CONCEPTUALISING PHYSICAL CONSCIOUSNESS

James Tartaglia

Philosophy Programme, SPIRE

Keele University

Staffordshire ST5 5BG, United Kingdom

Tel: +44 (0)1782 734 315 Fax: +44 (0)1782 733 592

j.tartaglia@keele.ac.uk

Abstract

Theories which combine physicalism with phenomenal concepts abandon the phenomenal irrealism characteristic of 1950s physicalism, thereby leaving physicalists trying to reconcile themselves to concepts appropriate only to dualism. Physicalists should instead abandon phenomenal concepts and try to develop our concepts of conscious states. Employing an account of concepts as structured mental representations, and motivating a model of conceptual development with semantic externalist considerations, I suggest that phenomenal concepts misrepresent their referents, such that if our conception of consciousness incorporates them, it needs development. I then argue that the Phenomenal Concept Strategy (PCS) of a purely cognitive account of the distinction between phenomenal and physical concepts combines physicalism with phenomenal concepts only by misrepresenting physical properties. This is because phenomenal concepts carry ontological commitment, and I present an argument to show the tension between this commitment and granting ontological authority to physical concepts only. In the final section, I show why phenomenal concepts are more ontologically committed than PCS theorists can allow, revive U.T. Place's notion of a 'phenomenological fallacy' to explain their enduring

appeal, and then suggest some advantages of functional analyses of concepts of conscious states over the phenomenal alternative.

Keywords

Consciousness, Physicalism, Phenomenal Concept Strategy, Identity Theory, Conceptual Development, Misrepresentation

1. 1950s Physicalism vs. 1990s Physicalism

U.T. Place's 'Is Consciousness a Brain Process?' was primarily addressed to 'logical', or as would now be said 'conceptual' objections to physicalism. Having previously remarked that he considered these objections 'no greater than the logical objections which might be raised to the statement "lightning is a motion of electrical charges"' (Place 1954: 255), Place now proceeded to defend his 'identity theory' as the scientific hypothesis that statements about consciousness and brain processes converge on the same referents *a posteriori*. The key to his theory was what J.J.C. Smart later called 'topic-neutral analysis' (Place 1956: 49; Smart 1959: 149), the idea being that senses of terms for sensations are fixed extrinsically by the circumstances in which they typically occur rather than by anything intrinsic to them, so that we conceptualise only a 'something' recognised from its circumstance of occurrence. This explained how we might grasp concepts of sensations without realising they have physical referents, the explanation being that we are conceiving brain processes under contingent modes of presentation based on their causal role; by a contingent mode of presentation, I mean a way of conceiving a referent based on its contingent features, the causal role of a brain process being a contingent feature of it.

Place's conception of the identity theory as an empirical hypothesis and of concepts of conscious states as topic-neutral are two of the most distinctive features of 1950s physicalism. Since the 1990s, however, many physicalists have taken their lead from Brian Loar's 'Phenomenal States' (Loar 1990 / 1997) by retaining the view that physicalism must be defended on the basis of *a posteriori* identifications[1], but abandoning the view that concepts of conscious states refer under topic-neutral modes of presentation. In fact, according to the 'Phenomenal Concepts Strategy' (henceforth PCS)[2] our concepts of conscious states are *phenomenal concepts*, and as such do not refer under contingent modes of presentation at all; as Loar originally put it, 'phenomenal concepts conceive physical-functional properties "directly", that is, not by way of contingent modes of presentation' (Loar 1997: 602).

PCS theorists wanted to dispense with contingent modes of presentation because they were persuaded by Nagel and others that physicalism must hold there is 'something it is like, intrinsically, to undergo certain physical processes' (Nagel 1974: 445-6). As such they thought our concepts of conscious states needed to be based on something essential and intrinsic to those states, their 'what-it's-like' or phenomenal properties, rather than contingent and extrinsic features. And the reason PCS theorists felt able to dispense with contingent modes of presentation was they had an alternative explanation for the *a posteriori* status of physicalist identities, namely that mental and physical concepts have different conceptual roles, with their independence deriving from the different circumstances in which they are formed and deployed rather than any difference in what they conceptualise. This conceptual role account is often presented as a neglected alternative, with its origins credited to Nagel, who in a

footnote to ‘What Is It Like to Be a Bat?’ appealed to distinct types of imagination rather than contingent modes of presentation to explain the apparent contingency of mental-physical relations (Papineau 1993: 117; Hill 1997: 65; Hill & McLaughlin 1999: 447-8; McLaughlin 2001: 321).

This ‘neglected alternative’ narrative suggests that the only reason 1950s physicalists invoked contingent modes of presentation was to explain the conceptual independence of mental and physical concepts, and that they unfortunately chose the wrong explanation. However what was important about topic-neutral analysis was not simply that it explained conceptual independence, but that it did so without commitment to the existence of phenomenal properties, understood as introspectively known, intrinsic properties of conscious states such as painfulness and phenomenal-red. Place rejected phenomenal properties as ‘mythological’ (Place 1956: 49), and when Smart wrote that sensations are ‘described indirectly in material object language, not in phenomenal language, for there is no such thing’ (Smart 1959: 151), he was in effect rejecting phenomenal concepts as they are envisaged by PCS. In short, Place and Smart held that sensations exist but that phenomenal properties do not: this phenomenal irrealism was the third distinctive feature of 1950s physicalism.

PCS’s talk of ‘phenomenal concepts’, then, is not simply a terminological variation on earlier talk of concepts of sensations, but rather marks a substantive doctrinal commitment. The commitment is to phenomenal concepts and properties, where a phenomenal concept is a concept used to think about a phenomenal property *qua* phenomenal property, that is, *qua* a subjective, ‘what-it’s-like’ property of an experience known in introspection. As such phenomenal concepts purport to refer to

phenomenal properties, just as the concept BEAR, for example, which we use to think about bears *qua* bears, purports to refer to bears; the qualification 'purports' leaves open the possibility that reference fails. These characterisations seem broad and uncontroversial enough to apply to any of the numerous competing theories of phenomenal concepts that have been proposed, and allow PCS to be characterised in a similarly all-accommodating manner as the view that phenomenal concepts refer to phenomenal properties, that phenomenal properties are physical properties, and that this physicalist identification is philosophically unproblematic.[3]

The basic idea of PCS, then, is that physicalism can accommodate phenomenal properties without generating philosophical difficulties. Thus Loar thinks physicalism is free to 'take the phenomenology at face value' by holding that 'the property of *its being like this* to have a certain experience is nothing over and above a certain physical-functional property of the brain' (Loar 1997: 602). This stands in stark contrast to 1950s physicalism, since topic-neutral analysis was specifically designed to explain how we might conceptualise conscious states without commitment to phenomenal properties; the guiding assumption was that if physicalism was to stay, phenomenal properties had to go. Between 1950s physicalism and PCS, then, a major shift in physicalist thinking had taken place, namely from phenomenal irrealism to phenomenal realism, and an underlying metaphilosophical difference emerges if we grant that an acceptance of phenomenal properties is the intuitive default, just a matter of taking the phenomenology 'at face value' as Loar puts it. The difference is that while 1950s physicalists were trying to develop the way we think about consciousness by finding an alternative to phenomenal concepts, PCS accepts them as an inviolable

datum and then tries to work the properties they commit us to into a physicalist framework.

