

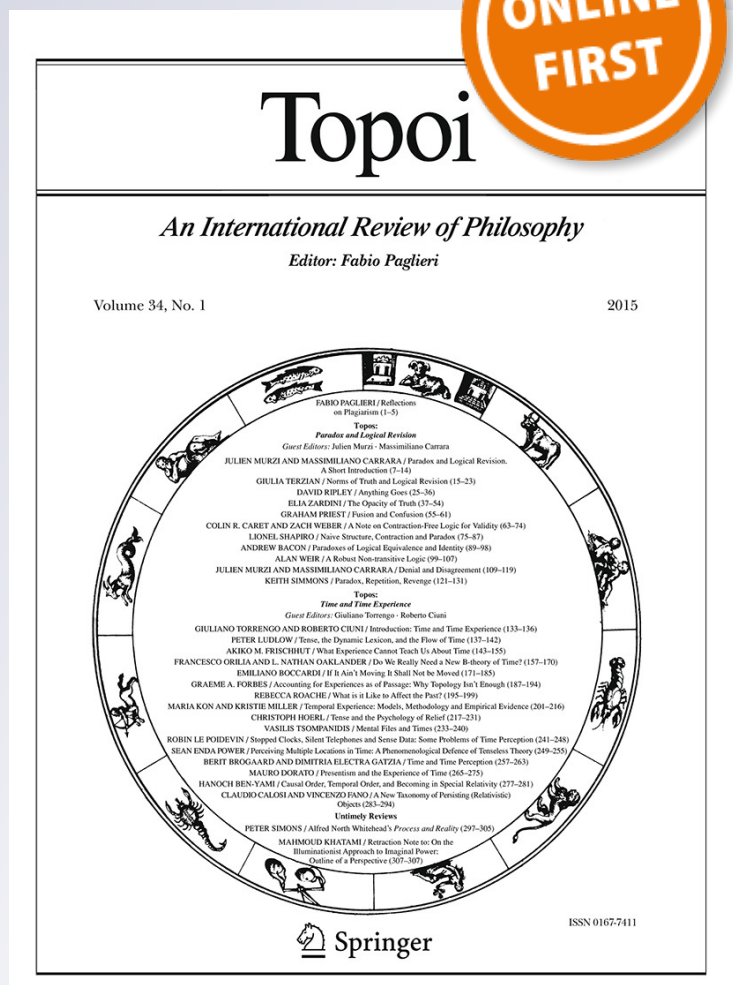
On a Confusion About Which Intuitions to Trust: From the Hard Problem to a Not Easy One

Miguel Ángel Sebastián

Topoi
An International Review of Philosophy

ISSN 0167-7411

Topoi
DOI 10.1007/s11245-015-9326-3



Your article is protected by copyright and all rights are held exclusively by Springer Science +Business Media Dordrecht. This e-offprint is for personal use only and shall not be self-archived in electronic repositories. If you wish to self-archive your article, please use the accepted manuscript version for posting on your own website. You may further deposit the accepted manuscript version in any repository, provided it is only made publicly available 12 months after official publication or later and provided acknowledgement is given to the original source of publication and a link is inserted to the published article on Springer's website. The link must be accompanied by the following text: "The final publication is available at link.springer.com".

On a Confusion About Which Intuitions to Trust: From the Hard Problem to a Not Easy One

Miguel Ángel Sebastián¹

© Springer Science+Business Media Dordrecht 2015

Abstract Alleged self-evidence aside, conceivability arguments are one of the main reasons in favor of the claim that there is a Hard Problem. These arguments depend on the appealing Kripkean intuition that there is no difference between appearances and reality in the case of consciousness. I will argue that this intuition rests on overlooking a distinction between cognitive access and consciousness, which has received recently important empirical support. I will show that there are good reasons to believe that the intuition is misguided—at least on the reading that the conceivability arguments require—and hence that the arguments are unsupported. This, in turn, alleviates the Hard Problem but leaves us with what I think is a *not easy problem*.

Keywords Hard problem · Phenomenal consciousness · Conceivability argument · Cognitive access · Not easy problem

1 The Hard Problem and the Conceivability Argument

The Hard Problem of consciousness is presented by Chalmers (1996) as the problem of explaining why some processes that take place in the brain are “accompanied by an experienced inner life”, by conscious experiences. The

idea behind the Hard Problem is that structure and function do not suffice for explaining consciousness and so, if this is all our science can tell us about, we are left with a lack of scientific explanation. We might need some other kinds of resources if we want to provide an explanation, but we have no idea where we should begin looking. And thus, the Hard Problem.

Of course, the claim that structure and function cannot explain consciousness requires further support, for many would disagree with it. Some philosophers have faced this challenge and attempted to provide arguments in its favor. A paradigmatic example is the conceivability argument. According to those who defend the argument, we can conceive of a being that satisfies whatever function and structure we postulate that is sufficient for having certain conscious experience, but who, nonetheless, has a different one or lacks consciousness at all. And this poses a unique problem. The reason why such conceivability is problematic in the case of consciousness but not in the case of other *a posteriori* necessities is the assumption, perfectly expressed by Searle’s famous quote, that “where appearance [consciousness] is concerned we cannot make the appearance/reality distinction because the appearance is the reality” (Searle 1997, p. 456). I will refer to this assumption as ‘the Kripkean intuition’, for it is the core idea that supports Kripke’s articulation of the conceivability argument—and also, as we will see, posterior refinements of it by Chalmers himself.

I will argue that this intuition is only acceptable if we assume that there is no distinction between the cognitive access we have to our conscious experiences and the experiences themselves, a distinction that seems to be conceptually sound (Block 1995–2002) and empirically supported (Block 2007, 2011b; Sebastián 2014). This will undermine the conceivability argument and the Hard

✉ Miguel Ángel Sebastián
msebastian@gmail.com

¹ Instituto de Investigaciones Filosóficas, Universidad Autónoma de México, UNAM, Circuito Mtro Mario de la Cueva, Ciudad Universitaria, 04510 Del. Coyoacán, D.F., Mexico

Problem with it. I will conclude that, if consciousness can be disentangled from cognitive access, we would be left with a “not easy problem”.

