

Karol Polcyn

The conceivability argument and the intuition of dualism

Diametros nr 24, 90-106

2010

Artykuł został opracowany do udostępnienia w internecie przez Muzeum Historii Polski w ramach prac podejmowanych na rzecz zapewnienia otwartego, powszechnego i trwałego dostępu do polskiego dorobku naukowego i kulturalnego. Artykuł jest umieszczony w kolekcji cyfrowej bazhum.muzhp.pl, gromadzącej zawartość polskich czasopism humanistycznych i społecznych.

Tekst jest udostępniony do wykorzystania w ramach dozwolonego użytku.

THE CONCEIVABILITY ARGUMENT AND THE INTUITION OF DUALISM

- Karol Polcyn -

1. INTRODUCTION

The standard assumption in the literature is that Kripke's argument against the identity theory presupposes that an identity statement can be true a posteriori only if at least one of the concepts flanking the identity sign refers contingently (Chalmers [1996]; Garcia-Carpintero and Macia [2006]; Loar [1997, 2003]; Papineau [2002]). Mind-brain identities do not satisfy this criterion and therefore they cannot be true. According to David Papineau (Papineau [2007a, 2007b]), however, this is a wrong picture of Kripke's argument which results from conflating Kripke's argument with the two-dimensional argument developed by Chalmers (Chalmers [1996, 2006a]) and Jackson (Jackson [1993, 1998]). While the a posteriority thesis mentioned above is crucial to the two-dimensional argument, it is not the basis of Kripke's argument.¹ According to Papineau, then, Kripke's argument does not lead to the metaphysical conclusion that mind-brain identity is not true. Instead, the conclusion it supports is that we cannot believe mind-brain identity at some very intuitive level. As Papineau puts it, Kripke shows that we are all in the grip of *an intuition of dualism*.

For Papineau the conclusion that we are in the grip of the intuition of dualism is an important result. It shows that there isn't any conceptual problem of understanding how mind-brain identities can be true even in the light of our theoretical evidence for their truth but only a psychological problem of disbelieving them. The alleged conceptual problem is known as the problem of the so-called explanatory gap.² This problem is supposed to be independent of the two-dimensional argument, so even if one thinks that the two-dimensional argument does not work, one still needs to come to grips with the problem of the explanatory gap and Papineau now offers a solution of that problem. The explanatory gap

¹ The role this thesis plays in Kripke's argument under its orthodox interpretation as well as in the two-dimensional argument will be clarified in the subsequent sections.

² See Levine [1983].

is simply a manifestation of our intuitive conviction that dualism is true and therefore does not pose any conceptual difficulty for physicalism.

Here I will argue that Papineau's assumption that we are all in the grip of the intuition of dualism strengthens physicalism further by leading to the refutation of the entailment from conceivability to possibility (in short, CP) which is the key premise of the version of the two-dimensional argument developed by Chalmers. I will argue that if we can assume that the intuition of dualism is consistent with the truth of materialism, that intuition provides the best explanation of why our intuitions about the separability of mind and body in other possible worlds are unreliable. Contemporary physicalists, including Papineau himself, assume that they already have a good independent way of seeing why CP is unjustified. However, in my view, this previous criticism of CP is not successful. So therefore the new objection to CP that will emerge from Papineau's interpretation of Kripke will bring an entirely new perspective to the current debate.

I will discuss the weaknesses of this previous criticism of the two-dimensional argument in section 4. In section 3 I will explain how the intuition of dualism undermines the entailment from conceivability to possibility and in section 5 I will explain more specifically how the intuition of dualism undermines the dialectic of Chalmers' version of the two-dimensional argument. But first let me give a brief summary of the orthodox interpretation of Kripke's argument and the new interpretation suggested by Papineau.

2. TWO INTERPRETATIONS OF KRIPKE'S ARGUMENT AGAINST THE IDENTITY THEORY

On the standard interpretation, Kripke's argument against materialism goes as follows:

- (1.) Mind-brain identity claim, if true, is necessary.
- (2.) The identity appears contingent.
- (3.) Unless this appearance can be explained away, mind-brain identity is possibly false.
- (4.) The appearance that mind-brain identity is contingent cannot be explained away.
- (5.) So the identity is not necessary and therefore not true.

(1.) follows from the fact that the concepts of the corresponding mental states and physical states are rigid and (2.) is something that Kripke takes for granted. Contingency is a matter of possible falsity combined with actual truth and so the claim that mind-brain identity appears contingent amounts to the claim that the identity appears possibly false even if we assume that it is true. However, one might also

think that mind-brain identity appears contingent in the sense that it appears possibly false even though conscious and physical states are perfectly *correlated* in the actual world and it is not entirely clear which of those two interpretations Kripke has in mind. In fact, he seems to speak of the appearance of contingency in both of those senses. I will come back to this issue later. The key point to emphasize here is that the appearance that mind-brain identity is contingent involves the appearance that the identity is possibly false and that under the orthodox reading of Kripke's argument *explaining* the appearance of mind-brain contingency *away* means for Kripke explaining away this appearance of the possible falsity of mind-brain identity, that is, explaining how this appearance arises consistently with assuming that mind-brain identity is necessary. Thus (3.) can be read this way: unless the appearance of mind-brain contingency can be explained away, this appearance entails that mind-brain identity is possibly false and hence not true.

