

# Using Phenomenal Concepts to Explain Away The Intuition of Contingency

Nicholas Shea

## Abstract

Humans can think about their conscious experiences using a special class of 'phenomenal' concepts. Psycho-physical identity statements formulated using phenomenal concepts appear to be contingent. Kripke argued that this intuited contingency could not be explained away, in contrast to ordinary theoretical identities where it can. If the contingency is real, property dualism follows. Physicalists have attempted to answer this challenge by pointing to special features of phenomenal concepts that explain the intuition of contingency. However no physicalist account of their distinguishing features has proven to be satisfactory. Leading accounts rely on there being a phenomenological difference between tokening a physical-functional concept and tokening a phenomenal concept. This paper shows that existing psychological data undermine that claim. The paper goes on to suggest that the recalcitrance of the intuition of contingency may instead be explained by the limited means people typically have for applying their phenomenal concepts. Ways of testing that suggestion empirically are proposed.

## Contents

- (1) Introduction
- (2) Phenomenal Concepts Generate an Intuition of Contingency
  - 2.1 Defining phenomenal concepts
  - 2.2 How phenomenal concepts explain the standard intuitions
- (3) Existing Accounts of the Recalcitrance of the Intuition of Contingency
  - 3.1 Evidence against a phenomenological difference
  - 3.2 Inadequacies in other proposals
- (4) An Alternative Hypothesis
  - 4.1 A source of resistance to identities
  - 4.2 Testing the hypothesis empirically
- (5) Conclusion

Word count: 8,700, inclusive

## (1) Introduction

Although there are good arguments for physicalism, an obstacle to its acceptance is its incompatibility with some deeply-held intuitions. Whatever physical-functional property is offered as a candidate for a psycho-physical identity claim, whether it involves neural substrates, patterns of brain activity (synchrony, recurrence, resonance), informational integration / binding, neurofunctional roles, sensorimotor loops, feedback circuits, or even social interactions, the proposed identity elicits a strong intuition that *that* property could, at best, only be contingently connected to the subjective character of experience. If the connection really were contingent, then there would of course be no identity. Nor is this just a problem with the currently-available theories. The intuition runs deeper: that no putative scientific characterisation could do the job.

These residual dualist intuitions present an obstacle to the acceptance of physicalism. There is a long tradition of seeing it in the first instance as a dualism of concepts.<sup>1</sup> As with theoretical identities like heat = molecular kinetic energy, while we retain separate concepts we can still formulate the possibility claim: it is possible that heat  $\neq$  molecular kinetic energy. Of course, if the identity is true, then it is necessarily true, and the claim about metaphysical possibility is false. Before the identity is known, the dual concepts can be used to express an epistemic possibility (heat  $\neq$  molecular kinetic energy). Even when the identity is known (so that heat  $\neq$  molecular kinetic energy is neither metaphysically nor epistemically possible), an intuition of contingency arises, which can be explained by a closely-related genuine possibility: that something other than molecular kinetic energy might have been responsible for causing the same sensations. But that move is not available to the physicalist who wants to explain away the apparent contingency of a psycho-physical identity, since the sensation itself features in the identity. There is no gap between the referent and the signs by which it is identified, so there is no place in which to locate a nearby genuine possibility. That is Kripke's challenge (Kripke, 1972). Unable to answer it, Kripke draws a dualist conclusion.

The challenge to the physicalist is to explain why, when faced with a psycho-physical identity claim,<sup>2</sup> people should have an intuition of contingency. Physicalists try to discharge that burden by pointing to special features of phenomenal concepts, the class of concepts we

---

<sup>1</sup> (Locke, 1690), the ideas of thought and of solidity are 'independent of one another' (II.xxiii.32), i.e., logically independent ('It is no more a contradiction that thinking should exist separate and independent from solidity than that solidity should exist separate and independent from thinking.' *Ibid*). For Locke this was the basis of property dualism, but he thought that the view was compatible with substance monism (IV.iii.6).

<sup>2</sup> 'Psycho-physical identity claim' is used as shorthand to cover both thoughts and sentences. At the level of thought it involves a phenomenal concept and a physical-functional concept or concepts, connected by the relation usually expressed by 'is identical to' or '='. At the level of language, it is a sentence expressing such a thought.

use for thinking about our own experiences from the first-person point of view. The ‘phenomenal concept strategy’ is to show that we should expect people to have an intuition of contingency, even if physicalism is true and there are true psycho-physical identity claims. The importance of the phenomenal concept strategy is widely acknowledged, even by critics:

‘This extremely interesting strategy is perhaps the most attractive option for a physicalist to take in responding to the problem of consciousness.’

(Chalmers, 2007, p. 168)

The strategy is not intended as an argument for physicalism from neutral premises. Instead, the aim is to diffuse anti-physicalist arguments that are based on the intuition of contingency. The strategy is to presuppose physicalism and show that, if physicalism is true, we should expect people to have a recalcitrant intuition of contingency. The positive support for physicalism derives from other sources, not discussed here. Nevertheless, if a proper understanding of phenomenal concepts can indeed be deployed to explain away the intuition of contingency, then that constitutes significant progress towards a satisfactory metaphysics of mind.

The physicalist argument comes in two stages. The first is to show that a dualism of concepts creates an intuition of contingency. Section 2 individuates a class of phenomenal concepts that people plausibly possess (§2.1). That allows a contingency claim to be formulated in the first place (§2.2). The second stage is to explain why intuition of contingency is not displaced by good evidence for physicalism, unlike with other putative theoretical identities. Section 3 argues that existing accounts fail. Leading proposals rely on there being a general phenomenological difference between tokening a physical-functional concept and tokening a phenomenal concept, the phenomenal concept being phenomenologically richer or more phenomenologically akin to the experience to which it refers. Section 3.1 shows that there is good empirical evidence against that claim. These features are far from peculiar to phenomenal concepts. Other accounts make phenomenal concepts too special, holding that they put the thinker in an especially direct epistemic relation with their referents. Section 3.2 points to inadequacies in such proposals. Section 4.1 argues that the recalcitrance of the intuition of contingency could be explained by a rather prosaic limitation of the phenomenal concepts possessed by most thinkers. Section 4.2 goes on to suggest ways of testing that hypothesis.

