



# A Posteriori Physicalism and the Discrimination of Properties

Philip Woodward<sup>1</sup> 

Received: 4 April 2017 / Accepted: 24 July 2017 / Published online: 2 August 2017  
© Springer Science+Business Media B.V. 2017

**Abstract** According to a posteriori physicalism, phenomenal properties are physical properties, despite the unbridgeable cognitive gap that holds between phenomenal concepts and physical concepts. Current debates about a posteriori physicalism turn on what I call “the perspicuity principle”: it is impossible for a suitably astute cognizer to possess concepts of a certain sort—viz., narrow concepts—without being able to tell whether the referents of those concepts are the same or different. The perspicuity principle tends to strike a posteriori physicalists as implausibly rationalistic; further, a posteriori physicalists maintain that even if the principle is applicable to many narrow concepts, phenomenal concepts have unique features that render them inferentially isolated from other narrow concepts (a dialectical move known as “the phenomenal concept strategy” (PCS)). I argue, on the contrary, that the case for the perspicuity principle is quite strong. Moreover, not only have versions of the PCS repeatedly failed, likely, all versions will, given the strange combination of lucidity and opacity that the PCS has to juggle (it requires that we come up with a lucid explanation of an irremediable cognitive blindspot). I conclude that a posteriori physicalists currently lack a principled objection to classic anti-physicalist arguments.

## 1 Introduction

Clarifying the mind-body problem has provided occasion for philosophers to clarify much else besides. This paper is in keeping with that trend. The paper covers two issues: first, it is about prospects for a certain version of physicalism about consciousness, viz., a posteriori physicalism. Second, it is about the extent to which a subject’s possession of concepts affords her the ability to tell whether the properties picked out by those concepts are the same or different—a matter

---

✉ Philip Woodward  
Philip.woodward@valpo.edu

<sup>1</sup> Department of Philosophy, Valparaiso University, 1400 Chapel Dr., Valparaiso, IN 46383, USA

that has remained in the background of the mind-body problem for some time. (The relationship between the two issues, if not initially obvious, should presently become clear.) I will argue that it is impossible for a suitably astute cognizer to possess concepts of a certain sort—viz., narrow concepts—without being able to tell whether the referents of those concepts are the same or different. I call this claim “the perspicuity principle.” If it is true, it rules out standard versions of a posteriori physicalism. But it is hard to see how a posteriori physicalists can motivate denying it.

In Sects. 2 and 3, I introduce the perspicuity principle in the context of two decades’ worth of debate about phenomenal concepts. In Sects. 4 and 5, I provide arguments in favor of the perspicuity principle. In Sect. 6, I explain why a posteriori physicalists’ attempts to resist these arguments do not, and probably cannot, succeed.

## 2 From the Semantic Premise to the Perspicuity Principle

In his famous paper *Phenomenal States*, Loar (1997) observes that influential anti-physicalist arguments (such as Frank Jackson’s Knowledge Argument and Saul Kripke’s Modal Argument) turn on a disputable assumption about phenomenal concepts. Concepts, for Loar, are psychological particulars with sub-propositional semantic contents. These semantic contents come in at least two types: a *denotation* and a *connotation*. A concept’s denotation is the referent of the concept—the property such that, if an entity instantiates that property, it falls under the concept’s extension. A concept’s connotation is the property, or cluster of properties, whereby the subject represents to herself the concept’s denotation. Loar uses these semantic categories to formulate the assumption he disputes, which he calls “the Semantic Premise”:

**The Semantic Premise** A statement of property identity that links conceptually independent concepts is true only if at least one concept picks out the property it refers to by connoting a contingent property of that property.