Consider now the premise (4.). Kripke thinks that (4.) is true because he thinks that we could explain the apparently possible falsity of mind-brain identity away only if we could show that our impression of this possible falsity is a *misrepresentation* of what we really find conceivable. Kripke argues then that this cannot be done. Our impression of the possible falsity of mind-brain identity is not an illusion in this sense.

It is not difficult to see that Kripke's assumption that there is only possible model for explaining the appearance of mind-brain contingency away stems from assuming that conceivability entails possibility, that what is genuinely conceivable is always possible. Kripke does not endorse this entailment explicitly but that is the assumption that he is committed to by the acceptance of (4.). If conceivability entails possibility, then the conceivable falsity of mind-brain identity can be explained away in the sense of being shown to be consistent with assuming that the identity is necessary only if our impression of the conceivable falsity of mind-brain identity is a misinterpretation of what is, in fact, conceivable.

Kripke assumes that the paradigm cases of intuitions of contingency that can be explained away are the intuitions that we have with respect to theoretical identities, such as the identity of water and H₂O or the identity of heat and molecular motion. The appearance that those identities are contingent (and hence the appearance of their possible falsity) can be shown to be a result of misrepresenting what really appears possible. For example, the appearance that water might not have been H₂O is really the appearance that watery stuff might not have been H₂O. Similarly, the appearance that heat might not have been molecular motion is really the appearance that the sensation of heat might not have been caused by

molecular motion but some other phenomenon. However, there is no room for extending this kind of explanation to the mind-brain case, according to Kripke.

The explanation of the appearance of contingency in the case of theoretical identities turns on the fact that natural kind terms, such as 'water' or 'heat', refer contingently. So, for example, 'water' has its reference fixed by some *contingent* properties of water, such as being liquid, colorless, drinkable (in short, watery properties), and this explains why the appearance that water might not have been H₂O is an illusion – why we do not have the appearance that water might not have been H₂O but rather the appearance that watery stuff might not have been H₂O which we misrepresent as the former. However, this explanatory model does not work in the mind-brain case because phenomenal concepts do not refer contingently. Thus there is no appearance of pain that would be contingently related to pain. Pain is what it appears to be. So we cannot assume that the appearance that pain might not have been c-fiber stimulation, say, results from misrepresenting the appearance that some pain appearing state might not have been c-fiber stimulation. The two appearances come down to the same. So if the latter is not an illusion, the former is not an illusion, either.

In response to Kripke, physicalists typically reject the entailment from conceivability to possibility (CP) that is crucial to Kripke's argument.³ Even if Kripke is right that our intuition about the possible falsity of mind-brain identity is not an illusion in the sense of being a misrepresentation of what we find conceivable, physicalists argue that that intuition is not reliable as a guide to possibility or, at least, that it is not inconsistent with assuming that mind-brain identity is necessary. However, according to a new interpretation of Kripke's antimaterialist argument that has recently been offered by David Papineau, Kripke's argument does not depend on CP and does not lead to the conclusion that mind-brain identity is false. According to Papineau, the challenge that Kripke's argument raises is not so much to explain how the apparently possible falsity of mind-brain identity could be unreliable but rather to explain why mind-brain identity appears possibly false even to committed physicalists who *believe* that the identity is true. In other words, the challenge is to explain how mind-brain identity can appear contingent to begin with. With other identities, there is no corresponding sense of puzzlement. Thus there is no difficulty in understanding how one can think that water, say, might not have been H₂O. Given that 'water' refers contingently, via properties that are only contingently true of H₂O, the identity of water and H₂O can appear possibly false even to someone who assumes that it is true. But the

³ The details of the standard critique of CP will be discussed in section 4.

concepts of conscious experiences do not refer contingently and hence no such corresponding explanation is available in the case of mind-brain identity. So, for example, if you believe that pain is c-fiber stimulation, you are committed to holding that the identity is necessary and there is no room for the appearance that some pain-appearing state might not have been c-fiber stimulation since pain is, again, identical with what it appears to be. However, the identity of pain and c-fiber stimulation does appear possibly false even to committed physicalists who are persuaded that the identity is true at a theoretical level and this calls out for explanation.

As Papineau points out, the rejection of CP that physicalists advocate in response to Kripke's orthodox argument does not remove the difficulty raised by this new challenge. Even if we assume that mind-brain identity is necessary despite being conceivably false, this does not explain how the identity can appear possibly false to someone who believes that it is true.

