

In C. Gillett & B. Loewer, eds., *Physicalism and Its Discontents* (Cambridge University Press, 2001)

DECONSTRUCTING NEW WAVE MATERIALISM

Terence Horgan and John Tienson

University of Memphis.

In the first post World War II identity theories (e.g., Place 1956, Smart 1962), mind-brain identities were held to be contingent. However, in work beginning in the late 1960's, Saul Kripke (1971, 1980) convinced the philosophical community that true identity statements involving names and natural kind terms are necessarily true and furthermore, that many such necessary identities can only be known a posteriori. Kripke also offered an explanation of the a posteriori nature of ordinary theoretical identities such as that water = H₂O. We identify the kinds and substances involved in theoretical identities by certain of their contingent properties. What we discover when we discover a theoretical identity is the underlying nature of the kind that we identify by those contingent properties.

Now, of course, it was being a posteriori, not being contingent, that mattered to the identity theorists anyway, so the necessity of identity is not, in itself, damaging to mind-brain identity theories. However, Kripke also argued persuasively that the alleged mind-brain identities could not be treated in the same way as ordinary theoretical identities. We "identify" pain by feeling it, and surely how it feels is an essential property of pain, not a contingent property. Thus, a mind-body identity theory must provide a different explanation of why its identities are a posteriori.

A new wave of materialists has appeared on the scene with a new strategy for explaining the a posteriori nature of its alleged identities.¹ The strategy is to locate the explanation for the a posteriori nature of mind-body identities, not on the side of the world, but on the side of the mind--in different ways of thinking about or imagining, or in different concepts. Thus, on this new view, there is only one property—this brain process type, which is identical with this pain type—but we conceive of it under two different concepts, one phenomenal, one theoretical.² And these concepts are of such different types that it is not possible to know a priori that they are concepts that pick out the very same thing, and furthermore, it is not surprising that it is not possible to know this a priori.

We believe that this on-the-side-of-the-mind strategy is self defeating. As far as we can see, differences on the side of the mind of the sort that the new wavers invoke imply different properties on the side of the world. At any rate, the new wavers have not given us an account of the intentional connection by virtue of which a concept or way of conceiving is of one property rather than another. Giving such an account in a way that avoids commitment to different properties is essential to their purposes.

We begin, in section 1, by briefly articulating the problem to which the new wave solution is directed. In section 2 we lay out what we take to be the essentials of the solution the new wavers offer. In section 3 we explain why, on our view, the new wavers have not given an account of the intentionality of concepts that is sufficient for their purposes, and why it appears to us unlikely that such an account can be given. In section 4 we illustrate the problem with reference to a representative version of new wave materialism, Loar (1997).

1. Conceivability Arguments.

Various kinds of conceivability considerations, such as the conceivability of inverted spectra, have been urged against physicalism for a long time. In the post-Kripke era, conceivability arguments have taken a new form, particularly in the hands of David Chalmers (1995, 1996, in press). However, Chalmers' formulations of the argument typically depend upon a fairly elaborate semantic apparatus. We here offer a minimalist formulation of the conceivability issue to which new wave identity theories are addressed. Let us use '*M*' for a particular mental state type (say a particular type of pain), and '*P/F*' for the physical or physical/functional state type to which *M* is allegedly identical. Consider the following argument.

- (1) It is conceivable that *P/F* exist and *M* not exist (or vice versa).
- ∴ (2) It is possible that *P/F* exist and *M* not exist.
- ∴ (3) *M* is not (identical with) *P/F*.

Now, at one time a physicalist might have responded by bringing up the following, apparently parallel argument.

- (1w) It is conceivable that water exist but H₂O not exist.
- ∴ (2w) It is possible that water exist but H₂O not exist.³
- ∴ (3w) Water is not H₂O.

But of course, water is H₂O. So, the response goes, (2w) does not imply (3w); the identity of water and H₂O is contingent. And thus, we can say that the identity of *M* and *P/F* is also contingent and that (2) does not imply (3).