In other words, if properties F and G are identical, then if some subject does not know a priori that they are identical, the subject must be conceptualizing F or G in a semantically indirect way. For example, the proposition *water is H<sub>2</sub>O*, a necessarily true identification, is not a conceptual truth. And this is because our concept of *water* is “wide”: it denotes H<sub>2</sub>O *indirectly*, by connoting contingent properties of H<sub>2</sub>O (viz., the constellation of macro-level features—appearance, canonical location, etc.—whereby we recognize water as water). The semantic premise generalizes: *any* case of an inability on the part of a suitably astute subject to “just see” the truth of a property-identification is to be explained in the same way. But this is a prima facie problem for those who would identify phenomenal properties with physical properties, because phenomenal concepts are not wide concepts; we conceptualize phenomenal properties *directly*. Loar explains:

Phenomenal concepts and theoretical expressions of physical properties both conceive their references essentially. But if two concepts conceive a given

property essentially, neither mediated by contingent modes of presentation, one ought to be able to see a priori—at least after optimal reflection—that they pick out the same property. Such concepts' connections cannot be a posteriori; that they pick out the same property would have to be transparent.<sup>1</sup>

But there are counterexamples to the semantic premise. As Perry (1979) observes, it is possible to know all the goings on in the grocery store without knowing that *I* am the one making a mess; to know the full sequence of the day's events without knowing that the meeting begins *now*; and to know every trail depicted on the map without knowing that *this* is the way out of the woods—and all of this despite an obvious sense in which the unknown indexical/demonstrative facts are “nothing over and above” the facts already known. Indexicals and demonstratives, it is plausible to say, pick out their referents directly, unmediated by any descriptive content; consequently, they cannot be known a priori to co-refer with any descriptive expressions.

Loar proposes that phenomenal concepts function similarly: they are *type demonstratives*. The referents of such concepts are determined by two elements: (a) a demonstrative component and (b) the *type* of thing in the causal presence of which a subject is disposed to deploy the concept. We might translate the content of such concepts into English as “yet another one of *those*,” or if we restrict the relevant referent-type to the subject's own phenomenal states, “yet another one of those internal states of mine.” Such concepts—if there are any—will refer directly to phenomenal states, express phenomenal kinds, and be inferentially isolated from any physical descriptions. Thus physicalists can grant epistemological dualism—there are no a priori connections between physical concepts and phenomenal concepts—while denying ontological dualism. Physicalism is necessarily an a posteriori truth.<sup>2</sup>

Loar's proposal has been enormously influential.<sup>3</sup> But it has also been subjected to persuasive criticisms. According to one objection, the type-demonstrative view implies, implausibly, that normal perceivers and blindsighted perceivers conceptualize their experiences identically.<sup>4</sup> According to another, it implies that

<sup>1</sup> Loar (1997, p. 600). It is controversial that such theoretical expressions of physical properties *do* conceive their referents essentially. If, for example, physical properties are the *realizer* properties of roles specified by our scientific theories, then we do not conceive physical properties essentially when we conceive them under scientific concepts. Philosophers who affirm that phenomenal concepts conceive phenomenal properties essentially, but who deny that scientific concepts conceive physical properties essentially, can insist that phenomenal-physical identifications would be a priori for us were we to grasp physical properties essentially. (Included in this camp are Russellian panpsychists, who maintain that dynamic, physical properties have qualitative, phenomenal intrinsic natures; and type-C physicalists, who maintain that our currently deficient scientific concepts obscure the a priori link between physical and phenomenal properties.) I shall set this option aside in what follows: the debate I am interested in adjudicating in this paper takes place among those who grant Loar's view that we are already able to conceptualize physical properties as they are essentially, e.g., because physical properties *just are* role properties (including the high-level role properties invoked by the special sciences).

<sup>2</sup> I leave open the strength of the modal. I suspect that most a posteriori physicalists have in mind something like *psychologically necessity for beings relevantly like us cognitively*. Some a posteriori physicalists might go in for a stronger or a weaker modal.

<sup>3</sup> Perry (2001), McLaughlin (2001), and Levin (2007) have all developed versions of it.