There is also another standard response to Kripke's orthodox argument that does not work as a response to Kripke's new argument. To see the point of this response notice first that in the light of Kripke's analysis of theoretical identities and mind-brain identities CP has the following consequence: a statement of identity that is conceivably false is possibly false (and hence not necessary and not true) unless one of the concepts flanking the identity sign refers contingently (or has a descriptive content) (Levine [2001]; Yablo [2000]). And further, once we realize that conceivable falsity is a feature of statements that are a posteriori (not true a priori), the thesis just mentioned comes down to the following: an identity statement can be true a posteriori only if at least one of the concepts flanking the identity sign refers contingently. So this is the thesis that Kripke is committed to (Chalmers [1996]; Garcia-Carpintero and Macia [2006]; Loar [1997, 2003]; Papineau [2002]) and then the challenge he raises for physicalists can be interpreted as the challenge to explain how mind-brain identity statements can be true a posteriori assuming that both concepts flanking the identity sign in those statements refer directly and not contingently.⁴ Physicalists assume that there is a response to that challenge, too. We can explain how mind-brain identities can be true a posteriori even though the corresponding phenomenal and physical concepts refer directly.⁵ So this is another standard response to Kripke's orthodox argument. But this response does not, again, remove the difficulty that is raised by Kripke according to

⁴ The a posteriority thesis mentioned here is the thesis that Kripke's argument shares with the two-dimensional argument developed by Chalmers and Jackson. So this thesis, according to Papineau, is the root of the conflation of those two arguments.

⁵ See section 4 for the details of this response.

the new interpretation suggested by Papineau. To explain the a posteriority of mind-brain identities is to explain how one can be ignorant about their truth and to this extent find those identities conceivably false. But in response to Kripke's new challenge we need an explanation of how mind-brain identities can appear possibly false to someone who believes they are true. So ignorance about the truth of mind-brain identities is now out of the question.

Papineau argues that there is, in fact, no explanation of why mind-brain identity should appear contingent given that phenomenal concepts do not refer contingently. To illustrate this difficulty, Papineau appeals to proper names identities, such as the identity of Cicero and Tully. If you do not know that Cicero is Tully, you may think it is possible that they are two different people. But you cannot think this if you believe that Cicero *is* Tully. This is because 'Cicero' and 'Tully' refer directly and not descriptively. You cannot think that Cicero might have failed to be himself. Of course, there is room for the thought that Cicero might not have been Tully if one assumes that 'Cicero' and 'Tully' refer indirectly so that they have references fixed by such descriptions as 'the greatest Roman orator' and 'the greatest Roman statesman'. Obviously, the greatest Roman orator might not have been the greatest Roman statesman. But this does not help us see how mind-brain identities could appear contingent since there is no descriptive reading of mind-brain identities.

Now, the fact that we have no explanation of why mind-brain identity should appear contingent raises a difficulty for materialism because mind-brain identity does appear possibly false even to committed physicalists. Given that the identity appears contingent and given that we have no explanation of how this is possible physicalism turns out to be an unstable position.

On Papineau's interpretation, then, Kripke's argument can be summarized as follows:

- (1.) An identity can appear contingent only if at least one of the concepts flanking the identity sign refers contingently.
- (2.) Phenomenal concepts do not refer contingently.
- (3.) So mind-brain identity cannot appear contingent.
- (4.) But mind-brain identity appears possibly false even to committed physicalists.
- (5.) So physicalism is not a stable position.

Having proposed this new interpretation of Kripke, Papineau argues further that there is a solution of the difficulty that Kripke's argument raises. We simply have to admit that we *do not believe* mind-brain identity at some very *intuitive level*. This intuitive disbelief in the truth of mind-brain identity explains why

the identity appears possibly false while being perfectly consistent with our commitment to physicalism at the theoretical level.⁶

As I mentioned at the beginning, Papineau thinks that the intuition of dualism generates the feeling that there is some troublesome explanatory gap and that once we realize that this feeling is simply a manifestation of the intuition of dualism the problem of the explanatory gap disappears. Here I won't be interested in the issue of the explanatory gap. Instead, I will argue that the intuition of dualism undermines another challenge to physicalism, namely the conceivability argument which infers from the conceivable distinctness of mind and body that the relation between them is not necessary. I will show that the fact that we are in the grip of the intuition of dualism explains why our modal intuitions about the separability of mind and body are unreliable. I will also argue that the previous attempts to explain this unreliability that have been offered in the literature are not successful.

The entailment from conceivability to possibility may not be a premise (or the key premise) of Kripke's antimaterialist argument if Papineau's reading of that argument is right but it is the key premise of the two-dimensional argument against physicalism developed by Chalmers. So if we can see what is wrong with that entailment, this will certainly be an important result.