I shall be siding with the 1950s physicalists: if thinking of consciousness is thinking of something physical, our concepts need development. This is because *contra* the PCS view, the phenomenal properties posited by phenomenal concepts cannot be accommodated by physicalism.[4] Given that phenomenal concepts are used to refer to conscious experiences, then, and also that physicalism is true, the best conclusion to draw, I shall argue, is that phenomenal concepts misrepresent the experiences they refer to: they attribute phenomenal properties to experiences, but as Place recognised, these properties are 'mythological'. Thus phenomenal concepts misrepresent their physical referents by attributing to them properties they do not possess; phenomenal concepts can only be combined with physicalism by misrepresenting physical states. As such our concepts of experiences, which are best construed as incorporating phenomenal concepts as elements, need to be reworked and developed in line with the demands of physicalism to remove these elements.

If I am right that physicalism cannot accommodate phenomenal properties, and that the phenomenal concepts that posit these properties are misrepresentations, physicalism's return to phenomenal concepts was a mistake. This message has strong contemporary resonances. Michael Tye (2009) and Derek Ball (2009), for instance, have both argued along very similar lines to each other that physicalists should abandon phenomenal concepts, since concepts of consciousness are not a unique kind of concept which can only be possessed once the relevant subjective experience has been had.[5] An even stronger affinity can be seen in Derk Pereboom's suggestion that

introspective representations may fail to accurately represent their referents by ascribing to them qualitative natures they do not in fact possess; Pereboom calls this the 'qualitative inaccuracy hypothesis' (Pereboom 2011: 14) and in effect I shall be defending a version of it. However, the affiliation with Pereboom ends there, because he does not want to abandon phenomenal concepts or to reject the phenomenal properties they posit, and it is integral to my view that we do both.[6] Moreover, I think a rejection of phenomenal concepts and properties needs to be accompanied by historical awareness of why we should mistrust them if the same mistakes are not to be repeated, which is a matter the above authors barely touch upon.[7]

The concept of mind originated alongside the concept of breath and went through radical transformations in the Greek, medieval, and modern periods; like other rich and interesting concepts it is a thoroughly contingent deliverance of history. Given that dualism was central to this history, and that physicalist theories of consciousness have persistently met with intuitive opposition, I think there is good reason to suspect taking the phenomenology 'at face value', as Loar advocates, involves thinking about consciousness through the lens of dualist theory. In short, I think there is good reason to suspect that phenomenal concepts are dualist concepts. If this old historicist point were fully taken on board, and it has been urged many times (e.g. Ryle 1949, Matson 1966, Rorty 1979, Dennett 1991), philosophers might think twice about devising ingenious reconciliations of phenomenal concepts with physicalism such as PCS.

At the beginning of his history of the concept of mind, Paul S. Macdonald writes,

> The history of the concepts of mind and soul is a complex and twisted network of many paths, each path strewn with obstacles, dead ends, false or hidden beginnings, relapses into old ways of thinking and forward leaps of imaginative projection. (Macdonald 2003: 1)

PCS, it seems to me, is one of the 'relapses': its advocates misread anti-physicalist intuitions as an insight into the immutable nature of our concepts, when in reality they are simply a reflection of our intellectual heritage. My argumentative strategy from now on will be to first demonstrate the attractions of my misrepresentation view over the PCS alternative (in sections II and III) before arguing directly for the misrepresentation view and against PCS by showing that physicalism cannot accommodate phenomenal properties (in section IV) and then defending Place's reasons for thinking there are no such properties (in section V).

2. Developing our conception of consciousness

Concepts develop. Conceptual development takes place in individuals, as for instance when a child's concepts of love and money develop; to learn that love involves commitment or that money is a medium of exchange is to develop the way you think about love or money. Conceptual development also takes place in communities. Thus our concepts of the sun, air, and justice have all significantly developed in the last couple of millennia, so that it is now widely known that the sun is a star, the air is a composite of gases, and trial by ordeal is not just. Since most concepts are deferential, whether or not there are exceptions (see Loar 1990a), individual conceptual development is generally to be understood as a matter of the individual acquiring a

greater knowledge of the application conditions for the concept-type endorsed by those members of the community regarded as experts. Communal conceptual development on the other hand requires the concept-type endorsed by experts to itself change.

I shall follow most philosophers, psychologists and cognitive scientists in regarding concepts as mental representations, such that to possess the concept CAT is to possess a mental representation of a certain type that allows us to recognise and think about cats by employing the representation within more complex representations, such as beliefs and other propositional attitudes (cf. Laurence and Margolis 1999). So long as concepts are structured representations, which is another orthodoxy I shall follow, this model is conducive to the idea of conceptual development, which can be understood as the process of building up or revising the internal structure of the representation.[8] Thus a child whose concept WHALE comprises BIG and FISH, and who then learns that whales are not fish, can be thought of as revising the internal structure of their representation by replacing FISH with other concepts such as AQUATIC and MAMMAL. Conceptual development within a community can be similarly understood as revisions taking place to the internal structure of the concept-type approved by experts.

Individuals may possess concepts even though their grasp of them is inaccurate or vague in comparison to experts, where by the 'grasp' of a concept I understand the degree of conformity an individual shows to the application conditions endorsed by experts. Thus I can employ ELM to think about elm trees even if I cannot tell the difference between elms and beeches (Putnam 1970), and I can employ ARTHRITIS to think that I have arthritis in my thigh even though arthritis is found only in the joints

(Burge 1979). Any attempt to deny this and hold that these are cases in which I lack the concepts ELM and ARTHRITIS would face familiar obstacles. For if I lacked the concepts I could not form beliefs incorporating them, in which case the expert and I could not both believe that elms are trees or that arthritis is a disease. And neither could we straightforwardly disagree: if I say 'that's an elm' about a beech, and the expert tries to correct me, I could justifiably object that my utterance is correct given what I mean by 'elm'. Since these consequences are implausible as an account of how we actually use most if not all of our concepts, we should conclude that concepts can be shared even when the degree to which they are grasped varies.

Parallel considerations apply at the level of communities. Thus if ancient communities did not share the concepts SUN and AIR with us, they cannot have shared more basic beliefs such as that the sun shines or that we breathe the air. And neither can we straightforwardly disagree with ancient statements such as 'the sun is a deity' or 'the air is an element' if these expressed beliefs incorporating obsolete concepts. Once more, it is considerably more plausible to interpret these cases as involving shared concepts. We cannot this time speak of a differential grasp of the same concept, but communal conceptual development can instead be understood in terms of differential representational precision and accuracy: the ancient concept-type provided a vague or inaccurate representation of its referent in comparison with the concept-type endorsed by experts within our community, with both counting as developmental stages of the same concept according to a wider typing. Thus the modern concept of the sun revised aspects of the ancient concept while remaining the concept SUN, for instance.