## (2) Phenomenal Concepts Generate an Intuition of Contingency

### 2.1 Defining phenomenal concepts

Physicalism comes in several varieties, all of which admit variants of the dialectic discussed here. For present purposes I take physicalism to be the claim that all properties supervene, with metaphysical necessity, on physical properties. Subjective or phenomenal character is sometimes described as a property of experiences, but having an experience with a certain subjective character is a property of a person. For example, it is a property of Jane that she is having an experience as of the colour red. So I will focus on psycho-physical identity claims concerning such *phenomenal properties* of conscious people.

Suppose our best scientific theory tells us that the property of having an experience as of red is complex physical/functional property A (or, if there are relevant phenomenal differences between people, that Jane's having an experience as of red is complex physical/functional property A of Jane). A is a phenomenal property and we can pick it out with a scientific physical-functional description, and also with the folk psychological term 'having an experience as of red' (I use 'Q-red' as shorthand for the folk psychological term). We can also deploy concepts in thought that refer to phenomenal property A. Phenomenal concepts form a proper subclass of the class of all concepts that refer to phenomenal properties.

A first pass at defining phenomenal concepts is that they are the sort of concepts that Jackson's colour scientist Mary could not have in her black and white room (Jackson, 1986). A person who actually has colour experiences can classify them together in virtue of what it's like to undergo those experiences. Loar argues that a subject can thereby acquire a disposition to classify together experiences in virtue of their phenomenal character, and so acquire a concept (Loar, 1990, 1997). If the sort of identity claim at issue here is true, then the phenomenal property shared by these instances is in fact a physical-functional property. Papineau has a different view of phenomenal concepts, but his view also meets the constraint that a thinker can only acquire a phenomenal concept if she has experiences of the type to which the concept refers (Papineau, 2002).

Focusing on this common feature, I will adopt the following definition:

(PC) A person X's concept C is a *phenomenal concept*

iff<sub>df</sub>

- (i) acquiring C requires X to instantiate or to have instantiated a phenomenal property F; and
- (ii) on acquiring C, X has the disposition to apply C to herself non-inferentially in virtue of instantiating the phenomenal property F; and
- (iii) C refers to F.

(PFC) Concept C is a *physical-functional concept*

iff<sub>df</sub>

C refers to a physical-functional property and is not a phenomenal concept.

The definition makes having a phenomenal concept a matter of how it was acquired. It does not rely on phenomenal concepts having a special mode of presentation or canonical evidence for their application (e.g. Sturgeon, 1994). Indeed, the definition is consistent with the rejection of Fregean sense for concepts. Nor is it committed to the subject having the experience to which the phenomenal concept refers every time they token the concept (only on acquisition).

Condition (ii) requires that the concept has a non-inferential mode of application. For example, inferring that I must be angry by looking at my own face in the mirror would be non-inferential within condition (ii). This means that exercises of a phenomenal concept in the mode of application mentioned in PC show up at the personal level as immediate – the move from having the phenomenal experience to applying the phenomenal concept does not seem to be mediated by any intervening mental process. Nevertheless a thinker's having, on acquiring a phenomenal concept, the mode of application mentioned in PC is consistent with her being disposed to apply the concept in other ways as well. Jane can apply ANGER<sub>PHEN</sub> to herself non-inferentially when she feels angry, but she can also apply it to others, perhaps because they show outward signs and behave in ways that are similar to her own reactions when she feels angry.

Condition (iii) requires that, according to the correct theory of the reference of concepts, a phenomenal property F is the referent of C. No claims are made here about what makes it the case that C refers to F. The account is consistent both with phenomenal properties being ways that experience represents the world as being (intentionalism), and with their being raw feels. Concepts of phenomenal properties are clearly representational, but the phenomenal properties themselves need not be.

An experienced blindsight patient might be able to learn to report certain information presented in his blind hemi-field without first performing an action to see how he will in fact behave (de Gelder et al., 1999; Milner & Goodale, 1995; Weiskrantz, 1986). If so, he could deploy a concept HORZ-SLOT<sub>VISUAL</sub> that applies to cases where his dorsal visual processing represents that there is a horizontal slot (without an accompanying experience). Applications of this concept could be non-inferential within PC(i), but HORZ-SLOT<sub>VISUAL</sub> would not be a phenomenal concept because it is not applied in virtue of instantiating a *phenomenal* property.<sup>3</sup>

---

<sup>3</sup> Thanks to Daniel Stoljar for raising this query about the definition.

Definition PC draws the category of phenomenal concepts narrowly, in that concepts falling outside the definition might still generate a recalcitrant intuition of contingency according to the account offered below. However, since it is relatively clear that many people do have phenomenal concepts within the terms of PC, it will be enough for present purposes to show why the intuition of contingency is recalcitrant in those cases. It would be a bonus if the same scheme of explanation could be applied to a slightly wider class of concepts.

The class of phenomenal concepts is not drawn so as to make them too special. I have not made a claim of transparency for phenomenal concepts – that they give the thinker any special or direct way of grasping the nature of their referents. Nor does their application involve infallibility or incorrigibility. Epistemically, they are just like any other concept that admits of non-inferential application (e.g. on some views concepts like WATER and JOHN can be applied non-inferentially on seeing water or John). If concepts have modes of presentation, then the experience referred to by a phenomenal concept will itself be one of the modes of presentation of that concept. But the account is consistent with theories according to which concepts do not have modes of presentation.

Recall that phenomenal properties are properties of the subject of experience, not properties of external objects (like the redness of a fire engine). It is a nice question how people manage to acquire a disposition to classify together their own experiences in virtue of their subjective character, as distinct from classifying external objects in virtue of how those objects appear. One test is that phenomenal concepts are applied to experiences even in the absence of an experienced external object. So John is disposed to token his phenomenal concept of red experiences, RED-EXPC<sub>PHEN</sub> when he recalls an experience of red in his visual memory or imaginatively recreates an experience of red.