3. THE INTUITION OF DUALISM AND THE CONCEIVABILITY-POSSIBILITY PRINCIPLE

How does the intuition of dualism undermine the entailment from conceivability to possibility that is crucial to the conceivability argument? Papineau himself suggests this in his new response, inspired by the new reading of Kripke, to the orthodox reading of Kripke's argument. Before coming to his new interpretation of Kripke, Papineau assumed that the premise (4) in Kripke's argument is false because CP is unjustified and CP is unjustified because we can explain the appearance of the possible falsity of mind-brain identity (and thereby the appearance of its contingency) by appealing to conceptual differences between phenomenal and physical concepts (Papineau [2002]). Now Papineau argues that there is yet another approach available. The appearance that mind-brain identity is possibly false amounts simply to the intuition of the actual falsity of this identity (Papineau [2007b]) and we can 'explain it away' in so far as we can show that there is an explanation of the intuition of dualism consistent with assuming that the intuition is false. It remains to be seen what an adequate explanation of this intuition should be but it is reasonable to suppose that physicalists can succeed in finding

⁶ According to Papineau, this move is not ad hoc since there are plenty of other cases where our intuitive and theoretical beliefs conflict with each other.

such an explanation. Presumably there is an explanation of why the intuition arises even if it is false (Bloom [2004]; Melnyk [2003]; Papineau [1993, 2002, 2006]).

I take it that this way of explaining the apparently possible falsity of mind-brain identity away undermines the entailment from conceivability to possibility in the mind-brain case. If the appearance of the possible falsity of mind-brain identity results from the intuition of dualism, then this appearance does not entail possible falsity. By assumption, the intuition of dualism does not imply that mind-brain identity is false and hence the intuition of dualism does not imply that mind-brain identity is possibly false, either. But if the intuition of dualism does not imply the possible falsity of mind-brain identity, then the appearance of the possible falsity of this identity that stems from the intuition of dualism cannot have this consequence, either.

This objection to CP turns on the fact that phenomenal concepts do not refer contingently since this is the feature of phenomenal concepts that plays the key role in Papineau's proof that we are in the grip of the intuition of dualism. As we saw, this is also the feature of phenomenal concepts that plays a prominent role in Kripke's argument against materialism. Kripke's point was that given that phenomenal concepts do not refer contingently the appearance of mind-brain contingency cannot be explained away consistently with CP so that, assuming that CP is true, this appearance is the evidence that mind-brain identity is possibly false. It is somewhat ironic to realize now that the very same feature of phenomenal concepts undermines CP in the mind-body case.

CP is the key premise not only in Kripke's orthodox argument but also in the two-dimensional argument developed by Chalmers. So the objection to CP articulated here undermines the two-dimensional argument, too. When analyzed in two-dimensional terms, CP comes down to two claims: that *secondary conceivability* entails *secondary possibility* and that *primary conceivability* entails *primary possibility*, where primary conceivability and possibility are restricted to possible actual worlds and secondary conceivability and possibility to possible worlds understood counterfactually. In accordance with those two entailments, it is neither secondarily conceivable nor possible that water is not H₂O but that is both primarily conceivable and possible. The secondary inconceivability and impossibility of worlds in which water is not H₂O follows from the *rigidity* of the term 'water'. 'Water' is understood as the term that picks out the same referent across all possible counterfactual worlds. At the same time, 'water' is *semantically unstable* in the sense that it picks out different kinds across different actual worlds and this ex-

plains why it is conceivable and possible that water might not be H₂O in other actual worlds.⁷

When CP is employed in the two-dimensional argument it leads to the conclusion that there can be no a posteriori necessitation of the phenomenal by the physical due to the semantic stability of phenomenal concepts. The argument goes like this. A posteriori necessity is primarily conceivably false. So, assuming that primary conceivability entails primary possibility, a posteriori necessity must be primarily possibly false and hence primarily contingent. But this means that a posteriori necessity must be semantically unstable.⁸ And since phenomenal concepts are stable, the connection between the phenomenal and the physical cannot be necessary and a posteriori.

Now we can see that the key premise of this argument, the entailment from primary conceivability to primary possibility, is unjustified. That entailment is consistent with Kripkean a posteriori necessities because the natural kind terms, such as 'water' or 'heat', which are involved in those necessities, are unstable. However, the entailment turns out to be unjustified in the mind-brain case due to the fact that phenomenal concepts are stable. Given that phenomenal concepts are stable, mind-brain identity should not appear possibly false to someone who assumes that it is true. That is, it should not appear possibly false even across possible actual worlds. So given that even committed physicalists have this appearance nonetheless, this appearance must be an effect of the intuition of dualism and therefore, for the reasons explained above, cannot entail possibility.

It is again ironic to realize that the semantic stability of phenomenal concepts that the two-dimensional argument relies on is the very feature of those concepts that ultimately undermines that argument.