<sup>4</sup> See Horgan and Tienson (2001) and Demircioglu (2013). For responses, see Levin (2007) and Holman (2012). Loar (1997) himself acknowledges this objection, though his response to it has puzzled readers.

phenomenal concepts are “radically opaque,” revealing nothing of the nature of their referents.<sup>5</sup> The trouble seems to be this: on the type-demonstrative view, phenomenal concepts denote the items in their extension without any mediating reference-fixing properties; they *blindly* denote. We might say that demonstrative “concepts”—if it is appropriate to treat them as having any conceptual content at all—are “thin”: they denote without connoting. But phenomenal concepts are *thick*. To possess a phenomenal concept is to have a substantive grasp of the property connoted by that concept.<sup>6</sup> So, there are two ways that a concept can be non-wide, and thus “conceive its referent essentially”: it can denote without connoting (i.e., it can be thin, as type demonstratives are), or it can denote the very property it connotes. I will call concepts of the latter sort *narrow* concepts.

Phenomenal concepts are *narrow* concepts, then. And while Loar is right that the semantic premise is false (as his discussion of type demonstratives demonstrates), a weaker principle, *restricted to narrow concepts*, remains plausible. Variants of such a principle can be found throughout the literature, some explicitly framed in terms of phenomenal concepts and some framed more generally.<sup>7</sup> Here is the most general principle that captures what these proposals have in common:

**The Perspicuity Principle** For any suitably astute cognizer S and any properties F and G, if S possesses narrow concepts of properties F and G then S is able to know a priori whether F and G are identical.

Given that phenomenal concepts (along with physical concepts) are narrow, the perspicuity principle (or PP for short) entails that if phenomenal-physical identifications are not knowable a priori—a doctrine that a posteriori physicalists share with dualists—then physicalism is false.

### 3 The Perspicuity Principle in Dispute

Those philosophers who have used a version of PP to argue against a posteriori physicalism have not done much arguing for PP itself. Instead, they have merely invited readers to share their intuition that PP is true. For example, Horgan and Tienson (2001) assert that something like PP is “virtually tautologous”; its denial Goff (2011) declares “of dubious intelligibility” and Demircioglu (2013) pronounces “prima facie inconsistent.”

What is it about PP that these philosophers find intuitively compelling? (It is worth nothing that PP is no *formal* tautology; it can be denied without logical contradiction.) Their position, I suspect, is that the truth of PP falls out of the concept of a narrow concept. Here is the idea: narrow concepts essentially play a dual role in thought: (a) they *hook onto the world semantically*, thus determining truth values for particular cognitive states of which they are constituents; and (b) *they integrate with other concepts inferentially*, thus binding the subject’s cognizing

<sup>5</sup> See Goff (2011).

<sup>6</sup> For further criticism of the type-demonstrative view, see Papineau (2002), Tye (2009), Holman (2013), and Demircioglu (2013).

<sup>7</sup> Cf. Demircioglu (2013), Goff (2011), and Horgan and Tienson (2001).

into a rational whole. Something like this view lies behind Frege's conception of senses,<sup>8</sup> and most analytic philosophers since (at least those who countenance narrow content at all) have probably thought about narrow concept possession in much the same way, though the picture has rarely been made explicit.

But if there are other ways to think about the relationship between the two roles (semantic and inferential) that concepts play, it could turn out that it is not *essential* to narrow concepts that they play both roles perfectly well. This is an option that a posteriori physicalists have begun to explore explicitly. For example, Diaz-Leon has suggested that a subject S's knowledge of the following conditional is sufficient for possessing a narrow concept of pain: *If x is pain, then x is a mental state that feels like pain.*<sup>9</sup> But if that is all narrow concept possession comes to, then it is possible to possess a narrow concept of pain without knowing whether pain is a physical property or not. Similarly, Damnjanovic proposes that a subject possesses a narrow concept if she is in a position to know its counterfactual extension a priori.<sup>10</sup> The counterfactual extension of a phenomenal concept can be framed using phenomenal or physical vocabulary. So long as S knows whether a counterfactual scenario *described phenomenologically* satisfies a phenomenal concept, she possesses that phenomenal concept. She does not need to be able to tell whether or not the same counterfactual situation satisfies physicalistic vocabulary as well.