However, Kripke convinced us that there are no contingent identities involving proper names or natural kind terms; (2w) does imply (3w). The trouble with the water argument is that (1w) and (2w) are false; you cannot really conceive of a world with water but without H₂O. At a certain level of reflection, it might seem that you can conceive such a world, but you cannot because anything that isn't H₂O would not be the same stuff, would not be water. The reason it may seem that you can conceive of a world with water but without H₂O is that we identify water by certain of its contingent properties—look, taste, feel,

behavior. It has turned out that the stuff we identify in this way is H₂O. If it seems that you conceive of a world with water but without H₂O, what you are really conceiving is a world without water in which something other than water has the (contingent) properties by which we identify water in this world. To call that a possible situation with water but without H₂O would be to misdescribe the situation.

New wavers agree that the physicalist cannot respond to argument (1)-(3) in this way. They agree both that (2) implies (3), just as (2w) implies (3w), and that we should not deny (1). We do not identify our conscious states by their contingent properties. The basic way we have of identifying conscious states is by being in them, and the nature or essence of a conscious state is what it is like to be in it. Thus, we are not mistaken in thinking we conceive a world with *P/F* but lacking *M* (or vice versa); there are no contingent properties involved to confuse with the thing itself. A common alternative way to put the need to accept (1) is that we can conceive a world with *P/F* but lacking *M* because there is no a priori connection between phenomenal concepts and the theoretical concepts that would represent *P/F* as a physical/functional state type.

Thus, what the new wavers must say—and do say—is that (1) does not imply (2): conceivability does not imply possibility, at least in the special case where phenomenal concepts are involved. The heart of the position that we are calling new wave materialism is the account they offer of the a posteriori nature of (necessarily true) mind-brain identities, to which we now turn.

2. The Generic New Wave Position.

Generically described, and apart from differences of detail and nuance among various specific versions, the new wave position goes as follows. Phenomenal properties are identical to physical properties, in a broad sense of ‘physical’ that includes, for example, functional properties. (Following Loar 1997 we will speak of ‘physical-functional’ properties.) Phenomenal concepts, the concepts we apply to phenomenal properties on the basis of introspection, refer directly to those properties as they are in themselves; they do not refer via any contingent features of these referent-properties. Nevertheless, phenomenal concepts are distinct from the physical-functional concepts that characterize the physical-functional nature of phenomenal properties. In fact, phenomenal concepts are conceptually independent of the coreferential physical-functional concepts, since the two kinds of concepts have such different conceptual roles in human thought. Because of this conceptual independence, identity statements reflecting the coreference of phenomenal and physical-functional concepts are a posteriori. Also because of this conceptual independence, “separability scenarios” involving phenomenal and coreferential physical-functional concepts are coherently conceivable; for instance, it is coherently conceivable that physical-functional property *P/F* is instantiated without phenomenal property *M* being instantiated (or

vice versa), even though in fact, $P/F = M$; it is also coherently conceivable that could exist creatures physically just like ourselves, in an environment just like ours, whose internal states have phenomenal character that is systematically inverted relative the phenomenal character of our own experiences (“inverted qualia”); it is also conceivable that there could exist such duplicate-creatures, in a duplicate-environment, whose internal states have no phenomenal character at all (“absent qualia”). Nevertheless, both phenomenal concepts and the associated physical-functional concepts are rigid: they are coreferential not only in the actual world, but in all possible worlds. Thus, such separability scenarios are metaphysically impossible—despite being coherently conceivable, and despite the fact that the relevant psycho-physical identities are a posteriori. Conceivability does not imply possibility.

New wave materialists acknowledge three explanatory tasks that a credible version of materialism should simultaneously accomplish:

- (A) Explain the differences between phenomenal concepts and associated physical-functional concepts in a way that renders them conceptually independent, and thereby renders separability-scenarios coherently conceivable.
- (B) Explain the differences between phenomenal concepts and associated physical-functional concepts in a way that fully respects the phenomenology of conscious experience.
- (C) Explain the differences between phenomenal concepts and associated physical-functional concepts in a way that is consistent with the claim that phenomenal properties are identical to physical-functional properties.

New wavers claim to be offering a philosophical position that does meet all three of these constraints. Separability scenarios, they say, are coherently conceivable because phenomenal concepts and coreferential physical-functional concepts are so different that they are conceptually independent. Phenomenology is fully respected by acknowledging that phenomenal concepts refer directly to phenomenal properties as they are in themselves, while also acknowledging that these concepts are not physical-functional concepts. And consistency with materialism is maintained because nothing in the proffered account of phenomenal concepts precludes the possibility that the properties to which they refer are physical-functional properties. So say the new wavers.