2 The Kripkean Intuition and the Conceivability Argument

Many agree with the idea that there is no difference between appearance and reality in the case of consciousness. As Kripke (1980) presents the idea:

To be in the same epistemic situation that would obtain if one had a pain is to have a pain; to be in the same epistemic situation that would obtain in the absence of pain is not to have a pain.... Pain ... is not picked out by one of its accidental properties; rather it is picked out by its immediate phenomenological quality.... If any phenomenon is picked out in exactly the same way that we pick out pain, then that phenomenon is pain. (ibid., pp. 152–153)

Based upon this idea, Kripke argues against identity theories of consciousness on the assumption that the terms involved in the identity are rigid designators. He maintains that the corresponding identity claim, like ‘pain is C-fiber stimulation’, must be necessary if true, but it appears contingent. The problem is that, contrary to other scientific *a posteriori* identities, the appearance of contingency cannot be explained away. Consider a case like ‘water is H₂O’. The appearance of contingency in cases like this is due to our recognition of the genuine possibility that water might not have its appearance properties or that such properties might be had by some other substance. But, conscious states cannot be pulled apart from their appearance, according to the intuition above. Thus, given that there is no alternative explanation of the appearance of contingency, identity theorists should accept that the claim is not necessary and hence not true. I will argue that the Kripkean intuition is ungrounded. But it will be useful to focus on later and more sophisticated articulations of the conceivability argument like the one presented by Chalmers.

Although many materialists are not identity theorists, they all commit themselves to the claim that if certain physical facts obtain then so will certain mental facts, where this entailment holds with metaphysical necessity. Based on the Kripkean insight, Chalmers (1996) presents a conceivability argument against materialism and not only against identity theories. In a nutshell, the conceivability argument holds that (1) we can conceive that there is a possible world (w_z), which is a microphysical duplicate of the actual world ($w_{@}$) but such that some phenomenal truth in $w_{@}$ is not true in w_z and that (2) if we can conceive that

there is w_z , then w_z is possible. But if w_z is possible then (3) materialism is false, insofar as we take materialism to be committed to the claim that any *minimal duplicate* of $w_{@}$ —a world that satisfies all the physical truths in $w_{@}$ and “that’s all”—is a duplicate *simpliciter*.

Many philosophers grant the first premise (cf. Dennett 1991; Dretske 1995; Lewis 1990). However, the entailment from conceivability to possibility involved in (2) is often rejected. The interesting counterexamples for my current purposes are *a posteriori* necessities, because some philosophers hold that, for example, ‘water is not H₂O’ is conceivable while not metaphysically possible. In order to deal with this cases, Chalmers (2002; 2010) offers a two-dimensional analysis of conceivability:¹

There is a sense in which ‘water is not H₂O’ is not conceivable, we can call it ‘secondary conceivability’. In this sense, a situation in which one considers that water is not H₂O should better be understood as a situation in which one considers a substance that *seems* to be water—for it has the appearance properties of water—but it is not, because water is necessarily H₂O. Secondary conceivability seems to be a good guide to metaphysical possibility but hardly one usable in *a priori* arguments like the conceivability one, for what is secondarily conceivable depends on empirical investigation. But, there is another sense of conceivability, primary conceivability, in which we can say that ‘water is not H₂O’ is conceivable: in the sense that it cannot be ruled out *a priori* by an ideal thinker. Chalmers also constructs two notions of possibility—parallel to the former notions of conceivability—corresponding to an epistemic and a metaphysical possibility. A sentence *S* is epistemically possible if and only if it is true in some world *w* considered as actual (*S*’s primary intension is true at *w*). On the other hand, a statement is metaphysically possible if and only if it is true in some world considered as counterfactual (*S*’s secondary intension is true at *w*).

With these tools in hand, we can present Chalmers’ argument (p. 152). Let *P* be the conjunction of all the microphysical truths of the universe and *Q* a phenomenal truth like ‘I am in pain’.

¹ In his attempt to single out the circumstances in which conceivability is a good guide to metaphysical possibility, Chalmers also distinguishes between a *positive* and a *negative* notion of conceivability and between *prima facie* from *ideal* conceivability. The notion of positive conceivability is a bit obscure and is characterized “in terms of what subjects can form a positive conception of” (Chalmers 2010, p. 144). In this paper I will have in mind negative conceivability, which is clearer and simpler and many philosophers concede that it suffices for the argument to go through. According to it, a sentence *S* is negatively conceivable for a subject if and only if she can entertain *S* and is unable to rule it out through *a priori* reasoning. To avoid counterexamples derived from the cognitive limitations of the agent on such reasoning, Chalmers considers ideal thinkers who lack such limitations (*ideal* rather than *prima facie* conceivability).

1. P and not Q is primarily conceivable.
2. If P and not Q is primarily conceivable, then P and not Q is epistemically possible.
3. If P and not Q is epistemically possible, then P and not Q is metaphysically possible.²
4. If it is metaphysically possible that P and not Q then materialism is false.

∴ Materialism is false.

The argument is valid but, is it sound? In the rest of the paper I will argue, by rejecting premise 3, that it is not. Many have attempted to attack this premise; the novelty of my proposal lies in undermining the Kripkean intuition that supports the entailment in 3 by calling attention to the possibility that the primary intension of phenomenal sentences depends on the cognitive access we have to our conscious experiences and the possibility that such a cognitive access can be completely disentangled from the experiences themselves.

3 It *Seems* Pain But, *Is* it Pain?

Let's grant, for the sake of the discussion, that a microphysical duplicate of our world lacking pain is conceivable and hence that the primary intension of 'P and not Q' is true in this world W: W is epistemically possible. Does it follow from that the secondary intension is also true; i.e. that W is metaphysically possible? This entailment is supported by the Kripkean intuition that appearances and reality coincide in the case of consciousness: contrary to what happens with concepts like WATER, the primary and secondary intension coincide in the case of phenomenal concepts, like those involved in Q. Making use of Kripke's example, if something *seems* to be pain then it *is* pain; or in other words, nothing can fail to be pain if it has the appearance of pain: the primary and secondary intension of the phenomenal concept PAIN coincide.