Some argue that the thinker must deploy a distinction between the subjective and the objective in order to acquire a disposition which is specific to phenomenal properties (Papineau, 2007a). Certainly such a distinction can be deployed to generate a phenomenal concept. But it is not clear that an explicit grasp of the distinction is necessary. For example, John might form a concept of the feeling of being angry, ANGER<sub>PHEN</sub>, applying it to himself when he is angry in virtue of his instantiating the phenomenal property of feeling angry. That seems possible without an explicit grasp of the subjective-objective distinction. He just tokens ANGER<sub>PHEN</sub> when he has that characteristic feeling. Since John's feeling anger has heterogeneous causes and effects, it seems less likely that ANGER<sub>PHEN</sub> should refer to some property of the objects which tend to cause John to feel angry, nor to the behavioural effects of the feeling. So perhaps John can succeed in referring to his feeling of anger without deploying the sophisticated conceptual apparatus of an explicit subjective-objective distinction. The everyday concept of pain is probably, for most people, a phenomenal concept, as are concepts of relatively general ways of classifying experience, like the concept of visual experience or of

aural experience. It may be that these too can be acquired without deploying an explicit subjective-objective distinction.

## **2.2 How phenomenal concepts explain conceivability**

The phenomenal concept strategy comes in two parts. The first step is very straightforward: a dualism of concepts implies that thinkers are able to formulate contingency claims (e.g. that  $Q\text{-red} \neq A$  is possible). The second, more substantial, task is to point to further considerations to explain why that intuited contingency should not be displaced by good evidence for physicalism, or even by the belief that the identity is true (therefore necessarily so), unlike in the case of theoretical identities like  $\text{heat} = \text{molecular kinetic energy}$  and  $\text{water} = \text{H}_2\text{O}$ . Sections 3 and 4 deal with the second step.

The first step only requires that people should have phenomenal concepts (within definition PC), and that they should deploy concepts in thinking about conscious experience third-personally or scientifically that were not initially acquired in virtue of having the experience in question (i.e. physical-functional concepts within definition PFC). Equipped with two different concepts, whether or not those concepts in fact co-refer thinkers can formulate the thought that it is possible that  $Q\text{-red} \neq A$ ; or the thought that it is possible that a creature that instantiates physical-functional property A nevertheless is a zombie, lacking all conscious experience (thus not instantiating Q-red).

Possession of phenomenal concepts can also explain the intuition that Jackson's Mary learns something new when she leaves her black and white room (Jackson, 1986). In the room she could not have the phenomenal concept  $\text{RED-EXPCE}_{\text{PHEN}}$ . On leaving the room and seeing various coloured objects she is able to acquire this new concept. Jackson's thought experiment works by thinking of Mary acquiring a new concept which she applies non-inferentially to herself in virtue of instantiating the experience (rather than enriching her existing physical-functional concept of red experience to add in that mode of application). Then, using her new concept, she can think thoughts that were not available to her before; thoughts like wondering whether  $\text{RED-EXPCE}_{\text{PHEN}} = \text{RED-EXPCE}_{\text{PHYS-FNAL}}$  or believing that the identity holds. So the phenomenal concept strategy can be deployed to explain the intuition elicited by the Mary thought experiment in a way which is compatible with physicalism.

### **(3) Existing Accounts of the Recalcitrance of the Intuition of Contingency**

#### **3.1 Evidence against a phenomenological difference**

So if people have separate phenomenal and physical-functional concepts, they have the resources to formulate the claim underlying the intuition of contingency. Can the existence of

two concepts be finessed into an account of that recalcitrance of that intuition in the face of good evidence for physicalism? In this section I argue that existing attempts to account for the intuition by means of phenomenal concepts are inadequate. In particular, the common tactic of appealing to phenomenological aspects of the exercise of concepts is undermined by some under-appreciated psychological evidence.

The intuition is doubtless partly due to the fact that we don't yet know how the correct theory of consciousness will go (Shea & Heyes, 2010; Stoljar, 2006). Although clearly important, such considerations can explain resistance to theoretical identities throughout the history of science. They are not peculiar to identities concerning the mental. If the intuition of contingency is particularly recalcitrant when it comes to psycho-physical identities, then the physicalist has some explaining left to do.

The most common tactic adopted by proponents of the phenomenal concept strategy is to claim that the exercise of the two kinds of concepts is phenomenologically different. The idea is that exercising a phenomenal concept has a distinctive experiential aspect – e.g. tokening RED-EXPCE<sub>PHEN</sub> involves having a visual experience as of red – whereas exercise of a physical-functional concept does not.<sup>4</sup> The claim cannot just be that there is a difference in the phenomenology evoked by tokening RED-EXPCE<sub>PHEN</sub> and RED-EXPCE<sub>PHYS-FNAL</sub> (say), because that would be true of all theoretical identities. To target a feature that is plausibly distinctive of psycho-physical identities the phenomenal concept theorist claims that there is a general experiential difference between exercising phenomenal concepts, on the one hand, and exercising physical-functional concepts, on the other.

For Loar, tokens of phenomenal concepts are caused by instances of the experience-type to which they refer. The phenomenology of entertaining a phenomenal concept is thereby so unlike the phenomenology of entertaining a physical-functional concept that, when they are brought together in thinking, 'the illusion may be created that their references must be different' (Loar 1997, p. 605).<sup>5</sup> For Papineau exercise of a phenomenal concept actually makes use of a token experience of the type to which it refers, and thereby feels like having the experience. By contrast, the exercise of physical-functional concepts "leaves out" ... the technical colour phenomenology' (Papineau 2002, p. 170). Tye's explanation of the recalcitrance of the intuition of contingency similarly appeals to the phenomenology of concept exercise

---

<sup>4</sup> The otherwise puzzling idea that a phenomenal concept *resembles* its referent may amount to the same thing. The claim cannot be resemblance at the level of content (it is trivially true that the property referred to by a concept resembles the property that features in the content of that concept), so it must be a claim about the conceptual vehicle, in which case it seems to be phenomenological properties of the vehicle which are in the frame.