4. PREVIOUS CRITICISM OF THE CONCEIVABILITY-POSSIBILITY PRINCIPLE

The current literature assumes that we already have a good response to the two-dimensional argument because it assumes that we already have an explanation of why CP is unjustified. This is also the view of David Papineau who is one of the leading defenders of materialism. According to Papineau, CP is unjustified because it has the unjustified consequence that a posteriority is due to descriptive

⁷ The notion of semantic instability is due to Bealer [2002].

⁸ This is a two-dimensional equivalent of the a posteriority thesis adopted by Kripke. This two-dimensional version of the a posteriority thesis is also adopted by Jackson [1993, 1998]. However, Jackson, unlike Chalmers, does not motivate it by the intuition that conceivability entails possibility. Instead, he simply assumes that this thesis is supported by the two-dimensional analysis of a posteriori necessity.

content (Levine [2001]; Papineau [2002, 2006]).⁹ This is, indeed, the sort of criticism that I mentioned earlier when I discussed Kripke's argument against the identity theory. Apart from Papineau the key proponent of this criticism is Brain Loar (Loar [1997, 1999, 2003]). And the key point of this criticism is that we can explain how mind-brain necessity can be a posteriori consistently with assuming that phenomenal concepts have no descriptive content, namely, in terms of purely conceptual differences between phenomenal and physical concepts.¹⁰

This approach is, indeed, one of the most influential criticisms of CP in the current literature. It appeals crucially to some special features of phenomenal concepts that distinguish them from physical concepts and for this reason is a version of a strategy that is often referred to as *the phenomenal concept strategy*.¹¹ I do not think, however, that this strategy offers a good criticism of the entailment from conceivability to possibility and this is what I will explain below.

But first let me say a bit more about the strategy itself. The key point is that Loar and Papineau account for the a posteriori status of mind-brain necessities specifically by appealing to the fact the phenomenal properties picked out by phenomenal concepts constitute their own modes of presentations. This is the unique feature of phenomenal concepts that explains their semantic stability and this is also the feature that explains why we intuitively assume that mind-brain identities cannot be true. Since phenomenal properties constitute their own modes of presentations, we *expect* that they should present themselves as physical properties if they are physical. So given that phenomenal properties do not present themselves to us as physical properties, we assume that they cannot be physical. But Loar and Papineau argue in response that the expectation that we should be able to grasp the nature of phenomenal properties a priori is an illusion. There is no incoherence in the thought that even though phenomenal properties do not present themselves as physical properties, they are for all we know identical with physical properties.

It may seem somewhat unclear in what sense phenomenal properties constitute their own modes of presentations. But that may be clarified by Papineau's suggestion that phenomenal concepts use the properties they pick out, that those properties are somehow built into phenomenal concepts themselves. According to

⁹ See also Levine [2001].

¹⁰ Among those who make this point is Hill [1997]. However, Hill makes this point independently of the question of the validity of CP.

¹¹ The phenomenal concept strategy appeals to the nature of phenomenal concepts in response to all sorts of antimaterialist arguments, including the knowledge argument and the argument from the explanatory gap. Here I am focusing on the version of the strategy that was developed in response to the conceivability argument.

Papineau, then, we can without any incoherence suppose that the properties used by phenomenal concepts are a posteriori identical with physical properties.¹² So this explains how mind-brain identity claims can be true a posteriori despite the fact that phenomenal concepts are semantically stable.

Now, in response to the phenomenal concept strategy some philosophers argue that the problem with this strategy lies precisely in its appeal to the special features of phenomenal concepts. The problem is that it is far from obvious whether those features can themselves be explained in physical terms. Thus it is arguable that phenomenal properties cannot constitute their own modes of presentations unless we assume that phenomenal concepts *acquaint* us *directly* with the intrinsic nature of their referents and the relation of direct acquaintance is hardly explicable in physical terms (Chalmers [2006b]; Levine [2006]).

In my view, however, we can see that the strategy is problematic even regardless of its dependence on the special nature of phenomenal concepts. It seems to me that the dialectic of this strategy is badly flawed. The strategy assumes that CP is unjustified *because* we can make sense of how mind-brain identities can be true a posteriori. But the assumption that we can make sense of how mind-brain identities can be true a posteriori is question-begging in response to CP. We cannot assume that mind-brain identities can be true unless we already have an independent argument against CP. This is simply because CP, if true, rules out the truth of the assumption that mind-brain identities can be true. From this point of view, then, the details of Loar's and Papineau's account of mind-brain identities are irrelevant. Phenomenal concepts may use the properties they pick out, as Loar and Papineau point out, but if CP is true, the conceivable distinctness of those properties from physical properties will entail their possible distinctness from physical properties and hence the properties used by phenomenal concepts will not be identical with physical properties. So by assuming that the properties used by phenomenal concepts can be physical properties Loar and Papineau simply beg the question.