3. A Deconstructive Argument.

But all is not well. Consider the following simple argument for dualism with respect to phenomenal properties:

Deconstructive Argument

1. When a phenomenal property is conceived under a phenomenal concept, this property is conceived otherwise than as a physical-functional property.
2. When a phenomenal property is conceived under a phenomenal concept, this property is conceived directly, as it is in itself.
3. If (i) a property P is conceived, under a concept C, otherwise than as a physical-functional property, and (ii) P is conceived, under C, as it is in itself, then P is not a physical-functional property.

Hence,

4. Phenomenal properties are not physical-functional properties.⁴

This argument is valid, and the new wave materialists are committed to premises 1 and 2. Yet premise 3 does not appear to be credibly deniable; on the contrary, it seems virtually tautologous. For, if indeed phenomenal properties, when conceived under phenomenal concepts, not only are conceived otherwise than as physical-functional properties but also are conceived as they are in themselves, then surely phenomenal properties must be otherwise than physical-functional properties. I.e., surely they must be properties that are not physical-functional.

So new wave materialism's doctrinal commitments, in combination with a tautologous-looking further claim that does not seem credibly deniable, lead inescapably to the conclusion that phenomenal properties are not physical-functional properties. This contradicts the new-wave claim that they are physical-functional properties. Thus, the argument evidently effects a deconstruction of the new wave position.

In order to fend off this deconstructive challenge, a new waver would need to articulate his or her account of phenomenal concepts in a way that makes credible the claim that premise 3, despite its initial air of tautologousness, is in fact false. So far the new wavers have not even recognized clearly that they face this dialectical burden, let alone shown how it might be met. We wish them well in trying to meet it, but we doubt that they can succeed.

We will now consider in more detail a version of new wave materialism due to Brian Loar, in order to see how it fares vis-à-vis the deconstructive argument.

4. Loar's Version.

According to Loar (1997), phenomenal properties are physical-functional properties, whereas phenomenal concepts are a species of what he calls recognitional concepts. Concerning the notion of recognitional concepts, he says:

Phenomenal concepts belong to a wide class of concepts that I call recognitional concepts. They have the form ‘x is one of that kind; they are type-demonstratives. These type-demonstratives are grounded in dispositions to classify, by way of perceptual discriminations, certain objects, events, situations.... These dispositions are typically linked with capacities to form images, whose conceptual role seems to be to focus thoughts about an identifiable kind in the absence of currently perceived instances. (pp. 600-01)

Concerning phenomenal concepts and phenomenal properties, he sketches his position as follows:

Here is the view to be defended. Phenomenal concepts are recognitional concepts that pick out certain internal properties: these are physical-functional properties of the brain. They are the concepts we deploy in our phenomenological reflections; and there is no good philosophical reason to deny that, odd though it may sound, the properties these conceptions *phenomenologically reveal* are physical-functional properties—but not of course under physical-functional descriptions.... [T]he property of *its being like this* to have a certain experience is nothing over and above a certain physical-functional property of the brain. (pp. 601-2)

As regards the conceptual independence of phenomenal concepts and the corresponding physical-functional concepts, he says:

What then accounts for the conceptual independence of phenomenal and physical-functional concepts? The simple answer is that recognitional concepts and theoretical concepts are in general conceptually independent... Concepts of the two sorts have quite different conceptual roles. It is hardly surprising that a recognitional conception of a physical property should discriminate it without analyzing it in scientific terms. (p. 602)

As so far articulated, Loar’s account addresses two of the three above-noted constraints on a credible version of materialism, (A) and (C). It addresses (A) by explaining the differences between phenomenal concepts and physical-functional concepts in a way that is intended to render them conceptually independent; and it addresses (C) with the claim that phenomenal properties are identical to physical-functional properties. And as so far articulated, the account also evades the challenge posed by the deconstructive argument, since premise 2 of that argument has not so far been embraced.