Once the connection between appearances and primary intension on the one hand and secondary intension and the nature of the phenomenon on the other in the bidimensionalist framework is made explicit, we can analyze premise 3 of the conceivability argument in more detail. Let

² Premise (3) should be "If P and not Q is epistemically possible, then either P and not Q is metaphysically possible or panpsychism is true" and hence the conclusion "Materialism is false or panpsychism is true". I have presented this abbreviated version for the sake of simplicity in the exposition. Interesting as panpsychism might be, it entails a revisionary metaphysics not required to block the argument if my reasoning in this paper is correct. I will nonetheless say a bit more about panpsychism in the next section.

'*Seems_w*(X)' stand for the claim that the primary intension of X is true at a world W (there is an epistemically possible world, W, in which X is the case) and '*Is_w*(X)' the secondary intension of X is true at W (there is a metaphysically possible world, W, in which X is the case). Let's also give the name 'WT' (Weak Transparency)³ to the inference from primary to secondary intension: *Seems_w*(X) → *Is_w*(X).

WT does not universally hold as cases of *a posteriori* necessities show. Should we accept it in the case of 'P and not Q'? Making use of the introduced notation, we can analyze the inference as follows:

- 3.1 *Seems_w*(P and not Q) Assumption
- 3.2 *Seems_w*(P) and *Seems_w*(not Q) By distribution over disjunction
- 3.3 *Is_w*(P) From 2 and WT
- 3.4 *Is_w*(not Q) From 2 and WT
- 3.5 *Is_w*(P and not Q) By aggregation
- (3) *Seems_w*(P and not Q) → *Is_w*(P and not Q)

As we see, the inference requires two different instantiations of WT. Some panpsychists (see Chalmers 2010) deny the one in 3.3, while accepting the one in 3.4 following the Kripkean intuition. Their route to reject the metaphysical possibility of P and not Q is to hold that microphysical entities have an intrinsic nature physics cannot tell us about: all physics can tell us about is function and structure. Hence, microphysical terms would have different primary and secondary intensions and so the metaphysical possibility of a world where P and not Q is the case does not follow from its epistemic possibility. Panpsychists maintain that the intrinsic nature of our fundamental entities is tied to consciousness. So, there are worlds in which the primary intension of 'P and not Q' is true—worlds in which *it seems that P and not Q is the case*—, namely those worlds sufficiently close to ours in which the fundamental microphysical entities have a different intrinsic nature from ours, one that is not tied to consciousness. But, there is no world in which the secondary intension of P is true and not so the secondary intension of Q; i.e. no metaphysically possible world in which *P and not Q is the case*.

I will argue that it is not necessary to endorse the commitments of panpsychism by focusing on 3.4. Many, following Kripke, have found the entailment from the primary intension to the secondary one acceptable in the case of phenomenal truths. I am going to show that this intuition should be rejected on the basis of a distinction

³ I am borrowing the term from Williford (2007), who coined it inspired by David Armstrong (1999) where he speaks about the doctrine of the "perfectly transparent mind", meaning that WT and its inverse both hold.

between consciousness and the cognitive access we have thereof.

4 On a Confusion About Which Intuitions to Trust

When Kripke gave his lectures, the immediate connection between what is conveyed by our experience and the informational resources on which we, at the personal level, navigate the world and form beliefs was taken for granted. I am going to argue that once this connection is dismissed, Kripke's intuition is left unsupported.

The conceivability argument attempts to *a priori* establish—panpsychism aside—the falsehood of materialism. There is no empirical step whatsoever involved in the argument; and the step that makes use of WT is not an exception. According to the argument, WT holds with conceptual necessity in the case of phenomenal truths. Such a necessity rests precisely upon the idea that there is no distinction between appearance and reality in the case of consciousness and no matter what we learn about our cognitive system this will remain so. Contrary to this idea, I think that there are good empirical reasons to be suspicious about such a claim *in the actual world*. But even if this cannot be established (I myself think that the evidence is far from being conclusive), I am going to argue that the alleged connection between primary and secondary intension cannot be established *a priori*. This in turn makes the conceivability argument unsound because it is a straightforward rejection of WT on which the argument relies. Briefly, the argument is the following:

1. WT is necessarily true in the case of phenomenal truths.
2. Cognitive access can be (at least conceptually) dissociated from consciousness.
3. It is possible that consciousness and cognitive access can be completely dissociated. From 2
4. It is possible that the primary intension of a phenomenal truth depends on the cognitive access we have to our own experiences: the primary intension of 'I am not in pain' might depend on my having cognitive access to my pain experience.
5. It is possible that WT is false: The primary intension of 'I am not in pain' might be true, in spite of the fact that I am in pain. From 3 and 4

∴ WT is false From 1 and 5

I have already discussed premise 1. What about the other premises? In the next subsection, I will defend premises 2 and 3 and in Sect. 4.2, I will deal with premise 4. Section 5 discusses some possible objections. This discussion will also help to clarify the details of the argument.

4.1 Cognitive Access and Consciousness

Premise 2 was defended by Block (1995–2002), who famously introduced a conceptual distinction between Phenomenal Consciousness and Access Consciousness. On the one hand, a mental state is Access-Conscious if and only if, roughly, its content is available for belief formation and rational control of action. On the other hand, a mental state is Phenomenally-Conscious if and only if there is something it is like to be in that state: Phenomenal Consciousness (or simply consciousness or conscious experience as I have been referring to it) is what lies at the heart of the Hard Problem. The conceptual distinction is clear and many philosophers grant it. What remains controversial is the conclusion to be derived from this conceptual distinction: do these two concepts pick up different properties? Block's conceptual distinction is all my argument requires, because, as we have seen, in order to show the conceivability argument unsound we just need to establish the possibility of WT being false. However, let me go a step further for those skeptical about the distinction, because I think there are good empirical reasons to believe that they do pick up different properties.

In the search for an answer to the previous question, the debate has recently moved away from the conceptual domain into the empirical one focusing on the possibility of phenomenal consciousness without access. The question turned into whether the neural basis of phenomenal consciousness can be disentangled "from the neural machinery of the cognitive access that underlies reports of phenomenal consciousness". There is good evidence supporting this claim.