At issue are the epistemic implications of concept-possession. Those who endorse PP think that a subject possesses a narrow concept only if she can a priori discriminate its referent from (and co-identify its referent with) the referent of all the other narrow concepts she possesses. (Let us say that a narrow concept is "perspicuous" just in case it has this epistemic profile.) Diaz-Leon and Damnjanovic deny this: they think that knowledge of certain application-conditionals is sufficient for narrow concept-possession.<sup>11</sup> Other philosophers are likely to propose even weaker criteria: (1) a subject possesses a narrow concept if she knows certain disquotational propositions (e.g., '*pain*' means *pain*) and (2) a subject possesses a narrow concept if she reliably deploys

<sup>8</sup> Frege (1892, p. 152) writes: "A difference [in the cognitive value of two signs] can arise only if the difference between the signs corresponds to a difference in the mode of presentation." If a fully rational subject affirms that *the Evening Star is a planet with a shorter period of rotation than the earth* but fails to affirm that *the Morning Star is a planet with a shorter period of rotation than the earth*, this can only be because 'the Evening Star' and 'the Morning star' have different senses—which is to say that the subject fails to have a narrow concept of the Evening Star. Frege's principle amounts to the claim that: if  $x = y$ , but a fully rational subject does not know that  $x = y$ , then the subject fails to have narrow concepts of  $x$  and  $y$ . PP is a little bit stronger; it adds that if  $x \neq y$ , but a fully rational subject does not know that  $x \neq y$ , then the subject fails to have narrow concepts of  $x$  and  $y$ , and it specifies that such knowledge is a priori. All of this is in a recognizably Fregean key.

<sup>9</sup> Diaz-Leon (2014, 2016). Her view of phenomenal concepts is a bit difficult to categorize (as thin vs. narrow). While she thinks that there is descriptive content associated with phenomenal concepts, such descriptive content does not fix their reference; rather, a "purely automatic, recognitional disposition" does that, per the type-demonstrative view (Diaz-Leon 2016, p. 1195).

<sup>10</sup> Damnjanovic (2012, p. 88).

<sup>11</sup> Diaz-Leon and Damnjanovic offer similar proposals regarding the possession conditions of phenomenal concepts, but there are some key differences between their accounts, as well. For example, they offer their proposals at different stages of the dialectic. Diaz-Leon thinks that her proposal demonstrates the falsity of PP. Damnjanovic, on the other hand, thinks that PP poses a prima facie challenge to his proposal, but one that can be overcome if some version of the phenomenal concept strategy is successful—a matter to which I return in Sect. 6 below.

it in relevant contexts (without necessarily possessing any particular *propositional* knowledge).

In sum, if anti-physicalists are going to invoke PP against their opponents, they need to argue for it. I take up that dialectical burden in what follows. I begin by providing an argument for a weaker, existentially quantified version of PP. The purpose of this argument is to rule out the most general reason one might have for rejecting PP, viz., skepticism about a priori conceptual knowledge (or, more specifically, skepticism that the mere possession of a concept could afford a subject propositional knowledge a priori). After laying out this argument, I turn to my arguments for the stronger, universally quantified version.

#### 4 The Regress Argument for Perspicuity

Suppose an aspiring radio host (let's call her "DJ") hears audio engineers talking about the amplitude of a wave and about the frequency of a wave, and wants to know whether these are two terms for the same thing. How might DJ come to discover that amplitude is distinct from frequency? The most direct way would be to determine whether amplitude has features that frequency does not, or vice versa. To this end, DJ opens an audio engineering textbook and learns the following: amplitude admits of different *size*-magnitudes (and that is the only dimension along which it can vary); frequency admits of different *speed*-magnitudes (and that is the only dimension along which it can vary). Is she now in a position to learn that amplitude is distinct from frequency? Not unless she already knows that size and speed are different properties. Suppose it is *also* an open question for her whether size is speed. She opens an elementary physics textbook and learns that size admits of different magnitudes in spatial extension (and that is the only dimension along which it can vary); speed admits of different magnitudes in distance divided by time (and that is the only dimension along which it can vary). Is she now in a position to learn that size is distinct from speed, and hence that amplitude is distinct from frequency? Not unless she already knows that spatial extension and distance divided by time are different properties. And so on.