However, not enough has yet been said to satisfy condition (B), which requires the explanation of phenomenal concepts to fully respect the phenomenology of conscious experience. One way to see this is to notice the conceptual possibility of creatures who have self-directed recognitional concepts that (i) refer to physical-functional properties, (ii) by virtue of their recognitional nature are conceptually independent of the physical-functional concepts that characterize the nature of these properties, and yet (iii) are phenomenally empty. Loar himself acknowledges this fact in the following passage:

Not all self-directed recognitional concepts are phenomenal concepts.... A...fanciful self-directed nonphenomenal concept can be conceived. To begin with, consider blindsight. Some cortically damaged people are phenomenally blind in restricted retinal regions; and yet when a vertical or horizontal line (say) is presented to those regions, they can, when prompted, guess what is there with a somewhat high degree of correctness. We can extend the example by imagining a blindsight that is exercised spontaneously and accurately. At this point we shift the focus to internal properties and conceive of a self-directed recognitional ability, which is like the previous ability in being phenomenally blank and spontaneous but which discriminates an internal property of one's own. If this recognitional ability were suitably governed by the concept 'that state', the resulting concept would be a self-directed recognitional concept that is phenomenally blank. (p 603).

So the account of phenomenal properties and phenomenal concepts needs bolstering, in a way that distinguishes genuine phenomenal concepts from self-directed blindsight concepts. Loar addresses this issue in two steps. First he considers what an antiphysicist could be expected to say about it:

What might an antiphysicist say about these various self-directed recognitional concepts? Let us make a good-faith attempt to present a reasonable version.... We might say [says the antiphysicist] that a phenomenal concept has as its mode of presentation the very phenomenal quality that it picks out. We might also say that phenomenal concepts have "token modes of presentation" that are noncontingently tied to the phenomenal qualities to which those concepts point: particular cramp feelings and images can focus one's conception of the phenomenal quality of cramp feeling.... As for self-directed blindsight concepts, the antiphysicist then ought to say, they differ from phenomenal concepts in the obvious way, whether one puts it by saying that they lack the noncontingent phenomenal modes of presentation (types) that phenomenal qualities have, or that they they lack their phenomenal "token modes of presentation." (p. 604)

Second, he claims that the physicalist can, and should, say such things too:

Whatever the antiphysicist has said about these cases the physicalist may say as well. The idea that one picks out the phenomenal quality of cramp feeling by way of a particular feeling of cramp (or image, etc.) is hardly incompatible with holding that that phenomenal quality is a physical property. The contrast between phenomenal concepts and self-directed blindsight concepts... finds physicalist and antiphysicist equally able to say something sensible. (pp. 604-5)

Loar thus takes on board the claim that phenomenal concepts, unlike self-directed blindsight concepts, have modes of presentation, and also the claim that the mode of presentation of a phenomenal concept is

the very phenomenal property to which the concept refers. He thereby embraces premise 2 of the deconstructive argument, as underscored by the following further remarks:

It is natural to regard our conceptions of phenomenal qualities as conceiving them as they are in themselves, that is, to suppose that we have direct grasp of their essence.... Phenomenal concepts, as we have seen, do not conceive their reference via contingent modes of presentation. And so they can be counted as conceiving phenomenal qualities directly. Calling this a grasp of essence seems to me all right, for phenomenal concepts do not conceive their references by way of their accidental properties. (pp. 608-9)

So Loar's overall account of phenomenal concepts is now committed not only to the claim that they are self-directed recognitional concepts, but also to the claim that they differ from self-directed blindsight concepts by conceiving their referents directly, as they are in themselves.

Where does this leave us, dialectically? Loar himself clearly thinks that this further claim about phenomenal concepts can be added to the pot without disrupting the erstwhile consistency between the account of phenomenal concepts and the physicalist claim that phenomenal properties are identical to physical-functional properties. But it is just here that he runs afoul of the deconstructive argument. For, insofar as a concept C is characterized merely as a self-directed recognitional concept, the claim

(a) C refers to a physical-functional property

is consistent with the claim

(b) C conceives its referent otherwise than as a physical-functional property

because C could be a concept that picks out its referent in a nakedly recognitional way—a purely demonstrative way, involving no mode of presentation at all. But once one adds that C directly picks out its referent as it is in itself (rather than picking it out nakedly), this explanation for the consistency of (a) and (b) is no longer available. So some different account is needed for how (a) and (b) could be consistent. But Loar offers no such account, and thus he fails to defuse the force of the deconstructive argument.⁵