Based on the results of partial report experiments, like those in Sperling (1960) and some more recent results—Landman et al. (2003), Sligte et al. (2008) etc—, Block (2007, 2011b), has argued that the capacity of the memory buffer in which the content of phenomenally conscious states is encoded is greater than that of cognitive access. For example, in Sperling's experiment participants are asked to look at a 3×4 array of letters quickly flashed on a computer screen and to recall them immediately afterwards. This technique, called 'free recall' showed that on average participants were able to recall 4–5 letters out of the 12 they were shown, corresponding to the capacity of the working memory on which cognitive access depends. In a second condition, participants were presented with the same matrix for the same amount of time, and then heard a tone at a particular pitch. They were to recall the letters in one of the rows, depending on the frequency of the pitch. On average, subjects were able to recall more during these cued recall trials than during free recall. Furthermore, by modifying the delay between the presentation of the matrix and the cue, Sperling was able to show that visual stimuli that are not added to working memory are discarded soon after their initial introduction. Block concludes that the best explanation for

this result is that “...the machinery of phenomenology is at least somewhat different from the machinery of cognitive accessibility” (Block 2007, p. 489).⁴

Very recently, relating the neural correlates of cognitive access to empirical research into the neurophysiology of dreams, I have presented further support in favor of the claim that cognitive access is not required for having experiences (Sebastián 2014). There is strong empirical evidence suggesting that cognitive access essentially depends on the activity of the dorsolateral prefrontal cortex—more precisely Brodmann’s area 46. This has been shown, for example, by Fuster (2008), Goldman-Rakic (1988), Lau and Passingham (2006) and Oliveri et al. (2001), and contemporary elaborations at the neural level of the most popular and empirically supported theory of cognitive access; the (GWS) Global Workspace (Baars 1988) explicitly states such a commitment (Dehaene and Naccache 2001). However, the dorsolateral prefrontal cortex is strongly deactivated during sleep (Braun et al. 1997; Maquet et al. 1996), a period in which, as common sense⁵ and independent evidence shows (Horikawa et al. 2013; LaBerge 1988; Leclair-Visonneau et al. 2010), we entertain conscious experiences: dreams. This evidence provides further support for the claim that cognitive access is not constitutive of our experiences.

⁴ Interestingly, some philosophers and scientists who want to defend the constitutive role of cognitive access in phenomenology attempt to accommodate the result of partial report experiments by appealing, *pace* Kripke, to a phenomenological illusion (Rosenthal 2007; Brown and Lau forthcoming; Kouider et al. 2010; Phillips 2011). In particular, they maintain, roughly, that our experience *seems* (*appears*) to be rich—it seems to us that we see all the letters in the array—but it is sparse—we only experience some of the letters. Furthermore, there is some evidence showing that there is in fact some level of illusion (Kouider et al. 2010), and the debate now focuses on the extent of the illusion and on whether unconscious processing might explain the result of partial report experiments (Block 2012; Kouider et al. 2012). A similar thought derives from Chalmers’ rejection of his dancing qualia argument (2010, p. 24 fn. 7). One might reply that this reading conflates appearances and the beliefs that express such appearances—if our experience is sparse, then it appears so but we judge it mistakenly to be rich. But then, what is completely obscure is what is meant by *appearance* and how to make sense of what the primary intension is supposed to pick up in a way that does not beg the question against those who dispute the intuition.

⁵ We do often recall the content of our dreams, and one might think that this suggests that we have cognitive access to our dreams. However, the kind of reportability relevant for the discussion is not retrospective reportability (as when we wake up) but rather the one associated with the working memory; i.e. cognitive access to the experience we are currently having (Sebastián (2014) further discusses to some extent the relation between cognitive access and working memory). If this reasoning is right, reports upon awakening do not depend on cognitive access but rather on other kinds of memory systems like long term memory. This would entail that information can be encoded in the latter without going through the former. Further empirical evidence is required to determine whether this corresponds to our cognitive architecture.

It might well turn out that new evidence tips the scale in a different direction and that cognitive access is required for consciousness *in the actual world*, but the lesson learned from this sketch of the current debate in cognitive sciences and philosophy is that, at the very least, they can be conceptually dissociated—and I don’t know of any response claiming that the empirical research makes no sense due to an *a priori* connection. So, at the very least, the possibility of lacking cognitive access to an experience has been established—a possibility not foreseen by Kripke when he formulated his argument. And hence:

nPNA Cognitive access is not necessary for consciousness

Moreover, there is another thesis, which will become relevant later, that cannot be ruled out *a priori* either:

nANP Consciousness is not necessary for cognitive access

An example from the actual debate in cognitive science about the role of cognitive access in consciousness might illustrate the idea in nANP—which might sound odd at first glance—without considering exotic worlds. According to the GWS theory, the information encoded in the GWS is made available to other processes for belief formation, report and rational control of action. So, one has cognitive access to, say, her pain experience just in case certain information is encoded in the GWS. This in turn requires certain brain activity in the center of the GWS, call this activity ‘GWS-pain’. Those who deny that consciousness requires cognitive access deny that GWSpain is necessary for a pain experience, and maintain that some form of reverberatory activity in areas not including those of the GWS suffice for consciousness. Call such an activity ‘NCpain’. If NCpain is not necessary for GWSpain, and consciousness depends on NCpain, then it is possible that we have cognitive access to pain without being in pain: nANP is true. Some empirical examples that I will consider in 5.2 suggest that this might be a possibility even in the actual world.

In the next section, I am going to argue that it is an open empirical possibility about the nature of our cognitive systems that the primary intension of phenomenal concepts depends constitutively on cognitive access. As we are about to see, this, conjoined with the possibility that one lacks access to consciousness entails that WT cannot be established *a priori* in the case of phenomenal truths and, hence, that the conceivability argument is unsound.

4.2 Seemings and Cognitive Access

Once we acknowledge the conceptual distinction between cognitive access and consciousness the question is: does

the primary intension of phenomenal concepts require cognitive access to the experience? If an affirmative answer to this question is possible and nPNA is true, then WT is not necessary and hence not true.