<sup>5</sup> This is a different argument from the considerations based on the claim that phenomenal concepts are necessarily connected to their referents, and/or based on the supposed transparency of experience, which are discussed in the next subsection.



(Tye, 2003, §17), and the thought traces right back to the footnote in which Nagel launched the modern phenomenal concepts literature: ‘we put ourselves in a conscious state resembling the thing itself’ (Nagel, 1974, fn. 11).

This tactic fails because the phenomenological difference between exercise of phenomenal and physical-functional concepts relied on to date is empirically unsustainable; and there are reasons to doubt that there is any relevant difference in the phenomenology of the exercise of the two types of concepts that could do the required explanatory work. My objection has two strands. First, exercise of physical-functional concepts can carry perception-like phenomenological properties. Second, exercise of phenomenal concepts need not.

A series of experiments show that perceptual phenomenology is generally involved in concept exercise, across a wide range of physical-functional concepts. For example, cueing a particular perceptual modality affects which properties items in a given category are stated to have (Barsalou, Solomon, & Wu, 1999) and how quickly and accurately subjects verify whether or not items have such properties (e.g. whether LEAVES rustle) (Barsalou et al., 2005; Solomon & Barsalou, 2004). Similar results are obtained whether the task is presented abstractly or subjects are asked to engage perceptual imagery, suggesting that perceptual imagery is involved in both cases. Functional imaging confirms that brain areas involved in perception and perceptual imagery are also active in tasks that would seem only to require non-perceptual exercise of a concept (Kan et al., 2003; Simmons & Barsalou, 2003).<sup>6</sup>

The second strand of the objection is that exercise of phenomenal concepts need not involve the particular phenomenology to which the concept refers. Consider the thought involving RED-EXPC<sub>E</sub><sub>PHEN</sub> which is expressed by the following: I am not now having a red experience in perception or imagination. That thought deploys the phenomenal concept RED-EXPC<sub>E</sub><sub>PHEN</sub>. And it seems that we can sometimes when we think such a thought, it is true. If so, the phenomenal concept can be deployed without instantiating the phenomenal property to which it refers. It seems clear that, once one has acquired a phenomenal concept through instantiating the experience to which it refers, one can then exercise that concept in a range of ways that do not involve re-activation of experiences of the original type. Such ‘non-recreative’ uses of a phenomenal concept form the second strand of the objection.<sup>7</sup>

Caught between these two strands, it is hard to see that there could be a general phenomenological difference between the exercise of phenomenal and physical-functional concepts. Both phenomenal concepts and physical-functional concepts can be tokened with and without accompanying perceptual-like phenomenology. That makes it hard to point to a phenomenological difference in concept exercise that could explain the recalcitrance of the

---

<sup>6</sup> Barsalou argues that modality-specific processing is constitutive of concept exercise; for our purposes it is enough that perceptual phenomenology is usually associated with exercise of physical-functional concepts.

<sup>7</sup> The objection is pressed by (Crane, 2005), but is independently credited to Kirk Ludwig by (Block, 2007, p. 282).

intuition of contingency; indeed, Papineau has given up on this tactic (Papineau, 2007a, 2007b).

One way to try to hang onto the phenomenal concepts strategy here is to focus on occurrent uses. The intuition of contingency is particularly compelling in the philosopher's situation of applying a phenomenal concept to an aspect of her current perceptual experience (the experience being produced now by looking at a red folder, say) and at the same time entertaining a potential identity with a physical-functional property. But there is a danger of that approach missing out on there being a *concept*, a re-usable element of thought and reasoning, which is applied to the experience. If there is a concept involved in the occurrent case, then that concept can be re-used in the ways discussed above. Without involvement of a concept, the occurrent case serves only to focus on the characteristic what-it's-likeness of the occurrent experience. That is not, on its own, a physicalist explanation of the intuition of contingency. Furthermore, if the existence of current phenomenology were enough to generate an intuition of contingency, we should expect to find one with perceptual demonstrative concepts too. But there seems to be no difficulty in identifying, for example, *that* shininess as being a property of the polished table surface, or indeed *that* redness as being a property of the surface of the folder.

Those relying on a special experiential feature of phenomenal concepts need to point to an experiential difference between the exercise of phenomenal concepts and the exercise of physical-functional concepts. Loar and Papineau appeared to claim that exercise of a phenomenal concept always involves re-activating the experience to which it refers, which is undermined by the clear existence of non-recreative uses of phenomenal concepts. The claim could instead be that exercise of phenomenal concepts sometimes involves an associated phenomenology whereas exercise of physical-functional concepts never does. That would be a clear difference between the two classes of concept. But it is very unlikely to be true, in the light of data about the extent to which perceptual experiences are recreated by conceptual thought (data from Barsalou and others mentioned above). So relying on an experiential difference to explain the recalcitrance of the intuition of contingency is unpromising.

### **3.2 Inadequacies in other proposals**

Before giving my positive account, I will mention, more briefly, some other variants of the phenomenal concept strategy. Papineau puts some weight on phenomenal concepts being indexical: the subject thinks about *this* \_\_\_ type of experience (where \_\_\_ is replaced in thought by an instance of the experience). If the concept could only ever be used indexically, then the problem of accounting for non-recreative uses arises. If the problem is supposed to be generated by the indexical way of introducing or acquiring the concept, it faces counterexamples. I might rely on a demonstrative to introduce a concept of a particular way

of dancing, say: it is to dance *thus*. That doesn't make me resistant to a physical-functional way of picking out that type of behaviour. Nor is the concept any more recalcitrant if I first use it in my own case. For example, I might think about a particular way I have of combing my hair, forming a concept by thinking demonstratively as I brush (to C is to comb *thus* \_\_\_) ((Heal, 1997), see also (Heal, 2004)). Such concepts do not seem to resist merger or enrichment with (further) physical-functional modes of application, nor do they seem to generate a recalcitrant intuition of contingency.