In short: Loar's initial, partial, account of phenomenal concepts addresses requirements (A) and (C) for an adequate version of materialism, by embracing premise 1 of the deconstructive argument without premise 2. But the initial account does not adequately address condition (B), which leads him to extend the account by embracing premise 2 as well. This expanded account now perhaps satisfies conditions (A) and (B). But, given premise 3 of the deconstructive argument, the expanded account evidently fails to satisfy condition (C). Loar has failed to meet the dialectical burden of explaining how premise 3 could be false.

We lack the space to consider in detail various other versions of new wave materialism, such as Hill (1997) and Hill and McLaughlin (in press). But in our view, these versions too fail to address the crucial question of how premise 3 could be false.

5. Conclusion.

Let us summarize this discussion by formulating in a slightly different way the challenge we are posing for new wave materialism. Consider the following general principle about concepts and properties:

- (i) If a concept C provides a direct grasp of the property P it refers to, i.e., if C conceives P directly as it is in itself (rather than conceiving P via a “mode of presentation” or “reference-fixing property” distinct from P itself), then P is as it is conceived by C.⁶

We fail to see how this principle could possibly be false. And as far as we can see, new wavers have said nothing that begins to explain how it could be false. Now consider these two claims about phenomenal concepts:

- (ii) A phenomenal concept, by phenomenologically presenting the property to which it refers, provides a direct grasp of that property.
- (iii) A phenomenal concept phenomenologically presents its referent-property otherwise than as a physical-functional property.

Both of these claims seem clearly true; and new wave materialists affirm them both.⁷ But statements (i)-(iii) jointly entail that the properties referred to by phenomenal concepts are otherwise than physical-functional properties, i.e., that phenomenal properties are not physical-functional properties.⁸

New wavers tend to emphasize, correctly, that phenomenal concepts differ significantly from physical-functional concepts in their conceptual role, and in the conditions under which their application is epistemically warranted. They also tend to emphasize, again correctly, that these kinds of differences are sufficiently great to render phenomenal concepts conceptually independent of physical-functional concepts. But these kinds of observations, correct though they are, simply don't address the question of how principle (i) could be false, and specifically how it could be false for phenomenal concepts and properties. That question is what needs to be answered; otherwise, deconstruction.⁹

References

- Chalmers, D. (1995). "The Puzzle of Conscious Experience," Scientific American 273, 80-86.
- Chalmers, D. (1996). The Conscious Mind. Oxford: Oxford University Press.
- Chalmers, D. (in press). "Materialism and the Metaphysics of Modality," Philosophy and Phenomenological Research.
- Graham, G. and Horgan, T. (2000). "Mary Mary, Quite Contrary," Philosophical Studies 99, 59-87.
- Hill, C. (1997). "Imaginability, Conceivability, Possibility and the Mind-Body Problem," Philosophical Studies 87, 61-85.
- Hill, C. and McLaughlin, B. (in press). "There Are Fewer Things Than Are Dreamt of in Chalmers's Philosophy," Philosophy and Phenomenological Research.
- Levine, J. (1998). "Conceivability and the Metaphysics of Mind," Nous 32, 449-80.
- Loar, B. (1997). "Phenomenal States," in N Block, O. Flanagan, and G Guzeldere, eds., The Nature of Consciousness, 597-616. Cambridge MA: MIT Press.
- Melnyk, A. (this volume). "Physicalism Unfalsified: Chalmers' Inconclusive Conceivability Argument."
- Kripke, S. (1971). "Identity and Necessity," in M. Munitz, ed., Identity and Individuation, 135-64. New York: NYU Press.
- Kripke, S. (1980). Naming and Necessity. Cambridge, MA: Harvard University Press.
- Place, U. T. (1956). "Is Consciousness a Brain Process?" British Journal of Psychology 47, 44-50.
- Smart, J. J. C. (1962). "Sensations and Brain Processes," in V. C. Chappell, ed., The Philosophy of Mind, 160-72. Englewood Cliffs: Prentice Hall.
- Tye, M. (1995). Ten Problems of Consciousness. Cambridge, MA: MIT Press.
- White, S. (1986). "Curse of the Qualia," Synthese 68, 333-68.