I fail to see any *a priori* reason for denying the possibility that the primary intension of phenomenal concepts depends on cognitive access. And there might be reasons for thinking that, in fact, the possession of phenomenal concepts depends on cognitive access. For example, one might think that phenomenal concepts are only required in rational control of action. So, if these concepts are not required at a lower level of cognition, then it seems reasonable to assume that the features we have cognitive access to are the only ones that are relevant for the possession of a phenomenal concept. So 1-COG seems to be a plausible principle:

1-COG The primary intension of phenomenal concepts constitutively depends on cognitive access.

The suggestion above is too weak to support the claim that 1-COG is true. But, in order to resist the conceivability argument we only need 1-COG to be possible, and whether 1-COG is true seems even to be an open empirical question regarding human cognitive architecture.

We have seen that the soundness of the conceivability argument depends on the entailment from the truth of the primary intension of *not Q* to the truth of its secondary intension. But if 1-COG and nPNA are possible, such entailment should be rejected. The reason is that the primary intension of a sentence like 'I am not in pain' might be true, while the secondary intension is false because I am in fact in pain but lack cognitive access—on which the primary intension depends—to my experience. In more intuitive terms, I might *be* in pain while it *seems* to me that I am not—while I am in a state with the *appearance* of one in which I am not in pain, *pace* Kripke. The entailment in step 3 of the conceivability argument is invalid and the conceivability argument unsound.

In the next section I will further elaborate on the argument against the entailment from primary to secondary intension in the case of phenomenal truths by presenting some possible objections.

5 Possible Objections and Further Elaboration

5.1 The Super-Experiential World

Unfortunately for my current purposes, it is unclear that the presented argument is sufficient to block the conceivability argument. The reason is that the conceivability argument can be reproduced in terms of a positive phenomenal truth.

Instead of considering a *zombie* world consider a *super-experiential* world. A super-experiential world is a microphysical duplicate of the actual one, in which more experiences obtain. So, for example, in a super-experiential world, I (or my counterpart) would have a headache that I do not have in the actual one. I think that the (primary negative) conceivability of such a super-experiential world is granted by the very same lack of *a priori* entailment between physical and phenomenal truths that grants the conceivability of a zombie world. Now, if Q_{se} is the conjunction of all phenomenal truths in the super-experiential world, then P and Q_{se} is primarily conceivable, and the super-experiential world is epistemically possible. If (3') is valid, then we can conclude the metaphysical possibility of the super-experiential world. This is a direct refutation of materialism as we have defined it, because the super-experiential world would be a microphysical duplicate of the actual one, which is not a duplicate *simpliciter*: it differs phenomenologically from ours.

(3') If P and Q_{se} is epistemically possible, then P and Q_{se} is metaphysically possible.

In order to show (3') invalid the possibility of 1-COG together with the possibility that we lack cognitive access to our experience is not sufficient, because what we need to show is that the truth of the secondary intension of Q_{se} does not follow from the truth of its primary intension. In other words, we also need to show that from the fact that it *seems* that I am in pain it does not *a priori* follow that I *am* in pain. So, I want to show next that the *a priori* entailment that (3') requires does not hold either.

If nPNA is the case, then it is plausible, for example, that the access we have to the phenomenal character of the experience is less fine grained than the phenomenal character itself. There might be essential features of consciousness we lack cognitive access to. Hence, there will be cases, for example, in which the cognitive access we have to two different experiences is the same one. So, in certain occasions, the access we have to the phenomenal character of our experiences does not allow us to tell whether two experiences have the same phenomenal character or not—even if we can typically distinguish experiences with different phenomenal character. Call ' Exp_a ' and ' Exp_b ' two such experiences. If 1-COG is true, then the concepts of these two experiences, EXP_a and EXP_b , share the primary intension. Both Exp_a and Exp_b *appear* to be the same, in spite of the fact that, *ex-hypothesi*, they are experiences of different kinds—and so they might elicit, for example, different behavioral dispositions, which we might also fail to tell apart—; that is, EXP_a and EXP_b have different secondary intensions. Consequently, from the fact that "it *seems to be* Exp_a " it doesn't follow that "it *is* Exp_a ", because Exp_b also *appears* the same way. The concepts

EXP_a and EXP_b would be similar to the concepts WATER and TWINWATER:⁶ if 1-COG is the case, EXP_a and EXP_b share the primary intension—because we lack cognitive access to the differences between Exp_a and Exp_b —but if nPNA is possible, they might have different secondary intensions. Thus, it is not the case that the truth of the primary intension of a phenomenal truth follows from the truth of its secondary intension. And such an *a priori* entailment is required for the soundness of the conceivability argument.

5.2 Subjective Character and the Conceivability Argument

In reaction, my opponent might be willing to give up the Kripkean intuition to some extent that allows her to embrace the soundness of the conceivability argument. For example by focusing on the subjective character of experience:⁷ I might have well established that there could be essential aspects of an experience that are not captured by the primary intension of the corresponding concept—that there is more to consciousness than appearances, if appearances depend on cognitive access—, but there is one essential aspect of experience, the subjective character, that is always captured by the primary intension. In other words, it is still true that if something *seems to be a conscious state*, then it *is a conscious state*.⁸ Super-experiences are conceivable and this connection between intensions suffices for the conclusion that they are metaphysically possible. So, the conceivability argument still goes through.

The problem with this reply is the following. If we accept the distinction between cognitive access and consciousness, not just nPNA but also nANP is possible. This, together with 1-COG, entails the possibility that the possession of phenomenal concepts depends only causally on

the experience but constitutively on cognitive access. I think that it is even an open empirical question about our cognitive architecture whether the mechanisms of consciousness and the mechanisms of cognitive access can be completely dissociated, so that it might *seem* to one that one is having a conscious experience when one *is not* (it is possible that the primary intension of ‘I am in pain’ is true while the secondary intension is false): cases like Anton’s syndrome or Charles Bonnet syndrome are worth considering as plausible candidates.