There may be cases in which rather than saying a concept has developed, it is preferable to say a new concept has emerged, and there are also cases in which we lack clear intuitions either way; the concept of heat employed in the thermal theory of the Academy of Florence, for instance, incorporated elements of current concepts of both heat and temperature (Carey 1999: 464-5), and so might be classified as a stage in the history of the concept of heat, temperature, both, or simply as an obsolete concept. However although possible exceptions exist there are a wide range of ordinary concepts, with application conditions based on core observations and a distinctive causal role, for which Putnam / Burge considerations count decisively in favour of saying that our concepts have persisted while undergoing changes to their internal structure. And the usual reason such changes have occurred is that the concept has been found to misrepresent its referent.[9]

My suggestion, then, is that certain concepts we employ to think about states of consciousness, namely phenomenal concepts, misrepresent their referents, and so to the extent that our conception of consciousness incorporates these concepts, it is in need of communal development. This offers a simple explanation of the oft-remarked recalcitrance of dualist intuitions, namely that the concepts we have inherited for thinking about our conscious states represent them as possessed of features incompatible with physicalism, and do so because of the historical influence of dualist theory. However I also want to argue that this influence can be removed without the need to eliminate concepts of consciousness altogether; the elements which conflict with physicalism are not essential to their identity, just as the elements of the concepts of sun and air we abandoned were not essential to their identity.[10]

3. **Conceptual dislocation or 'flat Earth' thinking?**

Much light is shed on PCS by considering it in relation to a classic 'logical' objection to identifying mental and physical properties that was raised by C.D. Broad:

> Let us suppose (…) that whenever it is true to say that I have a sensation of a red patch it is also true to say that a molecular movement of a certain specific kind is going on in a certain part of my brain. There is one sense in which it is plainly nonsensical to attempt to reduce the one to the other. There is a something which has the characteristic of being my awareness of a red patch. There is something which has the characteristic of being a molecular movement. (….) [W]hether these "somethings" be the same or different, there are two different *characteristics*. (…) If this be not evident at first sight, it is very easy to make it so by the following considerations. There are some questions which can be raised about the characteristic of being a molecular movement, which it is nonsensical to raise about the characteristic of being an awareness of a red patch; and conversely. About a molecular movement it is perfectly reasonable to raise the question: "Is it swift or slow, straight or circular, and so on?" About the awareness of a red patch it is nonsensical to ask whether it is a swift or a slow awareness, a straight or a circular awareness, and so on. (Broad 1925: 622-3)

This objection forms the basis of 'Objection 3' in Smart's 'Sensations and Brain Processes', where it is used to motivate topic-neutral analysis; it subsequently became known as 'Max Black's Objection'.[11]

Broad's point is that even if particular sensations are particular brain states, there remain two sets of properties that cannot be identified, namely the properties by which we recognise something as a sensation and as a brain state. He thinks these properties cannot be identified because our concepts of them are *a priori* unconnected, which he demonstrates by the fact that it makes sense to say certain things about a brain property which it does not make sense to say about a sensation property. For Broad, then, the lack of an *a priori* connection between mental and physical concepts does not rule out the possibility of identifying mental and physical particulars, but it does rule out identifying mental and physical properties. The underlying principle accounting for this distinction seems to be that *a priori* unconnected concepts require distinct properties for the concepts to be based upon; call this 'Broad's principle'. The principle allows that we may conceptualise the same particular in ways that are not *a priori* connected in virtue of that particular possessing distinct properties. However, when the principle is applied to an empirical statement of property identity, we must either conclude that the statement is false, which is Broad's conclusion, or else invoke an additional property for at least one of the concepts to be based upon. Either way, the principle commits physicalism to distinct properties associated with mental and physical concepts.[12]

Smart implicitly accepts Broad's principle by using topic-neutral analysis to argue, in effect, that the properties by which sensations are recognised are functional rather than phenomenal, and thus compatible with physicalism. Thus in accordance with Broad's principle, Smart explains the difference between co-referring mental and physical concepts by positing distinct functional properties associated with mental

concepts. PCS, by contrast, rejects Broad's principle, albeit in a variant form such as Loar's 'Semantic Premise':

> A statement of property identity that links conceptually independent concepts is true only if at least one concept picks out the property it refers to by connoting a contingent property of that property. (Loar 1997: 600)

This is the same as Broad's principle in the crucial respect that it requires an additional property to the denoted property to explain conceptual independence, and Loar rejects it because he thinks *a priori* unconnected concepts can frame true property identifications even when both concepts refer 'directly', that is, without an additional property to provide a contingent mode of presentation.[13] Ned Block makes the same point by distinguishing 'cognitive mode of presentation' from 'metaphysical mode of presentation', with the cognitive mode understood as 'something cognitive or semantic about a representation' and the metaphysical mode understood as 'a property of the represented referent' (Block 2007: 251). Block then denies that a difference in cognitive modes entails a difference in metaphysical modes, the upshot again being that the difference between mental and physical concepts can be explained solely in terms of their different conceptual roles: cognitively rather than metaphysically, and thus without the need for extra properties.

Now since Broad's principle purports to be generally applicable, some independent motivation for its rejection is required.[14] PCS theorists typically provide this with the anti-rationalist claim that all *a priori* inferences from how we conceive the world to how the world actually is must be rejected; many critics of PCS are also sympathetic

to this cause (e.g. Levine 1983). In addition, some PCS theorists have suggested specific counterexamples unconnected to issues about consciousness. Thus Block raises the Paderewski example, in which a subject forms the false belief that there are two Paderewskis, and then forgets everything he or she once thought distinguished them, arguing that the subject possesses two distinct concepts even though 'the referent is the same and every property associated by the subject with these terms is the same' (Block 2007: 266). There are simply two distinct 'mental files' (Perry 2001: 52) with no corresponding difference of associated properties.

Let us concede that PCS did find a loophole to respond to Max Black's Objection, since Broad's principle is false: a subject may possess *a priori* independent concepts without the conceptual distinction tracking any distinction of properties. Nevertheless if a subject does possess such concepts this strongly suggests they are in error. This is, after all, the situation in the Paderewski case, in which the subject is mistakenly conceptualising the world as if there were two Paderewskis. Moreover it seems generally reasonable to assume that if a subject possesses *a priori* independent concepts, then either the concepts are associated with distinct properties, or else some kind of error is involved. Thus if we were to encounter an alien culture with two independent concepts both unfamiliar to us, we would naturally assume some property distinction was responsible for the distinction, but if we then satisfied ourselves there was not, it is hard to see what else we might conclude except that the culture was conceptually confused. Concepts represent, so if there is no distinction in the properties of the represented referent to account for a distinction of *a priori* independent concepts, the distinction seems like nothing more than cognitive clutter.

The usual explanation for a distinction of *a priori* independent concepts without any corresponding distinction of properties, then, is error. This kind of error is easy to dispel in the Paderewski case, for learning that there are not really two Paderewskis should, if the subject is rational, be enough to immediately collapse the conceptual distinction. The case of consciousness is quite different, however, because physical and phenomenal concepts apparently refer to distinct properties, and this impression remains even for those convinced the properties are identical. Although PCS can explain the independence of these concepts without the need for a distinction of properties then, some further explanation is needed of why this conceptual doubling-up does not constitute error, and why the impression of a distinction remains even when we believe there is none.

PCS's distinctive explanation for how non-erroneous conceptual doubling-up is possible is to appeal to our hard-wiring: we both perceive and introspect, and since these faculties are cognitively independent we find ourselves with *a priori* independent perception- and introspection-based concepts. This picture is then often completed by saying that the special nature of phenomenal concepts generates a 'cognitive illusion' (Tye 1999) accounting for the 'intuitive pull of dualism' (Papineau 2002: 6). Loar's original illusion-hypothesis, on which there are many variations, was that since phenomenal concepts refer directly, we expect them to reveal rather than simply refer to the essences of their referents, and thereby fail to appreciate that 'there can be two conceptually independent "direct grasps" of a single essence' (Loar 1997: 609). Thus our possession of directly referring introspection-based concepts in addition to perception-based physical concepts generates the

persistent but mistaken conviction that the extra concepts capture the essences of extra, non-physical properties.