Loar and Papineau seem to be led to this question-begging by the mistaken assumption that the main difficulty in understanding how mind-brain identities can be true lies in the intuition that phenomenal concepts should reveal the essence of their referents given that they refer directly. Loar's and Papineau's account of mind-brain identities is meant primarily to explain that intuition away. However, from the point of view of the conceivability argument, what prevents us from understanding how mind-brain identities can be true is the intuition that

¹² A similar idea has been articulated by Ned Block. See Block [2006].

conceivability entails possibility which is independent of the intuition that phenomenal reference should be epistemically transparent. Loar and Papineau fail to take this into account and that is why they beg the question. Their account of mind-brain identities may well explain away the intuition of transparency considered independently of the conceivability argument but it begs the question as a response to that argument. The proper response to that argument cannot depend on assuming that there is nothing incoherent about the truth of mind-brain identities.

Loar argues that this assumption is ultimately justified by the intuitive distinction between concepts and properties. Concepts, unlike properties, are cognitively individuated and this alone entitles us to assume that phenomenal and physical concepts can be the concepts of the same properties. There may be the residual difficulty of accounting for this coreference consistently with the semantic stability of phenomenal concepts but this difficulty does not undermine the key point that there can, in principle, be two ways of conceptualizing physical properties, the physical and the phenomenal.¹³

I do not find this line of reasoning compelling. It may be, again, a good way of disarming the intuition of transparency considered independently of the conceivability argument but it has no force against that argument. The problem is that the distinction between concepts and properties is consistent with the intuition that conceivability entails possibility and that entailment together with the semantic stability of phenomenal concepts rule out the possibility that phenomenal concepts pick out physical properties. Thus we cannot really appeal to the intuitive distinction between concepts and properties in response to CP. To assume in response to CP that phenomenal and physical concepts can be the concepts of the same properties is, again, question-begging.

¹³ This line of reasoning is also explicit in Levine [2001]. Levine argues that we do not really understand how mind-brain identities can be true because we do not really understand how phenomenal concepts can pick out physical properties directly. This is so for reasons that have nothing to do with the expectation of the transparency of essence that was mentioned earlier. The problem is that phenomenal concepts conceive of their referents under substantive modes of presentation and we do not seem to have a materialistically respectable way of explaining how they can do that. All materialistically respectable accounts of direct reference seem to leave no room for such substantive modes of presentations. Papineau's idea that phenomenal concepts use the properties they pick out can be interpreted as one possible response to this difficulty but, as I already pointed out, Levine argues that the use-mention feature of phenomenal concepts can hardly be explained in physical terms. Still, despite this difficulty of accounting for phenomenal reference Levine assumes that the intuitive distinction between concepts and properties justifies the claim that phenomenal and physical concepts can be the concepts of the same properties. This will then block the entailment from conceivability to possibility since the identity of phenomenal and physical properties will be necessary despite being conceivably false.

I take it then that the phenomenal concept strategy falls short of showing that CP is unjustified in the mind-brain case. To avoid the charge of question-begging, the criticism of CP cannot depend on assuming that mind-brain identities can be true. No such assumption, of course, is involved in the new objection to CP that emerges from Papineau's interpretation of Kripke's argument against the identity theory and for this reason this new objection seems to me much stronger than the criticism that comes from the phenomenal concept strategy. Of course, given that CP is unjustified, it follows that the a posteriority thesis (that a posteriority is due to semantic instability), which is a consequence of CP, is unjustified, too.

5. IS THE CONCEIVABILITY OF ZOMBIES A PRIORI?

We have seen why CP is unjustified and without CP the two-dimensional conceivability argument against materialism collapses. But there is still one aspect of that argument that needs to be addressed. Whereas the semantic stability of 'consciousness' implies that the primary conceivability of zombies *should be* dependent on assuming whether or not psychophysical identity is true, Chalmers explicitly argues that the primary conceivability of zombies is a priori, that it does not depend on any a posteriori truths about the actual world and a posteriori identities, in particular. This is, indeed, the key assumption of Chalmers' argument and this is the assumption that leads Chalmers astray. Let me then clarify this aspect of Chalmers' reasoning. This will help us see better where Chalmers goes wrong.

Chalmers' justification of the assumption in question is based on the analysis of Kripkean a posteriori necessities. This analysis reveals that a posteriori identities affect the secondary conceivability but not the primary one. Thus consider the statement 'Water is not H₂O'. Whether or not the statement 'Water is not H₂O' is primarily conceivable does not depend on the fact that water is H₂O. But whether or not 'Water is not H₂O' is secondarily conceivable does depend on whether or not water is H₂O. Given that water is H₂O, 'Water is not H₂O' is not secondarily conceivable. This difference between the primary and secondary conceivabilities in the case of 'Water is not H₂O' results from the fact that 'water' has different primary and secondary *intensions*, that is, functions that pick out the referent of 'water' across possible actual worlds and possible counterfactual worlds, respectively. The secondary intension of 'water' depends on the fact that water is H₂O in the actual world. Given that this identity is true, 'water' picks out H₂O in all possible counterfactual worlds. The primary intension of 'water', on the other hand, does not depend on the fact that water is H₂O but is determined by the properties

that identify water or fix the reference of 'water' (watery properties). Thus the primary intension of 'water' picks out in all actual worlds any kind of stuff that is watery.