Anton’s syndrome is classified as a delusion in which subjects report that they can see in spite of the evidence to the contrary. The most popular explanation for this condition is that they are undergoing some kind of visual experience of the object, say a red apple, they claim to see. But, an alternative explanation is available. It might be the case that they are not undergoing a visual experience at all. They rather think that they are seeing an apple (see Macpherson (2010) for discussion of Anton’s syndrome and a defense of a proposal along similar lines). If nANP is true, then the mechanism of cognitive access corresponding to an experience as of a red apple might be active (say, there is the corresponding activity in the center of the GWS), without the subject having the experience, and she forms the thought on this basis: it *seems* to her that she is having an experience but she *is not*.

Another interesting condition might be Charles Bonnet Syndrome. In this condition subjects report vivid visual hallucinations of faces, persons, objects, and complex geometric forms. However, they are not delusional, they know that their hallucinations are not real—for comparison of Charles Bonnet and Anton’s syndrome see Macpherson (2010). Brown and Lau (forthcoming) suggest that subjects suffering Charles Bonnet syndrome lack first-order representations due to the absence of activity in early perceptual areas (V1). They also suggest that those patients might nonetheless undergo experiences due to the formation of the corresponding higher-order thought.⁹ Here also an alternative explanation is available if one believes that first-order representations are necessary for consciousness and nANP is the case. This explanation would agree with the one offered by Brown and Lau that these subjects form a higher-order thought to the effect that they are having an experience (in virtue of the activation of the corresponding cognitive access mechanisms)—so it *seems* to the subjects that they are having an experience—but deny that they have conscious experiences.¹⁰

⁹ Brown and Lau present Charles Bonnet syndrome as a case of an empty higher-order representation in defense of a higher-order theory of consciousness against Block’s (2011a) objection.

¹⁰ This alternative is neutral on whether the transitivity principle (the principle that in having an experience one is aware of oneself as being in certain state (Rosenthal 1997, 2005), which backs up higher-order

⁶ TWINWATER refers to another substance, XYZ, with the same “appearance properties” that water has (Putnam 1975)

⁷ Following Kriegel (2009), Levine (2001), we can make a conceptual distinction between two components of phenomenal character: the qualitative character and the subjective character. The qualitative character is what distinguishes different kinds of experiences; for example, the kind of experience I have while smelling a rose from the one I have while hearing *Minor Swing*. On the other hand, a theory of subjective character abstracts from the particular ways having different experiences feel and concentrates on the problem of what makes it the case that having a conscious experience feels at all. Hence, the qualitative character is what makes a state the kind of phenomenally conscious state it is, and the subjective character what makes it a phenomenally conscious state at all (Kriegel 2009).

⁸ We have already seen that it doesn’t follow that if something *seems not to be a conscious state*, then it *is not a conscious state*—if 1-COG is possible and it is possible to lack cognitive access to the experience—, which would support the version of the conceivability argument that appeals to zombies.

5.3 Phenomenal Concepts and Experience

The last objection I want to consider is based on the idea that the previous possibility would be at odds with the popular view that possession of phenomenal concepts depends on having undergone the relevant experience—an idea that is embraced by representatives of both sides of the debate about the nature of consciousness. I am going to argue that this objection is only tenable under a notion of phenomenal concepts constructed to meet the Kripkean intuition, and therefore one that begs the question against my proposal in the present discussion. Let me elaborate.

The first thing that should be noted is that the conjunction of nANP and the possibility of I-COG does not deny that experiences play a causal role in the formation of phenomenal concepts. A step further would be to claim that the possibility that they are conjointly true can be ruled out *a priori*, because experiences are somehow constitutive of phenomenal concepts and, as I have shown, if nANP and I-COG are possible, then the possibility of possessing a phenomenal concept without having undergone the relevant experience is left open. In reply, it is worth stressing that not all theories of phenomenal concepts agree with the claim that experiences are constitutive of phenomenal concepts. Balog (2009) argues that those theorists that develop Loar's (1990) idea that a phenomenal concept refers directly, like Tye (2003) or Aydede and Guzeldere (2005), conceive the relation between the phenomenal concept and their experience as different existences related by causation. So, they would also agree that one can conceive that a basic¹¹ application of, say, PAIN be tokened by someone in the complete absence of pain, as I have been arguing. What is more important is that those who defend the *constitutive view*, like Chalmers (2003), Block (2006), Papineau (2006), base their endorsement of such a view on the intimate connection between appearances and reality. This is very explicit in Balog's paper when she objects to direct-reference accounts of phenomenal concepts that "Anybody who tokens a basic application of PAIN is really in pain." (p. 305)¹² But assuming such a connection clearly

Footnote 10 continued

theories of consciousness, is true or not. For a proposal that respects the transitivity principle and that is committed neither to higher-order representations nor to the fact that cognitive access is constitutive of consciousness see Sebastián (2012).

¹¹ Balog (2009) distinguishes basic from non-basic application of phenomenal concepts, where the former have some ontological or explanatory priority over the latter. In the basic use we typically apply phenomenal concepts "to our experiences *directly* as they occur, merely on the basis of having the experience." (ibid. 294). This use is contrasted with cases in which we apply it to others' experiences or experiences we had in the past.

¹² Sure, one can restrict, by definition, basic applications of phenomenal concepts to cases that satisfy this condition. This would

begs the question against my argument in the current debate, for it presupposes the very same intuition that is under discussion. If this is correct, then the possibility that nANP and I-COG are true cannot be ruled out *a priori* and the conceivability argument remains unsound.

We are now in a position to straightforwardly accommodate the alleged conceivability of zombies or super-experiencers without accepting its metaphysical possibility, just as other *a posteriori* necessities. We can (primarily negatively) conceive a microphysical duplicate of the actual world where water is not H₂O. But there is nothing mysterious about it, what we are conceiving is a world with another substance with the same "appearance properties" of water but a world that, if water is H₂O, lacks water. This world would not be a microphysical duplicate of the actual one, so we are not committed to the existence of a microphysical duplicate of the actual world that differs with respect to any property (having water in this case). In the case of consciousness, the reply is similar. Although we can (primarily negatively) conceive worlds that are microphysical duplicates of the actual one that differ from it in the phenomenal properties that are instantiated, these worlds are not metaphysically possible. The reason is that what we are really conceiving is a world whose inhabitants have either the same "cognitive access properties" that we have but lacking consciousness or different "cognitive access properties" and the same conscious states. As in the case of water, such worlds would not be microphysical duplicates of the actual one, so materialism is saved—there is no metaphysically possible world in which P and not Q is the case. We can grant the primary conceivability of beings identical to us who have different experiences. The lack of a *a priori* entailment between cognitive access and consciousness, together with the possibility that the primary intension of phenomenal concepts depends on cognitive access, allows us to reject their metaphysical possibility without committing ourselves to any kind of strong necessities. If my reasoning is sound, the coincidence of appearances and reality is not granted and the entailment from primary conceivability to metaphysical possibility would hence be illegitimate. WT is not valid in the case of consciousness and the conceivability argument should be rejected.