It is not difficult for most people to learn that amplitude and frequency are distinct. What is preventing DJ from doing so? The answer is that she does not grasp a conceptual distinction that is readily evident to most people, viz., that size and speed are distinct properties. The lesson is that learning new discriminatory knowledge requires that a subject already possesses some discriminatory knowledge. Given that we regularly *do* learn new discriminatory knowledge, then it follows, on pain of regress, that some discriminatory knowledge is epistemically basic.

One way that DJ could acquire the requisite basic discriminatory knowledge is by being told that (say) size is not speed by someone who knows. But every testimonial chain needs to be anchored in non-testimonial knowledge at some point. So *somebody* needs to have basic, non-testimonial discriminatory knowledge, even if DJ does not.

Another way DJ (or her informant) could acquire basic discriminatory knowledge is via perceptual experience. Certainly, it is often readily perceptually apparent that two properties are distinct. Suppose, for example, that you are watching a baseball team warming up before an inning. You see two pairs of players throwing baseballs back and forth, one pair using rapid throws, the second using softer throws. It is apparent that the

two baseballs are the same size but are traveling at different speeds. Obviously, a pair of objects cannot be the same and also different with respect to the same property, so size and speed are apparently distinct.

But it is only possible to determine that the baseballs are traveling at distinct speeds on the strength of prior discriminatory knowledge. Suppose you come by your belief that the balls are traveling at distinct speeds because they *look to you* to be traveling at distinct speeds: ball A looks one way, and ball B looks another way. But unless you know that these two looks are distinct—that is, unless you possess discriminatory knowledge with respect to the phenomenal “looks” of different speeds—you will not be able to judge that they are traveling at different speeds. Or suppose you come by your belief in a more rigorous way, by measuring each of their speeds. Again, unless you know that the two measurements are distinct—for example, that 56 miles-per-hour is not the same speed as 65 miles-per-hour—then you will not be able to judge that they are traveling at different speeds. Thus, even when perception affords knowledge of the distinctness of two properties, prior discriminatory knowledge is involved. It cannot be the case that *all* of our discriminatory knowledge is empirically founded. So at least some such discriminatory knowledge will need to be a priori.

A parallel line of argument applies to the acquisition of identificatory knowledge. Suppose a city boy visits his cousin in the country. They drive past a field filled with a dark green crop. Count, the country cousin, tells Sid, the city cousin, that that crop is the same stuff as the stuff Grandpa smokes in his pipe. Their conversation proceeds as follows:

Sid: No way!

Count: Both are tobacco, Sid.

Sid: I know they're both *called* “tobacco”. Mary Jane and dandelions are both called “weed,” but they're not the same.

Count: But you can't smoke dandelions. The leafy stuff is smokable, as is the dried stuff.

Sid: I know they're both *smokable*. “Smoking tobacco” in Grandpa's sense is lighting dried leaves in a pipe and puffing on it; “smoking tobacco” in the farmer's sense is some sort of harvesting method.

Count: Sid, I'm telling you, there's only one kind of tobacco-smokability: it involves jamming leaves in a pipe or rolling them up into a cigarette or a cigar.

Sid: Right, I know all about that: a pipe in the farmer's sense is some sort of mechanical harvester and “cigarettes” and “cigars” are ways of rolling up tobacco leaves.

What is Sid's problem? Where most people would be able to recognize that two terms, or two deployments of the same term, pick out the same property, Sid is at a loss. The lesson is that learning new identificatory knowledge requires that a subject already



possesses some identificatory knowledge. Given that we regularly learn new identificatory knowledge, then it follows, on pain of regress, that some identificatory knowledge is epistemically basic.