Chalmers argues then by analogy that if we assume that the primary and secondary intensions of 'consciousness' are different, zombies will not be secondarily conceivable but will be conceivable primarily. If the secondary intension of 'consciousness' is different from the primary intension, the property that 'consciousness' picks out will be different from the property that identifies consciousness across possible actual worlds, that is, the property of *phenomenal feel*. If consciousness then turns out to be actually identical with some physical property, consciousness will be identical with that property in every counterfactually possible world and therefore zombies will not be secondarily conceivable. But, as Chalmers points out, this will not affect the primary conceivability of zombies. For the primary intension of 'consciousness' picks out in all possible actual worlds the property of phenomenal feel that is, by assumption, different from the actual referent of 'consciousness'. Thus even if we assume on a posteriori grounds that consciousness is physical, we are not committed to holding that the property of phenomenal feel is physical and hence we can assume that there are other actual worlds which are like our world in all physical respects but in which nothing feels one way or another. Such worlds will be zombie worlds in the primary sense of zombies (or in the primary sense of consciousness).

So, according to Chalmers, if the primary and secondary intensions of 'consciousness' are different, the primary conceivability of zombies is a priori. And if the intensions in question are identical, Chalmers assumes that the primary conceivability of zombies is a priori, too. For the primary conceivability of zombies is guaranteed by the fact that it is a priori conceivable that there are other actual worlds physically identical to our world but containing no phenomenal feel and such worlds are a priori conceivable regardless of whether or not the primary and secondary intensions of 'consciousness' are identical (regardless of whether the phenomenal feel fixes the reference of 'consciousness' or whether it is identical with the referent of 'consciousness'). If those intensions are, in fact, identical, argues Chalmers, the secondary conceivability of zombies along with their primary conceivability will be a priori.

It is clear though that Chalmers assumes too quickly that it is a priori conceivable that there are other actual worlds physically identical to our world but containing no phenomenal feel. That is, indeed, a priori conceivable but in the light of Papineau's analysis of our zombie worlds intuition this is something that *should not be* conceivable if we assume that the phenomenal feel itself is a posteri-

ori identical with some physical property. This follows from the semantic stability of 'the phenomenal feel'. The semantic stability of 'the feeling of pain' guarantees that 'the feeling of pain' will pick out the feeling of pain in all possible actual worlds. Thus if we assume that the feeling of pain is identical with some physical property P in our world, we are committed to holding that 'the feeling of pain' will pick out P in all actual worlds and then it should not be conceivable that the feeling of pain might not be P . So therefore to someone who assumes that the phenomenal feel is physical it should not be conceivable that some other actual world physically identical to our world might not contain the property of phenomenal feel.

This then creates a difficulty for Chalmers. The difficulty is that zombie worlds are conceivable even to committed physicalists who believe that the phenomenal feel is physical, which appears to be inconsistent. Given their theoretical commitment, physicalists should not find zombie worlds conceivable. And if we now follow Papineau's way of resolving this inconsistency, we will end up denying the key premise of Chalmers' two-dimensional argument, that the conceivability of zombies entails possibility. According to Papineau, in order to avoid the inconsistency in question we have to assume that the identity of the phenomenal feel with P is something that physicalists (and anybody else) cannot believe at an intuitive level. This solution consists in assuming that the intuition of dualism explains why zombies are conceivable while at the same time leaving room for our commitment to the truth of the identity of the phenomenal feel with P at the theoretical level. However, for the reasons explicated earlier, the effect of this solution is breaking the entailment from the conceivability to the possibility of zombies. If the conceivability of worlds that resemble our world physically but contain no phenomenal feel follows from the intuition that the phenomenal feel cannot be a physical property, then we have no reason to believe that the conceivable worlds in question are possible. And without the entailment from the conceivability to the possibility of those worlds the two-dimensional argument does not get off the ground.

Notice that this response to Chalmers does not, strictly speaking, depend on assuming that 'consciousness' is semantically stable but only that 'the phenomenal feel' is. If we assume that the primary and secondary intensions of 'consciousness' are different, the referent of 'consciousness' will be different from the property of phenomenal feel that identifies consciousness and hence 'consciousness' will be semantically unstable: it will pick out different kinds in different actual worlds. However, even in this scenario the 'phenomenal feel' will be semantically stable

and this alone will make the a priori status of the primary conceivability of zombies problematic.