One clear consequence of this view, namely that dualist intuitions are hard-wired, is highly implausible given that few would dispute that Descartes' conception of consciousness was a conceptual innovation of the 17th century. But even if we put these qualms to one side it seems that PCS really ought to be regarded as an unusually unattractive option for physicalists, despite the enthusiasm with which it is typically endorsed. For firstly, it gives up on the project of overcoming dualist intuitions and grants Descartes' obsolete theory an eternal lease of life. Secondly, it abandons any prospect of making conceptual sense of physical consciousness in the way we can make sense of other *a posteriori* identities, thereby making it a unique case we must simply learn to accept.[15] And thirdly, positing a cognitive dislocation between phenomenal and other concepts renders them more or less useless, thereby raising the question of how they ever came to be hard-wired in the first place. For without detailed empirical knowledge of which phenomenal properties are identical to which physical properties we could never be in a position to make inferences between phenomenal and non-phenomenal concepts; PCS must consequently appeal to non-phenomenal, 'psychological' concepts (Chalmers 1996) to explain the evident integration of introspective awareness with the rest of our reasoning about the world.[16]

A more plausible and constructive explanation of intuitive resistance to physicalism is that when we employ phenomenal concepts we are thinking like dualists and thereby misrepresenting consciousness; we can after all conceive of consciousness like a

dualist even if physicalism is true, rather as we can conceive of the Earth as flat even though it is roughly spherical. However for all that has been said so far, PCS remains an option, if a highly unattractive one; it is only when we reflect further on *how* we represent consciousness with phenomenal concepts that it becomes clear that these concepts must be misrepresentations if their referents are physical.

4. An argument for misrepresentation

Thus far I have motivated my misrepresentation thesis firstly by showing its fit with a plausible account of conceptual development, and secondly by demonstrating its advantages over the alternative PCS view of phenomenal concepts. I shall now argue directly both for the misrepresentation thesis and against PCS, by making the case that if we think about our experiences phenomenally and those experiences are physical, then we must be misrepresenting them. In making this case, it is enlightening to return to Kripke's famous argument against the identity theory, as Loar did in his original presentation of PCS, since the historical subtext of my argument is that the reason PCS is untenable is that it tried to accommodate a Kripkean, and ultimately Cartesian, intuition. Moreover returning to Kripke allows us to complete the account of the intellectual genesis of PCS begun in the last section, since I suggest as a plausible hypothesis that the context in which PCS was conceived as a loophole to Max's Black's Objection was that of responding to Kripke. It was through pursuing this loophole and thereby letting Kripke set the terms of the debate, it seems to me, that phenomenal concepts returned to mainstream physicalism through the back door.

The pivotal claim in Kripke's argument was that, 'in the case of mental phenomena there is no "appearance" beyond the mental phenomenon itself' (Kripke 1972: 341). This idea is central to PCS and explains its insistence on the directness of phenomenal concepts. Thus Loar, for instance, says that 'a phenomenal concept has as its mode of presentation the very phenomenal quality it picks out' (Loar 1997: 604) and hence that its reference is not 'phenomenally mediated' (ibid.: 608), and Hill and McLaughlin say that, '[w]hen one uses a sensory concept to classify one's own current experiences, the experiences that guide and justify one in applying the concept are always identical with the experiences to which the concept is applied' (op. cit.: 448). These are just different ways of making Kripke's point, namely that we do not conceptualise the underlying reality of a conscious experience indirectly on the basis of its appearance, in the way we conceptualise the underlying reality of a natural kind like gold indirectly on the basis of its golden appearance, because in the case of conscious experiences the appearance *is* the reality being conceptualised. For our concepts to capture the intrinsic, essential nature of conscious experience then, they must refer directly to the subjective appearance itself, rather than indirectly to something else for which the subjective appearance might provide a contingent mode of presentation.

It is this distinctive feature of phenomenal concepts that Kripke uses as the basis of his argument against the 'identity theory', by which he means the PCS position that phenomenal properties are identical to physical properties; the Place / Smart theory, which rejects phenomenal concepts, is summarily dismissed before the well-known argument begins.[17] The argument is that the 'identity theory' cannot explain away the apparent contingency of phenomenal-physical relations, since the standard

explanation for an illusion of contingency is unavailable. The standard explanation appeals to a sensation acting as 'an intermediary between the external phenomenon and the observer' (op. cit.: 339), but since phenomenal concepts refer directly, and hence there are no sensory intermediaries between phenomenal concepts and their referents, the appearance of contingency established through conceivability must be veridical, and so phenomenal and physical properties cannot be identical.

One way to short-circuit this argument is to reject phenomenal concepts, which was the move made by Place and Smart's identity theory. But for physicalists committed to phenomenal concepts an alternative explanation is required. Very few options are available. One would be if the physical concept referred under a contingent mode of presentation based on the perceptual appearance of the physical property (Boyd 1980), but this invokes an additional sensation that would subsequently need to be provided with a physicalist account, thereby simply postponing the explanation.[18] Another is if each of the concepts were based on distinct essential properties of a shared physical referent, with the apparent contingency explained by our lack of understanding of how these properties unite in a single essence. But this would mean that our understanding of physical nature is currently inadequate; this kind of position, defended by Nagel (1998) and Strawson (2003), is anathema to mainstream physicalists who are trying to demystify consciousness with physicalism, not mystify physicalism with consciousness. And the third option, discovered by Loar, is to explain the apparent contingency as due to the cognitive disparateness of perception- and introspection-based concepts, which becomes compatible with these concepts providing 'two conceptually independent "direct grasps" of a single essence' (Loar 1997: 609) once Broad's principle and its cognates are abandoned.[19]

PCS provides a loophole in Kripke's argument then, to show that phenomenal concepts can be combined with physicalism after all, and without the need to go down the Nagel / Strawson route. However in order for physicalists to be able to hold that phenomenal concepts provide a 'grasp of essence', as Loar puts it (ibid.: 608-9), they first need to be stripped of any ontological significance, given that physical concepts alone, according to physicalists, represent the essential nature of physical properties. Loar recognised this requirement and consequently distinguished the grasp of essence provided by a physical concept, which 'structurally analyses the property' and 'reveal[s] how it is internally constituted' (ibid.: 608-9), from the grasp of essence provided by phenomenal concepts.[20] The latter counts as a 'grasp of essence' solely because phenomenal concepts 'do not conceive their references by way of their accidental properties', and are thus direct, and Hill clarifies this crucial component of PCS by saying that the theory makes only the 'purely negative claim' that phenomenal properties are 'self-presenting' (Hill 2002: 887), and hence not presented by distinct, reference-mediating properties.

The main problem with PCS, however, is that the concepts it adopted from Kripke's avowedly 'Cartesian' argument (op. cit.: 334) are loaded with ontological significance: if you conceptualise conscious experience in this way you will inevitably move beyond a 'purely negative claim' and represent it as having a nature physicalism cannot accommodate (unless extended in the Nagel / Strawson manner; hereafter I omit this qualification). Kripke was half wrong and half right, then. On the one hand, physicalism could be true regardless of what we can or cannot conceive, and if consciousness is indeed physical this does not prevent us thinking about it with

phenomenal concepts. But on the other hand, Kripke rightly detected that there is a conflict here, for although physicalists can embrace phenomenal concepts they do can so only by resigning us to misrepresentation.

The reason phenomenal concepts must be misrepresentations if physicalism is true can be seen as follows:

(1) Pain = c-fibre stimulation
(2) Pain = the introspective appearance of pain
(3) Therefore, c-fibre stimulation = the introspective appearance of c-fibre stimulation

(1) and (2) are applications of core PCS commitments: (1) is the kind of physicalist type-identity claim PCS typically endorses and (2) is a consequence of conceiving sensations such as pain phenomenally. However (3), I shall argue, is incompatible with physicalism, and the best explanation of this incompatibility is that phenomenal concepts are misrepresentations.