Another tactic is to claim that phenomenal concepts stand in an especially tight epistemic relationship with their referents. For example, Loar argues that phenomenal concepts are transparent (Loar 1990, pp. 98-99). The notion of transparency can seem elusive to a physicalist. The rough idea is that exercise of the concept gets us on to the essential nature of the property to which it refers. But direct reference does not imply transparency. That we have concepts that we can apply non-inferentially to ourselves in virtue of instantiating their referents does not by itself show that exercise of those concepts puts us in any special epistemic relation to their referents. Indeed, if transparency is a matter of knowing facts about what is essential to or constitutive of being a phenomenal property, exercise of phenomenal concepts seems, taken alone, to deliver very little in the way of such knowledge. So if transparency is supposed to be a rich source of knowledge, it is implausible that any of our concepts display transparency. On the other hand, transparency may just be the idea that the concept is applied on the basis of a necessary connection to the referent. That would be true if there were phenomenal concepts which have the experience to which they refer as a their constitutive mode of presentation. But then we are back into the empirical difficulty with claiming that this is true in any sense which is distinctive of phenomenal concepts.

Furthermore, Chalmers argues that the phenomenal concepts strategy cannot succeed if it relies on phenomenal concepts having an epistemic feature that uniquely depends upon having experiences (Chalmers, 2007). If that were so, a zombie (a physical-functional duplicate) that lacked experiences would lack the special epistemic feature, undermining the claim that the special epistemic feature is physicalistically explicable. The phenomenal concept strategist might object that such a zombie is impossible, leading us back into familiar debates. But there is something to Chalmers' idea that, if the phenomenal concept strategist relies on an epistemic feature that is unique to phenomenal concepts and is unable to elaborate on that feature in non-phenomenal terms, his account of the recalcitrance of the intuition of contingency will be correspondingly mysterious from the physicalist point of view, leaving the door open to the dualist, who will seize on the supposed unique epistemic feature and offer a more satisfying explanation of it in terms of the ontological difference between phenomenal properties and physical-functional ones. My positive account of the recalcitrance of the intuition of contingency, given in the next section, relies on a rather prosaic and

contingent feature of our application of phenomenal concepts, which can be fully explained without relying on the phenomenality of their referents (indeed, which can arise in cases where the referent is not phenomenal at all, as in the blindsight case mentioned above). So it side-steps Chalmers' objection.

This section has not attempted a comprehensive review of all the variants of the phenomenal concept strategy. Instead, I have argued that the most common tactic, which is to rely on an experiential difference between entertaining phenomenal concepts and entertaining physical-functional concepts, is untenable. I have mentioned difficulties with other versions of the strategy, and while this falls far short of a decisive refutation, it should be sufficient to motivate searching elsewhere for a positive account of the recalcitrance of the intuition of contingency. The next section offers such an account.

#### **(4) An Alternative Hypothesis**

##### ***4.1 A source of resistance to identities***

This section puts forward an alternative explanation for the recalcitrance of the intuition of contingency. §4.2 suggests ways of testing that hypothesis empirically.

We start with an illustrative analogy. William Bateson famously rejected the idea that chromosomes could be the basis of the inheritance of heritable phenotypic features (Darden, 1977). Walter Sutton and Theodor Boveri had hypothesised that chromosomes are the basis of heredity on the basis of a close match between some of their properties, like the fact that they operate in pairs (Darden & Maull, 1977, p. p. 52). Morgan, who was initially sceptical, became convinced of the theory by work in his laboratory on linking of factors on chromosomes and crossing over (Darden, 2005). Bateson's particular resistance to the chromosome theory was at least partly because of his aversion to microscopy and the fact that he lacked a close colleague skilled in cytology (Cock, 1983). He accepted that chromosomes could be treated as if it were the basis of heredity, but he seemed to take this as a kind of theoretical posit that should be taken merely instrumentally. Without being able to see or manipulate the structure of nuclear chromatin, he just could not understand how it could be the basis of heredity, which led him to the claim that it was *inconceivable* that it could:

'In our present ignorance of the nature of life we cannot distinguish cause and effect in these phenomena and it is not possible to attach any satisfactory meaning to the expression that the sex-linked factors [i.e. heritable differences] are "carried" by a chromosome, but if any one wishes to describe the association or the phenomena in that way there is nothing to forbid it. The properties of living things are in some way

attached to a material basis, perhaps in some special degree to nuclear chromatin; and yet it is *inconceivable* that particles of chromatin or of any other substance, however complex, can possess those powers which must be assigned to our factors or genes [genes being the hypothesised basis of heritable differences]. The supposition that particles of chromatin, indistinguishable from each other and indeed almost homogeneous under any known test, can by their material nature confer all the properties of life surpasses the range of even the most convinced materialism.’

(Bateson, 1916, p. 542, emphasis added)

The correspondence with today’s scepticism about mind-brain identities is striking. Replacing heritable features with phenomenal properties and particles of chromatin with properties of the brain converts Bateson’s text into a recognisable example of current scepticism towards materialism about phenomenal properties, even paralleling claims about inconceivability. Of course, now that we know the structure of DNA and so much about its mechanism of operation, the identity – basis of heredity = chromosomes – is uncontroversial (at least in relation to very many heritable features).

Bateson had various ways of identifying chromosomes (chemical etc.) which together gave him a CHROMOSOME concept. He also had the concept of THE CAUSAL BASIS OF HEREDITY, identified in part through the outcomes of breeding experiments. But he didn’t believe the identity CHROMOSOMES = THE CAUSAL BASIS OF HEREDITY. Nor could he see any way that he could come to believe the identity. So he found it inconceivable that the identity could be true. What he lacked (in part because of his limited facility with microscopy) was a way of identifying features of chromosomes, under his CHROMOSOME concept, in a way that made it plausible that chromosomes could also be identified under his CAUSAL BASIS OF HEREDITY concept. Conceptual dualism made non-identity conceptually possible, and particular limitations in his ways of deploying those two concepts gave rise to a forceful intuition of contingency.