Bibliography

- Bealer [2002] – G. Bealer, *Modal Epistemology and the Rationalist Renaissance*, [in:] *Conceivability and Possibility*, T. Gendler & J. Hawthorne (ed.), Oxford University Press, Oxford 2002: 71-126.
- Block [2006] – N. Block, *Max Black's Objection to Mind-Body Identity*, [in:] *Phenomenal Concepts and Phenomenal Knowledge: New Essays on Consciousness and Physicalism*, T. Alter & S. Walter (ed.), Oxford University Press, New York 2006: 249-306.
- Bloom [2004] – P. Bloom, *Descartes' Baby*, Basic Books, New York 2004.
- Chalmers [1996] – D. Chalmers, *The Conscious Mind*, Oxford University Press, Oxford 1996.
- Chalmers [1999] – D. Chalmers, *Materialism and the Metaphysics of Modality*, "Philosophy and Phenomenological Research", 59 (2) 1999: 473-496.
- Chalmers [2002] – D. Chalmers, *Does Conceivability Entail Possibility?*, [in:] *Conceivability and Possibility*, T. Gendler & J. Hawthorne (ed.), Oxford University Press, Oxford 2002: 145-200
- Chalmers [2006a] – D. Chalmers, *The Two-Dimensional Argument against Materialism*, [in:] *The Character of Consciousness*, D. Chalmers (ed.), Oxford University Press, Oxford 2006.
- Chalmers [2006b] – D. Chalmers, *Phenomenal Concepts and the Explanatory Gap*, [in:] *Phenomenal Concepts and Phenomenal Knowledge*, T. Alter & S. Walter (ed.), Oxford University Press, New York 2006: 167-194.
- Hill [1997] – C. Hill, *Imaginability, Conceivability, Possibility and the Mind-Body Problem*, "Philosophical Studies" 87 (1) 1997: 61-85.
- Hill & McLaughlin [1999] – C. Hill & B. McLaughlin, *There are Fewer Things in Reality Than are Dreamt of in Chalmers' Philosophy*, "Philosophy and Phenomenological Research" 59 (2) 1999: 445-454.
- Jackson [1998] – F. Jackson, *From Metaphysics to Ethics: A Defence of Conceptual Analysis*, Oxford University Press, Oxford 1998.
- Kripke [1980] – S. Kripke, *Naming and Necessity*, Blackwell, Oxford 1980.
- Levine [1983] – J. Levine, *Materialism and Qualia: The Explanatory Gap*, "Pacific Philosophical Quarterly" (64) 1983: 354-361.
- Levine [2001] – J. Levine, *Purple Haze*, Oxford University Press, New York 2001.
- Levine [2006] – J. Levine, *Phenomenal Concepts and the Materialist Constraint*, [in:] *Phenomenal Concepts and Phenomenal Knowledge*, T. Alter & S. Walter (ed.), Oxford University Press, New York 2006: 145-166.
- Loar [1997] – B. Loar, *Phenomenal States*, [in:] *The Nature of Consciousness: Philosophical Debates*, N. Block, O. Flanagan, G. Güzeldere (ed.), MIT Press, 1997: 597-616.

- Loar [1999] – B. Loar, *David Chalmers' The Conscious Mind*, "Philosophy and Phenomenological Research" 59 (2) 1999: 465-472.
- Loar [2003] – B. Loar, *Qualia, Properties, Modality*, "Philosophical Issues" 13 (1) 2003: 113-129.
- Melnyk [2003] – A. Melnyk, *Papineau on the Intuition of Distinctness*, SWIF Philosophy of Mind Review (4) 2003; available at:
<http://lgxserver.uniba.it/lei/mind/forums/004-0003.htm> (accessed 25 May 2008).
- Nagel [1974] – T. Nagel, *What Is It Like to be a Bat?*, "Philosophical Review" 83 (4) 1974: 435-450.
- Papineau [1993] – D. Papineau, *Physicalism, Consciousness and the Antipathetic Fallacy*, "Australasian Journal of Philosophy" 71 (2) 1993: 169-183.
- Papineau [2002] – D. Papineau, *Thinking about Consciousness*, Oxford University Press, Oxford 2002.
- Papineau [2006] – D. Papineau, *Phenomenal and Perceptual Concepts*, [in:] *Phenomenal Concepts and Phenomenal Knowledge*, T. Alter & S. Walter (ed.), Oxford University Press, New York 2006: 111-144.
- Papineau [2007a] – D. Papineau, *Kripke's Proof That We Are All Intuitive Dualists*, Forthcoming.
- Papineau [2007b] – D. Papineau, *Kripke's Proof is Ad Hominem not Two-Dimensional*, "Philosophical Perspectives" 21 (1) 2007: 475-494.
- Yablo [2000] – S. Yablo, *Textbook Kripkeanism and the Open Texture of Concepts*, "Pacific Philosophical Quarterly" 81 (1) 2000: 98-122.