6 Conclusion: A "Not Easy Problem".

I have shown that the Kripkean intuition that there is no distinction between appearances and reality in the case of consciousness—at least as this claim is understood in

Footnote 12 continued

make this statement trivially true and would cut no ice in current discussion.

support of the conceivability argument—depends upon there not being a distinction between cognitive access and consciousness. The open conceptual possibility that (i) conscious experiences and the cognitive access there-to can be dissociated and (ii) that the formation of our phenomenal concepts depends constitutively not on experiences but on the cognitive access we have to them, blocks the conceivability argument. If these hypotheses are in fact true, then the necessary entailment from physical truths to phenomenal truths is similar to other *a posteriori* necessities.

Although I think that this alleviates the Hard Problem, it does not convert it into an “easy” one. Let me explain the reason:

Cohen and Dennett (2011) argue that theories that separate the neural correlates of consciousness and cognitive access cannot be confirmed or shown to be false and are thus outside the scope of science. Although I think they are, at least partially, wrong, there is an important insight in their remark that points to serious problems. The “method” employed by Block (2007, 2011b) seems to be a perfectly scientifically valid one, inference to the best explanation: we look for the framework that makes the most sense of all the data, not just report. The reason why Cohen and Dennett are wrong is even more clear the methodology I have proposed (Sebastián 2014): the areas involved in cognitive access are deactivated during dreams and support for trusting the subjects’ reports on their conscious experience during sleep has been provided by Horikawa et al. (2013)—they decoded the visual content of dream reports upon awakening from fMRI readings of the subject’s highly visual areas while they sleep. Now, one thing is to show that there is consciousness outside cognitive access—as Block and I among others have attempted to show—and a different one is to find the mechanisms that give rise to consciousness. And it is at this point where Cohen and Dennett’s argument backfires.

In order to study consciousness the first-person perspective and the subject’s reports thereof play an essential role. We should trust, at least under certain circumstances, the reports of subjects in order to determine when they are having a certain experience; as (Koch and Tsuchiya (2007), p. 510) remark, “in the absence of compelling, empirical evidence to the contrary... if the subject denies any phenomenal experience, this should be accepted.” Cases like Anton’s syndrome, where subjects are blind but believe and report that they see, or other less rare forms of anosognosia, where subjects deny, for example, their motor disabilities and make all kinds of false reports about their own experiences are clear cases of such “compelling empirical evidence”. Equally, it is not controversial that retrospective reports about experiences we had in the past are not always trustworthy. However, although the

separation of the mechanisms of consciousness and cognitive access does not entail the conclusion that a subject’s report cannot be trusted—a conclusion that might really take consciousness outside the scope of science—, it acknowledges that even *normal subjects* might provide false reports about the experiences they are *currently having* and not just about experiences they had (see fn. 5). This makes the question regarding the conditions under which the subjects’ reports can be trusted (if at all) more pressing. And it is not clear what kind of considerations we should follow in order to ground our answers. Equally significant are other questions in the vicinity like whether the process of cognitively accessing the experience modifies (and to what extent) its content—in other words, if I believe that something is conveyed by my experience, is it really conveyed by the experience or is it part of the result of cognitive access?—or how we are to determine at what level of information processing consciousness arises. I am not claiming that questions like these cannot be answered, just that they suggest a “not easy problem”.

Acknowledgments I am specially indebted to Ken Williford. This paper originates from very fruitful discussions with him and I would like to thank him in the first place. For their comments and suggestions I would like to thank also to Ned Block, Esa Díaz-León, Maite Ezcurdia, Manolo Martínez, Josh Weisberg and two anonymous referees. This paper was presented at the Tecuemepe seminar. I am grateful to the participants in this event for their thoughtful comments, specially to Marc Artiga, Axel Barceló, Eduardo García Ramírez and Ricardo Mena. Financial support for this work was provided by the UNAM-DGAPA-PAPIIT Programs IA400615 and IN401315.

References

- Armstrong D (1999) The mind-body problem: an opinionated introduction. Westview Press, Boulder
- Aydede M, Guzeldere G (2005) Cognitive architecture, concepts, and introspection: an information-theoretic solution to the problem of phenomenal consciousness. *Noûs* 39(2):197255
- Baars BJ (1988) A cognitive theory of consciousness. Cambridge University Press, Cambridge
- Balog K (2009) Phenomenal concepts. In: McLaughlin B, Beckermann A, Walter S (eds) *Oxford handbook in the philosophy of mind*. Oxford University Press, Oxford
- Block N (1995–2002) On a confusion about the function of consciousness. In: Block N (ed.) *Consciousness, function, and representation: collected papers*, vol 1. Bradford Books, Cambridge
- Block N (2006) Max black’s objection to mind-body identity. *Oxford Review of Metaphysics* 3
- Block N (2007) Consciousness, accessibility, and the mesh between psychology and neuroscience. *Behav Brain Sci* 30:481–548
- Block N (2011a) The higher order approach to consciousness is defunct. *Analysis* 71(3):419–431
- Block N (2011b) Perceptual consciousness overflows cognitive access. *Trends Cogn Sci* 12:567–575
- Block N (2012) Response to kouider et al.: which view is better supported by the evidence? *Trends Cogn Sci* 16(3):141–142