Concepts of physical properties can be enriched by acquiring new modes of application, as we learn more about the property. One important way of learning about a physical property is by spatio-temporally tracking objects that have the property. (We obviously learn about individuals by tracking them spatio-temporally, but it is less widely-recognised that spatio-temporal tracking is important for learning about properties and how they tend to be co-instantiated.) For example, I may apply my CAT concept on the basis of an object being medium-sized, furry and making a miaow sound. I can track this particular cat and learn that it purrs when stroked. I don’t need to check that the purring thing still has the features of a cat, including checking again that it miaows, since I have already heard it miaow, and then I have tracked the object spatio-temporally to the one on which I am now experimenting by stroking it. In this way, we can apply a property concept in part on the basis of hidden or dispositional properties of an object, track the object, and thereby later come to learn that

further properties tend to be co-instantiated with the initial property (Millikan, 2000). This spatio-temporal tracking also allows us to intervene to learn the effects of various actions: to see which properties give rise to which other properties in objects of that type. In this way, spatio-temporal tracking allows us to learn more, not just about the particular object, but also about things of its type, in particular to learn which other properties tend to be co-instantiated with that type. So a property concept that starts with a very thin set of modes of application can be enriched into one that we can apply across vary many different circumstances relying on a large variety of properties.

Tracking is more than reidentifying. If A is the physical-functional property that is identical to having an experience as of red (recall that we are presupposing physicalism), then a thinker can use his phenomenal concept RED-EXPC<sub>PHEN</sub> to reidentify physical-functional property A. He can use it to tell when he himself is instantiating A. Tracking requires something more: following that property instance spatio-temporally so as to be able to learn more about its patterns of co-instantiation. There are serious limitations in the extent to which we can enrich phenomenal concepts by spatio-temporal tracking, at least using everyday means, because of the rather banal fact that their referents are instantiated in our heads. The brain is responsible for the difference between being in one phenomenal state and another, if not wholly, then at least largely and importantly so. When a thinker applies a phenomenal concept non-inferentially to himself in virtue of instantiating its referent there is little he can do to keep track of that instance of the property spatio-temporally. He can keep track of himself, but that does nothing to differentiate one phenomenal property from another. He can learn about typical causes and effects (fire engines typically cause red experiences, feelings of anger typically cause shaking limbs, etc.). But these modes of application don't allow the thinker to pick out an instance of the phenomenal property which he can follow spatio-temporally with sufficient fineness of grain so as to learn more about *it*, as opposed to the multitude of other things that are going on in his head. The very feature that makes phenomenal concepts so easy to deploy (non-inferentially to oneself in virtue of instantiating an experience) makes it hard to enrich them with many third-person modes of application.

As a result, when we are faced with the properties identified by scientists that figure in putative physical theories of phenomenality, properties like brain scan traces, cyto-architectural features of neurons and electrophysiological recordings of firing rates, we have no way of seeing how those modes of identification could be ways of identifying the very same property as the property we pick out with a phenomenal concept. The same thing happened to Bateson with chromosomes. Practical limitations on his ways of visualising, identifying, manipulating and intervening on chromatin meant that he had no way of plausibly enriching his CHROMOSOME concept with the causes and effects of heredity. Just as limitations on spatio-temporal tracking disabled Bateson from enriching his CHROMOSOME concept, they also prevented him from being able to see how it could co-refer with this BASIS OF HEREDITY concept. That in

turn led to his claim that the identity was inconceivable.

When a set of modes of identification C2 stands apart and cannot be integrated with the modes of identification associated with an existing concept C1, then C2 cannot but form a second concept. The result will be conceptual redundancy, unknowingly making use of two concepts for the same thing, as Bateson did. One major way that we can come to notice the redundancy is by tracking an object that falls under the first concept and noticing that it also has the property picked out by the second concept. I learn something about the property of being a raconteur when I predicate RACONTEUR IN A CROWD of someone at a party and find, when I track them to a quiet corner later, that the same person falls under the concept TACITURN ONE-TO-ONE. In cases where co-instantiation is not apparent at a time, spatio-temporal tracking can give us evidence of co-occurrence. When there is a rich and detailed match in coinstantiated properties we can have good reason to conclude that we have been operating with two ways of thinking about the same property – that there are not really two different properties at all (as in the chromosome / heredity case). Discovering co-reference does not necessitate that we have only one concept, but discovering the identity licences us to merge the two concepts into a single concept, which conjoins the modes of application of the two.

Similarly with phenomenal properties, the limitations on spatio-temporal tracking that usually prevent us enriching our phenomenal concepts with third-personal and scientific means of identifying phenomenal properties also block us gathering the kinds of evidence that would ordinarily convince us of co-reference. To do that in the phenomenal case, we would have to exercise a phenomenal concept, RED-EXPC<sub>E<sub>PHEN</sub></sub> say, and then track the particular brain state that makes it the case that I am experiencing red while seeing that further properties (functional, neural, etc.) are predicable of it. Prosaic limitations on spatio-temporal tracking of phenomenal properties identified under a phenomenal concept prevent us from doing so. When we apply a phenomenal concept, we may succeed in picking out a phenomenal property, but we have no way of tracking it spatio-temporally, other than just by tracking the person who is having that experience, with all of her psychological properties. So limitations on spatio-temporal tracking may explain why we resist identities between phenomenal and physical-functional concepts, and thereby have a recalcitrant intuition of contingency.

The resources which account for the intuition of contingency can also form part of an explanation of the explanatory gap.<sup>8</sup> As usually understood, the explanatory gap arises for those who believe there are psycho-physical identities but think there is an obstacle to obtaining a satisfactory explanation of why they hold. My hypothesis is that limitations on the ability to enrich phenomenal concepts with third-personal or scientific modes of application, and the correlative absence of the kinds of evidence that would convince us that phenomenal

---

<sup>8</sup> Doubtless the absence of a plausible real theory to figure in the identity claims also has a role to play (cf Shea & Heyes, 2010), as mentioned above (Stoljar, 2001, 2006).

concepts and physical-functional concepts could co-refer, make us feel an explanatory gap. The felt gap is the absence of the kinds of considerations that would usually convince us that an identity claim is true. However, this limitation is not deep-rooted. It is practical and contingent. So the account makes empirical claims that may be falsified. The final subsection suggests how the hypothesis might be tested.