There are both intensional and extensional readings of 1-3 and to make these explicit we must make reference to the concepts employed as well as their referents. The intensional reading is:

(1a) The referent of PAIN = the referent of C-FIBRE STIMULATION
(2a) The referent of PAIN = the introspective appearance of pain-qua-pain

(3a) Therefore, the referent of C-FIBRE STIMULATION = the introspective appearance of pain-qua-c-fibre stimulation

The extensional reading is:

(1b) The referent of PAIN = the referent of C-FIBRE STIMULATION

(2b) The referent of PAIN = the introspective appearance of the referent of PAIN

(3b) Therefore, the referent of C-FIBRE STIMULATION = the introspective appearance of the referent of C-FIBRE STIMULATION

On the extensional reading 1-3 is valid, and shows that if (1) pain is c-fibre stimulation, and (2) we represent pain with a phenomenal concept, such that we identify pain with the introspective appearance of pain, as both Kripke and the PCS theorists do, then (3) c-fibre stimulation must be the same as the introspective appearance of c-fibre stimulation. This conclusion is incompatible with physicalism, however, because the concept of c-fibre stimulation, to which physicalists must grant sole ontological authority, has nothing to do with introspective appearance; a scientist would not have to mention introspection in order to explain what c-fibre stimulation is. Thus the physicalist who wants to accept the kind of type-identity claims made by (1) while rejecting (3) must also reject (2); generalised this amounts to a rejection of phenomenal concepts. This rejection is motivated because the argument shows that if pain is c-fibre stimulation and we represent pain with a phenomenal concept, we thereby represent c-fibre stimulation as essentially something it is not: an introspective appearance. As such, phenomenal concepts must be misrepresentations.

This conclusion needs clarification, but we first need to establish that PCS is committed to the extensional reading 1b-3b. To see this, consider the following, which assumes Venus always looks blue in the mornings when conceptualised as Phosphorus, and red in the evenings when conceptualised as Hesperus:

(4) Hesperus = Phosphorus

(5) The visual appearance of Hesperus is always red

(6) Therefore, the visual appearance of Phosphorus is always red

There are both intensional and extensional readings of this argument depending on whether our interest is in Venus under its two different modes of presentation or just Venus itself. The intensional reading can be made explicit as follows:

(4a) The referent of HESPERUS = the referent of PHOSPHORUS

(5a) The visual appearance of Venus-qua-Hesperus is always red

(6a) Therefore, the visual appearance of Venus-qua-Phosphorus is always red

The premises are true but the argument is invalid because the conclusion is reached by substituting into an intensional context: (5a) makes a claim not simply about the referent of HESPERUS, namely Venus, but about Venus as conceptualised as Hesperus, from which no conclusion can be drawn about Venus as conceptualised as Phosphorus. But there is also an extensional reading:

(4b) The referent of HESPERUS = the referent of PHOSPHORUS

(5b) The visual appearance of the referent of HESPERUS is always red

(6b) Therefore, the visual appearance of the referent of PHOSPHORUS is always red

This argument is valid, but (5b) and (6b) are false. This is not the natural reading of 4-6, the intensional reading is, since we ordinarily use HESPERUS to refer only to Venus in its evening aspect. But there is nevertheless a clear sense in which if Hesperus always looks red, then since Hesperus is Phosphorus, Phosphorus must always look red too, just as there is a clear sense in which if Superman can fly, then Clark Kent can fly too.

If we now turn back to the original argument 1-3, we see that whether it is read intensionally or extensionally depends on whether (2) is interpreted as a claim about pain as conceptualised in a certain way, namely as pain or as c-fibre stimulation, or rather as a claim about the property itself irrespective of how it is conceptualised. On the former, intensional reading (2a), pain is identified with the introspective appearance of pain-qua-pain, from which it would be invalid to infer anything about the introspective appearance of pain-qua-c-fibre stimulation. On the latter, extensional reading (2b), pain is identified with the introspective appearance of the referent of PAIN, from which a valid inference is then made about the introspective appearance of the referent of C-FIBRE STIMULATION.

Only the extensional reading is available to PCS, however, since PCS holds that PAIN is not associated with any distinct, reference-mediating properties, but rather directly conceives the essence of its referent, namely a kind of introspective

appearance called ‘pain’. For to conceive pain phenomenally is to conceive it as a kind of *phenomenon* or appearance, which is why Kripke and the PCS theorists deny any distinction between pain and the appearance of pain: they think pain is a kind of appearance. This commitment goes further than just claiming that pain has a certain kind of appearance as one of its contingent aspects, one it might lose while remaining pain, and it even goes further than claiming that a certain kind of appearance is necessary to pain but not all there is to it. For both these weaker claims allow pain and the appearance of pain to be distinguished; they adopt a phenomenal conception of one aspect of pain, not pain itself, and thus introduce distinct, reference-mediating properties for PAIN of the kind PCS theorists go out of their way to deny, it being the *raison d'être* of the view to combine such a denial with physicalism.

Given this commitment, then, PCS cannot appeal to distinct aspects of pain, akin to the Hesperus and Phosphorus aspects of Venus, to lay claim to an intensional reading of (2) that would invalidate the argument. And in any case, it seems clear that the Kripke-intuition expounded by (2) is not the claim that pain is the introspective appearance of pain-qua-pain, but rather that pain is the introspective appearance of pain itself, however conceptualised. For the claim that pain and the introspective appearance of pain are the same property only makes sense if pain is a kind of introspective appearance, one we might indifferently call ‘pain’ or ‘the introspective appearance of pain’. But if pain is a type of introspective appearance, then this is a fact about the property itself, not about the property as conceptualised in one way rather than another. The property could not, for instance, fail to be an introspective appearance when perceived on a neuroimaging device and conceptualised as c-fibre

stimulation, as Venus might fail to be red when perceived in the morning and conceptualised as Phosphorus.

On the only reading available to PCS, then, the argument is valid, and shows that if there is no difference between pain and the introspective appearance of pain, then within a physicalist framework there can be no difference between c-fibre stimulation and the introspective appearance of c-fibre stimulation. However we already know there is a difference because the physical concept does not represent the property as an introspective appearance. So if conceptualising c-fibre stimulation as a phenomenal property entails that being a certain type of introspective appearance is the essence of c-fibre stimulation, we must conclude either that the physical concept of c-fibre stimulation is inadequate since it fails to capture the essence of its referent, or that conceptualising c-fibre stimulation as a phenomenal property misrepresents it since being an introspective appearance is not the essence of c-fibre stimulation; the former is Nagel and Strawson's view, and the latter is my own.

PCS cannot respond that the physical concept of c-fibre stimulation tells us both what c-fibre stimulation and the introspective appearance of c-fibre stimulation is, but simply fails to conceptualise it as the latter, because it is committed to the claim that introspective appearance is the essence of c-fibre stimulation. For while it is true that the physical concept of gold, for example, tells us both what gold and the ancient world's most valued metal is without conceptualising it as the latter, the reason this does not show that the physical concept fails to capture the essence of gold is that being the ancient world's most valued metal is not the essence of gold. If it were, the physical concept would indeed fail to capture this essence. Similarly, PCS cannot

simply say that the scientific description completely describes c-fibre stimulation without mentioning its introspective appearance, even though it does have an introspective appearance that can only be conceptually entertained via phenomenal concepts. This is because if introspective appearance is the essence of pain, and hence of c-fibre stimulation, then introspective appearance would have to be mentioned in any complete description of c-fibre stimulation. If it is not mentioned, the description cannot have captured its essence and so cannot be complete.