- Braun A, Balkin TJ, Wesenten NJ, Carson R, Varga M, Baldwin P, Selbie S, Belenky G, Herscovitch P (1997) Regional cerebral blood flow throughout the sleep wake cycle. An H2(15)O PET study. *Brain* 120:1173–1197
- Brown R, Lau H (forthcoming) The emperor's new phenomenology? the empirical case for conscious experience without first-order representations. In: Pautz A and Stoljar D (eds) *Festschrift for Ned Block*, MIT Press, Cambridge
- Chalmers DJ (1996) *The conscious mind: in search of a fundamental theory*, 1st edn. Oxford University Press, Oxford
- Chalmers DJ (2002) Does conceivability entail possibility? In: Gendler TS, Hawthorne J (eds) *Conceivability and possibility*. Oxford University Press, Oxford, pp 145–200
- Chalmers DJ (2003) The content and epistemology of phenomenal belief. In: Smith Q, Jokic A (eds) *Consciousness: new philosophical perspectives*. Oxford University Press, Oxford
- Chalmers D (2010) *The character of consciousness*. Oxford University Press, Oxford
- Cohen M, Dennett D (2011) Consciousness cannot be separated from function. *Trends Cogn Sci* 15:358–364
- Dehaene S, Naccache L (2001) Towards a cognitive neuroscience of consciousness: basic evidence and a workspace framework. *Cognition* 79:1–37
- Dennett DC (1991) *Consciousness explained*, 1st edn. Back Bay Books, New York
- Dretske F (1995) *Naturalizing the mind*. MIT Press, Cambridge
- Fuster J (2008) *The prefrontal cortex*, 4th edn. Academic Press, London
- Goldman-Rakic PS (1988) Topography of cognition: parallel distributed networks in primate association cortex. *Annu Rev Neurosci* 11:137–156
- Horikawa T, Tamaki M, Miyawaki Y, Kamitani Y (2013) Neural decoding of visual imagery during sleep. *Science* 340(6136):639–642
- Koch C, Tsuchiya N (2007) Phenomenology without conscious access is a form of consciousness without top-down attention. *Behav Brain Sci* 30(5–6):509–510
- Kouider S, de Gardelle V, Sackur J, Dupoux E (2010) How rich is consciousness? the partial awareness hypothesis. *Trends Cogn Sci* 14:301–307
- Kouider S, Sackur J, de Gardelle V (2012) Do we still need phenomenal consciousness? comment on block. *Trends Cogn Sci* 16(3):140–141
- Kriegel U (2009) *Subjective consciousness: a self-representational theory*. Oxford University Press, Oxford
- Kripke SA (1980) *Naming and necessity*. Harvard University Press, Cambridge
- LaBerge S (1988) Lucid dreaming in western literature. In: Gackenbach J, LaBerge S (eds) *Conscious mind, sleeping brain: perspectives on lucid dreaming*. Plenum Press, New York
- Landman R, Spekreijse H, Lamme VAF (2003) Large capacity storage of integrated objects before change blindness. *Vision Res*, 43(2), 149–164. PMID: 12536137. URL: <http://www.ncbi.nlm.nih.gov/pubmed/12536137>
- Lau H, Passingham R (2006) Relative blindsight in normal observers and the neural correlate of visual consciousness. *Proc Natl Acad Sci* 103(49):18763–18768
- Leclair-Visonneau L, Oudiette D, Gaymard B, Leu-Semenescu S, Arnulf I (2010) Do the eyes scan dream images during rapid eye movement sleep? Evidence from the rapid eye movement sleep behaviour disorder model. *Brain* 133:1737–1746
- Levine J (2001) *Purple haze: the puzzle of consciousness*. Oxford University Press, Oxford
- Lewis D (1990) What experience teaches. In: Lycan WG (ed) *Mind and cognition*. Blackwell, Hoboken, pp 29–57
- Loar B (1990) Phenomenal states. *Philosophical perspectives* 4:81–108
- Macpherson F (2010) A disjunctive theory of introspection: a reflection on zombies and antons syndrome. *Philos Issues* 20:226–265
- Maquet P, Peters J, Aerts J, Delfiore G, Degueldre C, Luxen A, Franck G (1996) Functional neuroanatomy of human rapid-eye-movement sleep and dreaming. *Nature* 383:163–166
- Oliveri M, Turriziani P, Carlesimo GA, Koch G, Tomaiuolo F, Panella M (2001) Parieto-frontal interactions in visual-object and visual-spatial working memory: evidence from transcranial magnetic stimulation. *Cereb Cortex* 11(8):606–618
- Papineau D (2006) Phenomenal and perceptual concepts. In: Alter T, Walter S (eds) *Phenomenal concepts and phenomenal knowledge: new essays on consciousness and physicalism*. Oxford University Press, Oxford, pp 111–144
- Phillips I (2011) Perception and iconic memory: what spering does not show. *Mind Lang* 26:381–411
- Putnam H (1975) The meaning of 'meaning'. *Minn Stud Philos Sci* 7:131–193
- Rosenthal DM (1997) A theory of consciousness. In: Block N, Flanagan OJ, Guzeldere G (eds) *The nature of consciousness*. MIT Press, Cambridge
- Rosenthal DM (2005) *Consciousness and mind*. Oxford University Press, Oxford
- Rosenthal DM (2007) Phenomenological overflow and cognitive access. *Behav Brain Sci* 30:521–522
- Searle J (1997) Breaking the hold: silicon brains, conscious robots, and other minds. In: Owen NB, Flanagan J, Guzeldere G (eds) *The nature of consciousness*. MIT Press, Cambridge
- Sebastián MA (2012) Experiential awareness: Do you prefer “it” to “me”? *Philos Top* 40(2):155–177
- Sebastián MA (2014) Dreams: an empirical way to settle the discussion between cognitive and non-cognitive theories of consciousness. *Synthese* 191(2):263–285
- Slighte IG, Scholte HS, Lamme VAF (2008) Are there multiple visual short-term memory stores? *PLoS One* 3:1–9
- Sperling G (1960) The information available in brief visual presentation. *Psychol Monogr Gen Appl* 74(11):1–29
- Tye M (2003) A theory of phenomenal concepts. In: O'Hear A (ed) *Minds and persons*. Cambridge University Press, Cambridge
- Williford K (2007) The logic of phenomenal transparency. *Soochow J Philos Stud* 16:181–195