#### **4.2 Testing the hypothesis empirically**

An unusual but attractive feature of the positive story offered above is that it makes an empirical prediction. If the recalcitrance of the intuition of contingency depends partly on contingent features of phenomenal concepts which make it hard track their instances spatio-temporally, then that limitation may be surmountable. There is no in-principle reason why a phenomenal property cannot be tracked starting with a first-personal mode of application. If physicalism is true, the property causing me to predicate RED-EXPC<sub>PHEN</sub> of myself is some physical property of me (body and brain). Just as in a third-personal investigation of consciousness, I can in principle observe the changes in my own brain that are characteristic of occasions when I predicate RED-EXPC<sub>PHEN</sub> of myself. I can then keep track of that state so as to learn further things about it. The first-personal mode of application is not an insuperable obstacle to reliance on tracking to enrich a concept.

One situation in which we could track instances of phenomenal properties with greater precision is when we are ourselves the subject of an experiment. Then we could see in a fine-grained way how the referent of a phenomenal concept covaried with some physical-functional properties (e.g. via the read-out of a brain scan), both temporally (when the experience is happening) and spatially (where the brain activity is located).

Experimental techniques may help people to track phenomenal properties spatially and temporally. Event-related potentials (ERPs: measures of electrical activity on the scalp) and magnetoencephalography (MEG: measures of changes in magnetic field caused by currents in the brain) can keep track of brain properties at a very fine temporal resolution, allowing us to predicate properties our brain has at the same time as we are applying a phenomenal concept (e.g. of red experience). Correlatively, fMRI and PET deliver much better spatial tracking. And transcranial magnetic stimulation (TMS) allows us to intervene on phenomenal properties and learn for ourselves the effects of such interventions, conceptualising those effects both in scientific terms and under first-person applications of phenomenal concepts. The TMS pulse is characterised in the theory of electromagnetism and its effect on the brain in the theory of electrophysiology; but I simultaneously experience the effect myself (e.g. I see the phosphemes), enabling me to deploy my phenomenal concepts in relation to it. Neurobiological feedback is a further source of potentially relevant experience (Weiskopf et al., 2004). The positive story offered above predicts that subjects of these kinds of experience



will come to feel the intuition of contingency with less force.

This kind of evidence would give subjects good reason to merge phenomenal and physical-functional concepts, but it could also stop them forming separate physical-functional concepts in the first place. Instead they may proceed from the start of scientific enquiry by enriching their phenomenal concept with new third-personal or scientific modes of application. Their concept would still count as a phenomenal concept within definition PC, since it was acquired through predicating it of oneself while instantiating the phenomenal property, but it would gradually lose the feature that makes for the recalcitrant intuited contingency, since it would not be stuck with a limited set of modes of application that resist integration with other third-person modes of application. If a subject starts with two concepts and comes to merge them, then the merged concept would also arguably fall under PC, but would again have lost the (contingent) feature that leads many people's phenomenal concepts to generate an illusion of contingency.

Some preliminary confirmation of this hypothesis can be had by asking psychologists and cognitive neuroscientists about the intuition of contingency. Anecdotally, they do indeed tend not to feel the force of the intuition of contingency. Since that could be an effect of scientific instruction, or even because of selection effects, surveying the intuitions of scientists would not be a good way to test the hypothesis. An appropriate test would be to take two matched groups of naïve subjects, giving one group experiences of the type suggested above, while giving the other group an equal amount of additional knowledge about phenomenal properties, but doing so only via physical-functional concepts (as if they were Mary in the black and white room). If the positive story I have sketched is right, then the first group should, after training, be less wedded to the intuition of contingency, relative to the second group. (The extra scientific knowledge may have an effect of lessening the intuition of contingency even in the second group, but the link to first-personal applications in the first group should produce a larger effect.) Existing evidence makes this hypothesis sufficiently plausible to be worth testing, but only such tests will show whether it is correct.

## **(5) Conclusion**

The intuited contingency between the physical and the phenomenal is incompatible with physicalism. However, if physicalism is true, the contingency can be explained away. We have a special class of phenomenal concepts of our own experiences. This allows us to formulate contingency, despite its metaphysical impossibility. Physicalists are still faced with the challenge of explaining why the intuition of contingency should be recalcitrant in the face of good evidence that physicalism is true, and thus that some identities between phenomenal concepts and physical-functional concepts obtain, and so are necessarily true. Leading

attempts to do so rely on there being a phenomenal difference between thinking with physical-functional concepts and thinking with phenomenal concepts. Existing psychological data make that claim unsustainable.

Other suggestions for explaining the recalcitrant intuition of contingency are also problematic. But the conclusions of this paper are not wholly negative. It suggests an alternative explanation, based on prosaic limitations people usually face in tracking instances of phenomenal properties, when they are identified in the first-personal non-inferential way that is distinctive of phenomenal concepts. That hypothesis can be tested by seeing whether scientific techniques that enrich our phenomenal concepts with third-person modes of application also reduce the force of the intuition of contingency. The intuition has had a strong influence in philosophy of mind. Explaining it away would remove a substantial obstacle to physicalism about phenomenal properties.

### Acknowledgements

The author would like to thank Derek Ball, Tim Bayne, David Charles, Martin Davies, David Papineau, David Rosenthal, Helen Steward and Daniel Stoljar for discussion and comments on earlier drafts; audiences at the 10th annual meeting of Association for the Scientific Study of Consciousness, the Research School for Social Sciences at the Australian National University, the PPNB conference at the University of Bristol, and in Oxford for helpful discussion; and the editor William Bechtel and two anonymous referees for comments on a previous draft. This work was supported by: the Wellcome Trust (grant number 086041), the Oxford Martin School and the John Fell OUP Research Fund.