5. Where phenomenal thinking goes wrong

To think about a sensation with a phenomenal concept is to focus in introspection on the 'immediate phenomenal quality' (Kripke 1972: 340) or the 'property of *its being like this* to have a certain experience' (Loar 1997: 602), and thereby conceive of it as a type of subjective appearance. The problem with trying to combine this phenomenal conception of consciousness with physicalism, however, is that it leads physicalists to inadvertently start thinking of physical properties like c-fibre stimulation as a kind of subjective appearance: they peer introspectively at 'that' subjective appearance and believe themselves to be aware of c-fibre stimulation. This apparent conception of c-fibre stimulation, however, tells us something essential about the property that we could not glean from its physical concept, and it is this latter concept alone which physicalists must grant ontological authority.

PCS tries to avoid this conflict by offering accounts of phenomenal concepts drained of any potentially problematic content; this is why Loar construes phenomenal concepts as demonstrative concepts, and Hill insists that PCS is making a 'purely

negative claim'.[21] But this is a lost cause, because once we have started thinking phenomenally about consciousness, the commitments that cause the problems for physicalism have already been made no matter what denials are subsequently issued. For even if phenomenal concepts do have a demonstrative component they must also ascribe the property of being phenomenal, such that in thinking about redness phenomenally we do not conceptualise the redness as simply 'that', but rather as 'that subjective appearance' as opposed to 'that objective property of the object I am perceiving'. And the point generalises, for even conceptualising a pain as 'that pain' must ascribe the property of being phenomenal in order to secure the intended reference to a subjective appearance, rather than the bodily location in which the pain is felt.[22] This minimal interpretation is all that is required to generate the conflict with physicalism, however, for once I have conceived of a property phenomenally, I have a rival and very definite conception of what it is: it is what it subjectively appears to be.

What inclines us to think of conscious properties as subjective appearances? Part of the answer is surely that dualist theory, which conceptually and ontologically separates subjective appearances from objective reality, has been deeply entrenched within philosophical literature and thought for nearly four hundred years, making the phenomenal conception our accustomed way of thinking about consciousness. As such Nagel and Kripke needed only to issue reminders about this conception to gain the intuitive high-ground over physicalists trying to develop an alternative. However there is another more substantive explanation for the enduring appeal of phenomenal concepts that was proposed by the 1950s physicalists, and which can be illustrated as follows.

Suppose I close my eyes tightly and have an experience of a red patch against a black background. If physicalism is true, my subjective awareness of the patch must be identical to a physical state I am in; this is simply token physicalism.[23] However it does not follow that the properties I can apparently pick out in virtue of being in this state are properties of the state I am in. I can apparently pick out the redness from my first-person perspective, but not my subjective awareness of the redness, and it is only the property of being a subjective awareness of redness that has any claim to being a conscious property. If I focus on the redness and think of it as a conscious property, that is, if I form a phenomenal concept of the redness, then I have prised it apart from my subjective perspective. It will then seem that I am confronted with a distinct kind of property since I am conceiving the property as a conscious property rather than as the redness of a perceived object.[24] But phenomenal properties are neither properties of conscious states nor properties those states make us aware of: they are fictions dualists invented a non-physical reality to accommodate. Phenomenal concepts are formed, then, when we conceptualise the properties we are subjectively aware of as conscious properties, but in isolation from our subjective awareness; this is the thought-process that explains the appeal of dualism, but it is not a cognitive illusion, simply a philosophical mistake.[25]

The upshot of this unduly neglected analysis is that phenomenal properties are mistakenly thought to exist both because we misconceive perceived properties as subjective properties, and also because in cases where there is no perceived property, such as the above example in which my eyes are shut, we misconceive states in which we have only apparent awareness of perceived properties as states in which we have actual awareness of subjective properties. If this is right, then so is Place's conclusion

that phenomenal properties are ‘mythological’, and we might build on it by saying that phenomenal concepts misrepresent conscious states by ascribing phenomenal properties to them, a conclusion fully in accordance with the anti-rationalist stance that we cannot infer metaphysical matters such as whether phenomenal properties exist from the way we conceptualise the world.

Nevertheless, even if there are no phenomenal properties and phenomenal concepts are consequently misrepresentations, there seems little sense in denying that in thinking about pain as possessed of the phenomenal property of painfulness, for instance, you can thereby succeed in referring to pain. The best way to explain this is to say that phenomenal concepts are a dispensable part of the internal structure of our concepts of conscious states. Thus just as FISH might be an erroneous component of the internal structure of a child’s concept WHALE, the current suggestion is that PHENOMENAL-PAIN or PHENOMENAL-RED might be erroneous components of the internal structure of people’s concepts of their conscious experiences of pain and redness. This still leaves the question, if reference to conscious states is not secured in virtue of phenomenal properties, of what are the properties which do in fact secure reference. And here the best suggestion remains the one originally made by Place and Smart, and subsequently developed by Armstrong, Lewis, and others, namely that concepts of conscious states are at root concepts of the functional properties of those states.

Functional analysis claims that the only properties of conscious states of which we need be aware in order to conceptualise them as conscious states are functional properties. Thus my awareness of having a conscious experience of redness is, at a

minimum, my awareness that I am in the same type of conscious state I am typically in when perceiving red. But that does not mean the redness is to be identified with a functional property, or indeed with any property at all, for when my eyes are closed there is no redness: the reason it seems this way is that I recognise my present state in the same way I would recognise a conscious state of perceiving red, namely on the basis of its functional properties. Functional analyses thereby seek to explain phenomenology in terms of shared functional properties rather than shared phenomenal properties. This has three major advantages: it is compatible with physicalism, it can distinguish the representational function of concepts of conscious states from that of co-referring physical concepts, and it raises no obstacle to integrating reasoning about conscious states with the rest of our reasoning.

I am not suggesting that functional analyses of conscious states are without their problems.[26] Rather what I am suggesting is that they were abandoned prematurely in favour of a return to phenomenal concepts, and that the resulting quest to find some way of satisfactorily combining physicalism with phenomenal concepts should be abandoned. If phenomenal concepts are abandoned, however, as they were once before in the 1950s, then this time it needs to be done with historical awareness if the same mistakes are not to be repeated later on down the line. For if the phenomenal component within the concept of consciousness becomes recognised as an inheritance of dualist theory, then no physicalist should even want to revive it, however intuitive it might seem. Maybe then the process of conceptual development, which is bound to be counterintuitive at first, might begin to acquire a new resilience.[27]

References

Ball, D. (2009) 'There Are No Phenomenal Concepts', *Mind*, 118: 935-962.

Block, N. (2007) 'Max Black's Objection to Mind-Body Identity', in T. Alter and S. Walter (eds.) *Phenomenal Concepts and Phenomenal Knowledge*, pp. 249–306. Oxford: Oxford University Press.

Borst, C. (ed.) (1970) *The Mind-Brain Identity Theory*, London: Macmillan.

Boyd, R. (1980) 'Materialism without Reductionism: What Physicalism Does Not Entail', in N. Block (ed.) *Readings in the Philosophy of Psychology (vol. 1*, pp.67-106). Cambridge, MA: Harvard University Press.

Braddon-Mitchell, D. and Jackson, F. (2007) *The Philosophy of Mind and Cognition*, 2nd edn., Oxford: Blackwell.

Broad, C.D. (1925) *Mind and its Place in Nature*, London: Routledge.

Burge, T. (1979) 'Individualism and the Mental', *Midwest Studies in Philosophy*, 4: 73-122.

Carey, S. (1999) 'Knowledge Acquisition: Enrichment or Conceptual Change?', in S. Laurence and E. Margolis (eds.) *Concepts: Core Readings*, pp. 459-488. Cambridge, MA: MIT Press.