¹ Clear examples of what we are here counting as "new wave materialism" are Loar (1997), Hill (1997), and Hill and McLaughlin (in press). Versions of materialism that are somewhat similar in spirit, but do not clearly fall under this rubric, include Tye (1995), Levine (1998), and Melnyk (this volume). On the problems faced by such recent views insofar as they might fall outside the rubric, see note 5 below.

²New wavers use various terminology, sometimes speaking of modes of presentation or modes of access. For the most part we will stick with 'conceiving', 'concept of', and the like.

³Zombie worlds have become a common representative of conceivability arguments. So compare

-
- (1') A world that is physically identical to ours but lacking conscious states (a zombie world) is conceivable.
- ∴ (2') A world that is physically identical to ours but lacking in conscious states is possible.
- (1w) A world with water but without H₂O is conceivable.
- ∴ (2w') A world with water but without H₂O is possible.

(Again, the two arguments are now considered disanalogous. The intuition behind (1w) and (1w') was, of course, that the stuff we identify as water could have turned out to have some other composition—to be an element, or homogeneous.)

⁴ This argument is similar in spirit to the “property dualism argument” presented in White (1986), pp. 351-53. Note well that premise 1 does not say that phenomenal properties are conceived, under phenomenal concepts, as non-physical-functional properties. Conceiving a property otherwise than as a physical-functional property is different from, and weaker than, conceiving it as otherwise than a physical-functional property.

⁵ A dilemma looms, for those materialists who acknowledge that phenomenal concepts do not pick out their referents via a contingent feature of the referent-property. On one hand they can avoid premise 2 of the deconstructive argument, by effectively treating phenomenal concepts as nakedly referential. The trouble with this option, as illustrated by Loar’s discussion of generalized blindsight, is that it fails to respect the phenomenology of conscious experience. On the other hand they can embrace premise 2, thereby falling prey to the deconstructive argument. (Our rubric “new wave materialism” covers positions that do embrace premise 2.) Graham and Horgan (forthcoming) argue, in effect, that the first horn of this dilemma impales certain current versions of materialism, notably Tye (1995).

⁶ Premise 3 of the deconstructive argument is essentially a special case of this principle. Note that principle (i) does not say that property P is—or is in its essence—only as it conceived to be under concept C. Likewise, premise 2 of the deconstructive argument does not say that when a phenomenal property is conceived under a phenomenal concept, this property is thereby conceived as it completely is in itself. The possibility is left open that there is more to a property P, as it is in its essence, than is revealed when P is directly grasped under a concept C.

⁷ Recall what Loar says in a passage quoted earlier: “Phenomenal concepts, as we have seen, do not conceive their reference via contingent modes of presentation. And so they can be counted as conceiving phenomenal qualities directly. Calling this a grasp of essence seems to me all right.” (p. 609). Concerning the much-discussed problem of the “explanatory gap,” Loar goes on to say this: “What generates the

problem is not appreciating that there can be two conceptually independent “direct grasps” of a single essence, that is, grasping it demonstratively by experiencing it, and grasping it in theoretical terms.” (p. 609) We ourselves are among those who fail to appreciate this contention. For, given that phenomenal concepts conceive phenomenal properties directly, the contention can be true only if principle (i) is false for phenomenal properties; yet Loar has failed to explain how it could be false.

⁸ Even if certain mental properties with phenomenal content (for instance, pain) are physical-functional properties, the point would still hold. Phenomenal properties (for instance, the hurtfulness of pain) then would be non-physical properties that are possessed by the physical-functional properties themselves. Moreover, even if it should turn out that phenomenal properties have a total essence that is partly physical-functional—so that this aspect of their essence is not revealed when these properties are directly grasped under phenomenal concepts (see note 6)—nevertheless they still would be partly non-physical (since they still would be as they are conceived under phenomenal concepts, and they are so conceived otherwise than as physical).

⁹ This paper is entirely collaborative; order of authorship is alphabetical. We thank Ronald Endicott, Christopher Hill, Barry Loewer, and Brian McLaughlin for helpful discussion and/or comments.