### References

- Barsalou, L. W., Pecher, D., Zeelenberg, R., Simmons, W. K., & Hamann, S. B. (2005). Multi-modal simulation in conceptual processing. In W. Ahn, R. Goldstone, B. Love, A. Markman & P. Wolff (Eds.), *Categorization inside and outside the lab: Essays in honor of Douglas L. Medin* (pp. 249-270). Washington DC: American Psychological Association.
- Barsalou, L. W., Solomon, K. O., & Wu, L. L. (1999). Perceptual simulation in conceptual tasks. In M. K. Hiraga, C. Sinha & S. Wilcox (Eds.), *Cultural, typological, and psychological perspectives in cognitive linguistics: The proceedings of the 4th conference of the International Cognitive Linguistics Association, Vol. 3* (pp. 209-228). Amsterdam / Philadelphia: John Benjamins Publishing Company.
- Bateson, W. (1916). Review of *The Mechanism of Mendelian Heredity* by T. H. Morgan et al. *Science*, 44, 536-543.

- Block, N. (2007). Max Black's Objection to Mind-Body Identity. In T. Alter & S. Walter (Eds.), *Phenomenal Concepts and Phenomenal Knowledge* (pp. 249-306). Oxford / New York: OUP.
- Chalmers, D. C. (2007). Phenomenal concepts and the explanatory gap. In T. A. Alter & S. Walter (Eds.), *Phenomenal Concepts and Phenomenal Knowledge: New Essays on Consciousness and Physicalism* (pp. 167-194). Oxford / New York: O.U.P.
- Cock, A. G. (1983). William Bateson's rejection and eventual acceptance of chromosome theory. *Annals of Science*, 40(1), 19-59.
- Crane, T. (2005). Papineau on phenomenal concepts. *Philosophy and Phenomenological Research*, 71(1), 155-162.
- Darden, L. (1977). William Bateson and the Promise of Mendelism. *Journal of the History of Biology*, 10(1), 87-106.
- Darden, L. (2005). Relations among fields: Mendelian, cytological and molecular mechanisms. *Studies in History and Philosophy of Biology and Biomedical Sciences*, 36, 349-371.
- Darden, L., & Maull, N. (1977). Interfied theories. *Philosophy of Science*, 44(1), 43-64.
- de Gelder, B., Vroomen, J., Pourtois, G., & Weiskrantz, L. (1999). Non-conscious recognition of affect in the absence of striate cortex. *NeuroReport*, 10(18), 3759-3763.
- Heal, J. (1997). Indexical Predicates and Their Uses. *Mind*, 106, 619-640.
- Heal, J. (2004). Minds, Brains and Indexicals. Retrieved 2nd August, 2011, from <http://www.nyu.edu/gsas/dept/philo/courses/representation/papers/HealMinds.html>
- Jackson, F. (1986). What Mary Didn't Know. *Journal of Philosophy*, 83, 291-295.
- Kan, I. P., Barsalou, L. W., Solomon, K. O., Minor, J. K., & Thompson-Schill, S. L. (2003). Role of mental imagery in a property verification task: fMRI evidence for perceptual representations of conceptual knowledge. *Cognitive Neuropsychology*, 20, 525-540.
- Kripke, S. (1972). *Naming and Necessity*. Oxford: Blackwell.
- Loar, B. (1990). Phenomenal States. In J. Tomberlin (Ed.), *Philosophical Perspectives IV: Action Theory and Philosophy of Mind* (pp. 81-108). Atascadero: Ridgeview.
- Loar, B. (1997). Phenomenal states [second version]. In N. Block, O. Flanagan & G. Güzeldere (Eds.), *The Nature of Consciousness* (pp. 597-616). London / Cambridge MA: MIT Press.
- Locke, J. (1690). *An Essay Concerning Human Understanding*.
- Millikan, R. G. (2000). *On Clear and Confused Ideas*. Cambridge: Cambridge University Press.

- Milner, A. D., & Goodale, M. A. (1995). *The Visual Brain in Action*. Oxford: OUP.
- Nagel, T. (1974). What is it like to be a bat? *Philosophical Review*, 83, 435-450.
- Papineau, D. (2002). *Thinking About Consciousness*. Oxford: O.U.P.
- Papineau, D. (2007a). Phenomenal and perceptual concepts. In T. Alter & S. Walter (Eds.), *Phenomenal Concepts and Phenomenal Knowledge: New Essays on Consciousness and Physicalism* (pp. 111-144). Oxford: OUP.
- Papineau, D. (2007b). Phenomenal concepts are not demonstrative. In M. M. McCabe & M. Textor (Eds.), *Perspectives on Perception* (pp. 87-110). Lancaster / Frankfurt: Ontos Verlag.
- Shea, N., & Heyes, C. M. (2010). Metamemory as evidence of animal consciousness: the type that does the trick. *Biology and Philosophy*, 25, 95-110.
- Simmons, W. K., & Barsalou, L. W. (2003). The similarity-in-topography principle: Reconciling theories of conceptual deficits. *Cognitive Neuropsychology*, 20, 451-486.
- Solomon, K. O., & Barsalou, L. W. (2004). Perceptual simulation in property verification. *Memory & Cognition*, 32, 244-259.
- Stoljar, D. (2001). The conceivability argument and two conceptions of the physical. In *Philosophical Perspectives*, 15, *Metaphysics* (pp. 393-413). Oxford: Blackwell.
- Stoljar, D. (2006). *Ignorance and Imagination: on the epistemic origin of the problem of consciousness*. Oxford / New York: Oxford University Press.
- Sturgeon, S. (1994). The epistemic view of subjectivity. *Journal of Philosophy*, 61(5), 221-235.
- Tye, M. (2003). A theory of phenomenal concepts. In A. O'Hear (Ed.), *Minds & Persons: Royal Institute of Philosophy Supplement 53* (pp. 91-105). Cambridge: CUP.
- Weiskopf, N., Scharnowski, F., Veit, R., Göbel, R., Birbaumer, N., & Mathiak, K. (2004). Self-regulation of local brain activity using real-time functional magnetic resonance imaging (fMRI). *Journal of Physiology - Paris*, 98, 357-373.
- Weiskrantz, L. (1986). *Blindsight: A case study and implications*. Oxford: OUP.