Chalmers, D. (1996) *The Conscious Mind*, Oxford: Oxford University Press.

Chalmers, D. (1999) 'Materialism and the Metaphysics of Modality', *Philosophy and Phenomenological Research*, 59: 473-496.

Dennett, D. (1991) *Consciousness Explained*, Boston: Little, Brown and Company.

Fodor, J. (1998) *Concepts: Where Cognitive Science Went Wrong*, Oxford: Oxford University Press.

Hill, C. (1997) 'Imaginability, Conceivability, Possibility, and the Mind-Body Problem', *Philosophical Studies*, 87: 61-85.

Hill, C. (2002) ‘Review of *Purple Haze: The Puzzle of Consciousness*, by Joseph Levine’, *Mind*, 111: 882-8.

Hill, C. (2009) ‘The Identity Theory’, in T. Bayne, A. Cleeremans, and P. Wilken (eds.) *The Oxford Companion to Consciousness*, pp. 359-363. Oxford: Oxford University Press.

Hill, C. and McLaughlin, B. (1999) ‘There are fewer things in reality than are dreamt of in Chalmers’ philosophy’, *Philosophy and Phenomenological Research*, 59: 445-454.

Horgan, T. and Tienson, J. (2001) ‘Deconstructing New Wave Materialism’, in C. Gillett and B. Loewer (eds.) *Physicalism and Its Discontents*, pp. 307-318. Cambridge: Cambridge University Press.

Kripke, S. (1972) ‘Naming and Necessity’, in G. Harman and D. Davidson (eds.) *Semantics of Natural Language*, pp. 253-355, 763-769. Dordrecht: Reidel.

LaPorte, J. (2004) *Natural Kinds and Conceptual Change*, Cambridge: Cambridge University Press.

Laurence, S. and Margolis, E. (1999) ‘Concepts and Cognitive Science’, in S. Laurence and E. Margolis (eds.) *Concepts: Core Readings*, pp. 3-82. Cambridge, MA: MIT Press.

Law, S. (2004) ‘Loar’s Defence of Physicalism’, *Ratio*, 17: 60-7.

Levine, J. (1983) ‘Materialism and Qualia: The Explanatory Gap’, *Pacific Philosophical Quarterly*, 64: 354-361.

Levine, J. (1998) ‘Conceivability and the Metaphysics of Mind’, *Noûs*, 32: 449-80.

Lewis, D. (1980) ‘Mad Pain and Martian Pain’, in N. Block (ed.) *Readings in the Philosophy of Psychology* (*vol. 1*, pp. 216-222). Cambridge, MA: Harvard University Press.

Loar, B. (1988) 'Social Content and Psychological Content', in R. Grimm and P. Merrill (eds.) *Contents of Thoughts*, pp. 99-110. Tucson, AZ: University of Arizona Press.

Loar, B. (1990a) 'Personal References' in E. Villanueva (ed.) *Information, Semantics and Epistemology*, pp. 117-133. Oxford: Basil Blackwell.

Loar, B. (1990 / 1997) 'Phenomenal States', in J. Tomberlin (ed.) *Philosophical Perspectives, 4: Action Theory and Philosophy of Mind*: 81-108. Revised version (1997) in N. Block, O. Flanagan and G. Güzeldere (eds.) *The Nature of Consciousness*, pp. 597-616. Cambridge, MA: MIT Press.

MacDonald, P. (2003) *History of the Concept of Mind*, Aldershot, UK: Ashgate.

Matson, W. (1966) 'Why Isn't the Mind-Body Problem Ancient?', in P. Feyerabend and G. Maxwell (eds.) *Mind, Matter and Method: Essays in Philosophy and Science in Honor of Herbert Feigl*, pp. 92-102. Minneapolis: University of Minnesota Press.

McGinn, C. (1989) 'Can We Solve the Mind-Body Problem?', *Mind*, 98: 349-66.

McGinn, C. (2001) 'How Not to Solve the Mind-Body Problem', in C. Gillett and B. Loewer (eds.) *Physicalism and Its Discontents*, pp. 284-306. Cambridge: Cambridge University Press.

McLaughlin, B. (2001) 'In Defense of New Wave Materialism: A Response to Horgan and Tienson', in C. Gillett and B. Loewer (eds.) *Physicalism and Its Discontents*, pp. 319-330. Cambridge: Cambridge University Press.

Nagel, T. (1974) 'What is it like to be a bat?', *Philosophical Review*, 82: 435-450.

Nagel, T. (1998) 'Conceiving the Impossible and the Mind-Body Problem', *Philosophy*, 73: 337-352.

Papineau, D. (1993) *Philosophical Naturalism*, Oxford: Blackwell.

Papineau, D. (2002) *Thinking about Consciousness*, Oxford: Oxford University Press.

Pereboom, D. (2011) *Consciousness and the Prospects of Physicalism*, Oxford: Oxford University Press.

Perry, J. (2001) *Knowledge, Possibility, and Consciousness*, Cambridge, MA: MIT Press.

Place, U.T. (1954) 'The Concept of Heed', *British Journal of Psychology*, 45: 234-55.

Place, U.T. (1956) 'Is Consciousness a Brain Process?', *British Journal of Psychology*, 47: 44-50.

Place, U.T. (2002) 'A Pilgrim's Progress? From Mystical Experience to Biological Consciousness', *Journal of Consciousness Studies*, 9: 34-52.

Polger, T. (2011) 'Are Sensations Still Brain Processes?', *Philosophical Psychology* 24: 1-21.

Putnam, H. (1970) 'Is Semantics Possible?', *Metaphilosophy*, 1: 187- 201.

Rorty, R. (1979) *Philosophy and the Mirror of Nature*, Princeton: Princeton University Press.

Ryle, G. (1963) *The Concept of Mind*, Harmondsworth, UK: Penguin.

Smart, J.J.C. (1959) 'Sensations and Brain Processes', *Philosophical Review*, 68: 141-156.

Stoljar, D. (2005) 'Physicalism and Phenomenal Concepts', *Mind and Language*, 20: 469-94.

Strawson, G. (2003) 'Real Materialism', in L. Antony and N. Hornstein (eds.) *Chomsky and his Critics*, pp. 49–88. Oxford: Blackwell.

Tye, M. (1999) 'Phenomenal Consciousness: The Explanatory Gap as a Cognitive Illusion', *Mind*, 108: 705-725.

Tye, M. (2009) *Consciousness Revisited: Materialism without Phenomenal Concepts*, Cambridge, MA: MIT Press.

White, S. (1986) 'Curse of the Qualia', *Synthese*, 68: 333-68.

White, S. (2007) 'Property Dualism, Phenomenal Concepts, and the Semantic Premise', in T. Alter and S. Walter (eds.) *Phenomenal Concepts and Phenomenal Knowledge*, pp. 210–248. Oxford: Oxford University Press.

Notes

[1] Loar remains neutral on whether the identifications are with physical or functional properties; I shall simply refer to physical properties for brevity of exposition.

[2] I adopt this label from Stoljar 2005; others include 'New Wave Materialism' (Horgan and Tienson 2001) and 'Conceptual Dualism' (Papineau 2002; Hill 2009).

[3] Although a very broad characterisation, I am unaware of any positions able to satisfy it that would not intuitively be counted as a version of PCS. PCS theorists have various accounts of why physicalist identifications are not problematic drawing on their various accounts of the nature of phenomenal concepts, but appeals to the *directness* of phenomenal concepts, which will be my focus, are the most typical.

