* is a real division in nature. Now, what is to come of this racism when it is discovered that race is, in fact, *not* a real division in nature? Race, we might discover, is not a real *kind*. This doesn't mean that there are not genetic underpinnings or causes for the superficial characteristics that some point to when dividing persons into racial categories (for example, skin colour). But hair and eye colour have genetic underpinnings too, and we do not divide blue-eyed or red-haired people into different *races* solely because of their eye or hair colour. What if this racist persists in his separatist idea? Isn't this the paradigm of an irrationally held belief? Polger, I submit, is in the same position as our irrational racist. Shouldn't Polger be sceptical about whether type-materialism is true, given that we haven't a clue how to individuate brain states? Why would a sceptic about the identity conditions of x and y believe that x and y are identical? Polger's belief in type-materialism is, I submit, irrational.

5b *On the Identity Conditions of Brain States*

Polger has argued that brain-imaging studies cannot deliver us identity conditions. Even if this is true, it must be noted that brain-imaging studies play only a *tiny* part in a unified theory about brain states. Here's a proposal: The identity conditions of brain states are supplied by a very broad theory T, and T will describe more than just the observational *effects* of brain-state activation on fMRI films and the like. T is a huge conjunction of data concerning the effects and the causes of brain states. Essentially, what we will do is 'Ramsify' brain states, and this Ramsifying process will also deliver us the identity conditions of brain states. So the Ramsey sentence of brain states, 'theory T'', is a long conjunction of all the truths about brain states – truths concerning CFF's explanatory role in the relevant sciences (neurology, psychology, clinical studies, evolution, etc.) *as well as* their observational *consequences* (fMRI/PET scan films). I submit that the identity conditions for CFFs are well understood simply because we understand that T is the theory of CFF and CFFs are explanatorily and causally relevant entities. But Polger has only mentioned *one* of the conjuncts in the Ramsified CFF theory – imaging studies – which clearly makes up only a *tiny* part of the CFF theory which delivers us its identity conditions.

6 Concluding Remarks

Polger has replaced Kripke's epistemic model with his own, but at a huge price. Kripke's model allows us to explain away the perceived contingency of many necessary a posteriori identities. Kripke's model will not, however, allow us to explain away the contingency surrounding mind–brain identity statements. Though he speaks favourably of Kripke's model being able to explain the contingency of a wide variety of scientific identities, apparently Polger doesn't pause to ask whether his own epistemic model can not only explain the contingency of mind–brain identity statements but *also* explain the contingency of all those scientific identities explained by Kripke's model. But Polger's epistemic model is unable to explain the contingency of these familiar scientific identities, and it is easy to see why. Polger seems to accept a very myopic and idiosyncratic view about how we come to understand the identity conditions of theoretical entities. For example, what if geneticists genuinely believed that the identity conditions for genes were determined *solely* by the way DNA is presented on the display screen of an electron microscope? Certainly Polger would admit that in order to obtain these identity conditions we must *also* include data from evolutionary biology, population studies, etc. – whatever theoretical and applied disciplines require the explanatory power of DNA. It is puzzling, then, why Polger would not accept this view with respect to the identity conditions of brain states. To put it another way, Polger is flirting with a total scepticism about science, since science must make reference to theoretical entities and the identity conditions of theoretical entities are obtained through the grand theories in which they provide and, in turn, receive support. Polger's candidacy model, as we have seen, is untenable on multiple grounds.

Northwest Arkansas Community College, USA

Notes

1 Thomas Polger, *Natural Minds* (Cambridge, MA: MIT Press, 2004).
2 The same is true of sensations, though I will have nothing to say about their identity conditions here.
3 Saul Kripke, *Naming and Necessity* (Cambridge, MA: Harvard University Press, 1972/1980), p. 146.
4 Polger, p. 43.
5 Strictly speaking, we know that water could not have turned out to be anything other than H_2O. Sidelle's point is innocuous: we would have been willing to entertain the possibility of *discovering* that water was something other than H_2O.
6 Polger, p. 45.
7 Ibid., p. 46.
8 Ibid.

9 Ibid., p. 48.
10 Ibid., p. 49.
11 Ibid.
12 Ibid., p. 50.
13 Ibid., p. 51.
14 Ibid.
15 Ibid., p. 52.
16 Ibid.
17 Ibid.
18 Ibid., p. 55.
19 Ibid.

International Journal of Philosophical Studies Vol. 19(4), 603–617

Critical Notice

The Undeconstructed Sovereign

A critical notice of Jacques Derrida, *The Beast and the Sovereign, Volume I.* Trans. Geoffrey Bennington. Series editors: Michel Lisse, Marie-Louise Mallet and Ginette Michaud. University of Chicago Press, 2009. Pp. xvi + 352. ISBN 978–0–2261–4428–3. £24.00/$35.00 (hbk).

The Beast and the Sovereign is the first in a planned series from the University of Chicago Press making available the seminars of Jacques Derrida.[1] The thirteen sessions reproduced here took place between 2001 and 2002 (Volume II will cover sessions from 2002–3). Its appearance, and the further volumes it heralds, might be considered a philosophical event of the first order: indeed, a blurb nominates Derrida, along with Heidegger and Wittgenstein, as one of a trio of the century's most important and influential philosophers. The book arrives at almost the same moment as the first major biography of the thinker; at a door-stopping 736 pages, Benoît Peeters' *Derrida* (Flammarion, 2010) is unlikely to have any competitors for some time to come.[2] As they appear, these lectures will no doubt enhance Derrida's reputation and provide considerable aid to students of his published work, much as has happened with the published volumes of Michel Foucault's Collège de France courses. A website (http://derridaseminars.org/) has been set up for the project, which gives details of this and forthcoming titles. Derrida once emphasised, apropos of Nietzsche, that 'the name, to be distinguished from the bearer, is always and a priori a dead man's name, a name of death'.[3] The Derrida seminar project ensures that new titles will continue to be published under the name of this particular dead man far into the future, and is to be welcomed.

No indication is given in the introduction as to why this particular seminar was chosen to initiate the series – it is a decision of the editors, not taken in consultation with Derrida before his death, though it was dictated in part by the practical matter of beginning with post-1988 seminars which exist in electronic form, rather than typescript or handwritten material (p. x). We may guess, however, that it was selected for its particular relevance to a set of contemporary debates.

International Journal of Philosophical Studies
ISSN 0967-2559 print 1466-4542 online © 2011 Taylor & Francis
http://www.tandfonline.com
http://dx.doi.org/10.1080/09672559.2011.608567