[4] Unless the concept of physicalism is given a revolutionary extension; I discuss and reject this kind of view in Section IV.

[5] Ball denies that there are any phenomenal concepts, since our concepts do not conform to the characterisations given by phenomenal concept advocates. However since people do think phenomenally about consciousness this formulation seems unnecessarily paradoxical. Tye is more equivocal: he toys with the claim that there are no phenomenal concepts (op. cit.: 39), and indicates in the subtitle of his book that he rejects them, but his considered view is simply that phenomenal concepts 'do not stand apart from all other concepts' (ibid.: 63). My view is that our concepts of *conscious states* 'do not stand apart' since like other ordinary concepts they secure their reference by functional role, but that phenomenal concepts, which misrepresent those states, encourage the view that conceptualising consciousness is unique in requiring introspective acquaintance with phenomenal properties.

[6] Pereboom does not reject phenomenal properties since he thinks they could be identified with the physical properties which are the normal causes of experiences. Thus according to this 'dual-content view', phenomenal concepts succeed in representing phenomenal properties while nevertheless misrepresenting them as having exhaustively qualitative natures (op. cit.: 29-43). Against the objection that it is a conceptual truth that phenomenal properties have exhaustively qualitative natures, Pereboom argues, in effect, that phenomenal concepts may undergo conceptual development (ibid.: 38-40). It seems to me, however, that concepts like PHENOMENAL-RED are philosophical constructions rooted in dualist theory which are best abandoned.

[7] Tye ends his book by saying that phenomenal concepts are among the 'vestiges of Cartesianism' that must be 'eliminated from the materialist worldview' (Tye 2009: 199), and so does seem to agree with the line I am taking here, although he does not elaborate.

[8] Just as I shall not be considering concept atomism, neither shall I be considering the view, also associated with Fodor (e.g. Fodor 1998), that epistemological matters such as coming to know more about whales are orthogonal to facts about what concepts like WHALE represent. To engage such views would take us too far afield and so my ambition in this paper is limited to the claim that for those who reject them, the link I am making between conceptual development and representation should be attractive, and offers a plausible diagnosis of philosophical debates about consciousness.

[9] There may be other reasons, however. For example, when jadeite was first imported into China, the Chinese distinguished it from their traditional jade (nephrite), but ultimately decided both should count equally as jade (LaPorte 2004: 94 & ff.). This

was a change to the concept JADE not because it had misrepresented nephrite, but because it now needed to include jadeite.

[10] Conceptual development requires striking a balance between conceptual continuity and current theoretical needs. Place, for instance, judged that behaviourism had tipped the balance too far in latter direction by replacing the current reality of conscious experience with mere dispositions; the identity theory was specifically designed to correct this shortcoming (Place 2002).

[11] In correspondence, Smart confirmed that he used this passage when writing his paper (the examples are the same); the reason he credited Black was that he was the first to publically raise this kind of objection to the Place / Smart theory.

[12] For an argument that the properties associated with concepts of 'qualitative states' must either be topic-neutral or non-physical, see White 1986.

[13] The difference is that Loar's principle requires at least one of the properties to be contingent, which lets in possible counterexamples in which the concepts are based on distinct essential properties; Law (2004: 63) defends these counterexamples while White (2007: 228-32) rebuts them.

[14] Hill and McLaughlin claim that the principle only fails to apply to 'the joint exercise of sensory and physical concepts' (1999: 448-9) due to their distinct conceptual roles, but as Chalmers points out, this restriction is unmotivated since any pair of distinct concepts would have distinct conceptual roles; see Chalmers 1999: 486.

[15] McGinn's view that the cognitive dislocation between concepts formed on the basis of introspection and perception prevents us from making sense of the 'psychophysical nexus' is surprisingly similar (McGinn 1989: 352); the key difference is that McGinn thinks an alien subject with different cognitive faculties *could* make sense of it.

[16] None of these considerations count decisively against the view that phenomenal concepts are hard-wired, of course, and empirical evidence for hard wiring might yet undermine my view that we can develop our concept of experience away from phenomenality. However the main thesis of this paper, namely that phenomenal concepts misrepresent, could be true even if they are hard-wired, and given that PCS rather than hard-wiring is my primary target, and that no PCS theorist defends their theory on the basis of evidence for hard-wiring, I think it is enough for my purposes to point out that PCS's commitment to the hard-wiring of phenomenal concepts has problematic consequences, and that the history of the concept of mind makes it *prima facie* implausible; after all, as Wallace Matson pointed out many years ago, it is far from clear that even the Greeks had phenomenal concepts (Matson 1966). In section V below I shall argue directly for the view that we can overcome dualist intuitions by recognising the false reasoning that commits us to phenomenal properties; if dualist intuitions can be dismantled in this way, this would provide an additional and much stronger consideration against hard-wiring.

[17] David Lewis noted this (1980: footnote 2). Kripke's reason for discounting Place and Smart's theory is that since we can imagine any brain state existing without its typical causal role, it has the absurd consequence that a particular pain might have existed without being a pain (op. cit.: 335-6); anti-rationalists who reject *a priori* inferences from how we conceive the world to how it is should dismiss this reasoning as perfunctorily as Kripke used it to dismiss the real identity theory.

[18] See McGinn 2001: 293-4; most PCS theorists, but not all (Tye 1999: 722-3), grant that we can conceive of physical properties 'neat', as Block puts it (2007: 271).

[19] Working within the framework of Kripke's argument, PCS looks like an ingenious neglected alternative, and it is surely no coincidence that Loar, who was writing about Paderewski-type cases in the late 1980s (Loar 1988), sets up his rejection of the 'Semantic Premise' in the context of Kripke's argument (Loar 1997: 599-600).

[20] The details of how a physical concept can 'reveal' the essence of its reference, which according to Loar involves an element of *a priori* analysis, are controversial and not relevant to the present argument: all that matters here is that the grasp of essence provided by physical concepts is supposed to be ontologically significant, and that provided by phenomenal concepts is not.

[21] Nagel and Strawson pursue the opposite tactic: they drain physicalism of content by granting ontological authority to a future physics we know almost nothing of.

[22] This is why phenomenal concepts must be more than 'demonstrative arrows shot blindly that refer to whatever they hit', as Joseph Levine once rightly observed (Levine 1998: 467).

[23] Externalists about consciousness would have to add the stipulation that the subject be appropriately embedded in a physical environment. I shall not discuss the separate possibility of 'extended mind' theories in this paper.

[24] I am not presupposing any particular position on the nature of colours like redness here. However I do hold that colours are objective properties for which an appearance / reality distinction can be made; to describe a book as red is to ascribe an objective property to it, even if this turns out to be a highly disjunctive property, and any such property can appear to perceivers other than it is. As such the problems phenomenal properties create for physicalism do not arise for colours and other perceived properties.

[25] I am elaborating Place's idea of a 'phenomenological fallacy' (Place 1956: 49), and Smart's point that the identity theory does not claim that an 'after-image [e.g.] is a brain-process, but that the experience of having an after-image is a brain-process' (Smart 1959: 150). The latter is usually taken as merely a rejection of mental objects (e.g. Braddon-Mitchell and Jackson 2007: 99-100), but the point applies equally to phenomenal properties.

[26] See Borst (ed.) 1970 for some classics, although see also Polger 2011 for reasons to think these problems may not be as acute as is often thought.

[27] I would like to thank Tim Crane, Peter Fletcher, Rob Hopkins and Nils Kurbis for written comments on earlier drafts, Darragh Byrne for many conversations about PCS, as well as audiences at Birmingham, Keele, Oxford Brookes and Sheffield.