Here as above, the regress cannot terminate in testimony (I will not repeat why). Nor can it terminate in experience. Perception affords knowledge of property identities when one encounters instances of “two” properties, and those two instances share an overwhelming preponderance of features. But *recognizing* their shared features requires possessing identificatory knowledge already. Suppose you took samples of tobacco from a tobacconist and from a tobacco farmer, and studied the two samples to determine whether they are botanically identical. No matter how many features the two samples have in common, unless you know that some feature—combustion-odor, say—observed to be present in sample 1 is the same as a feature observed in sample 2, you are not in a position to learn that the samples are botanically identical. Thus, even when perception affords knowledge of the sameness of two properties, prior identificatory knowledge is involved. So at least some such identificatory knowledge will need to be a priori.

Thus, insofar as it is possible to acquire *any* discriminatory or identificatory knowledge about properties, some such knowledge must be a priori. And it is not possible to have a priori discriminatory/identificatory knowledge of properties conceived via *wide* concepts, because it is a live possibility that distinct wide concepts turn out to co-denote (as is the case for the concepts of *the morning star* and *the evening star*) or that a single wide concept denotes one property in one context and a distinct property in another (as is the case for the concept of *jade*).<sup>12</sup> Call this “the regress argument” for the perspicuity of some narrow concepts. We have established, in other words, that a weakened, existentially quantified version of PP is true: *some* of a subject’s narrow concepts are perspicuous.

Why should we think that *all* of a subject’s narrow concepts are perspicuous? Note that we have already overcome a major impediment to embracing PP just by establishing the weaker, existential claim. Once PP is explicitly stated, it might strike us as jarringly rationalistic—as requiring of us that we have magical powers of rational property graspings and property comparings. But if this is a reason to find PP objectionable, it is equally a reason to find the weaker, existential version objectionable. And we have found that the weaker, existential version is true. If *some* narrow properties are perspicuous, why should not we think that *all* of them are? At the very least, no general epistemological principle stands in the way of endorsing PP. Furthermore, there may be a route from the weakened version of PP to the stronger version (or at least a *ceteris paribus* version), by way of the *explanation* of the truth of the weaker version. For example, if the reason that some narrow concepts are perspicuous is that possessing narrow concepts affords a subject with rational insight into the nature of its

<sup>12</sup> Actually, it *is* sometimes possible to have a priori discriminatory/identificatory knowledge of properties conceived via wide concepts. For example, it may be a priori that the concept of *the shape of such-and-such a field* and the concept of *the color of the crop in such-and-such a field* denote distinct kinds. Such concepts are partially narrow: their connotations include essential properties of their referent kinds, *viz. being a shape* and *being a color*. Assuming that a subject can form narrow concepts of all the properties she grasps, the a priori distinction between the two wide concepts I just mentioned seems to be parasitic on the a priori distinction between two narrow concepts that the subject has or could have—the narrow concepts of *being a shape* and *being a color*.



referent, then we ought to expect that all narrow concepts afford rational insight of the same sort. But I shall not pursue that line of argument any further.

In the next section, I develop three arguments in favor of (full-strength) PP: an *inductive* argument, an *abductive* argument, and a *transcendental* argument. None is deductive (what non-question-begging argument for a universal generalization ever is?); nevertheless, together they amount to a strong case in favor of PP.

## 5 Three Non-deductive Arguments for the Perspicuity Principle

### 5.1 The Inductive Argument

According to the inductive argument, there are lots of uncontroversial cases of perspicuous narrow concepts, and no uncontroversial cases of non-perspicuous narrow concepts. So, probably, all narrow concepts are perspicuous.

In order to assess the inductive argument, we need a criterion for determining which of our concepts are narrow. Recall that narrow concepts are those concepts that connote the very property they denote; put a little less technically, narrow concepts are those that *reveal the essence of their referent*. Is there a way for a subject to tell, from the first-person perspective, whether a concept is narrow? Perhaps, by implementing the following steps. First, pick a concept (*dog*, say). Try to unpack its connotation (furry, four-legged, friendly, within such-and-such a size range, etc.).<sup>13</sup> Next, try to do the same with the constituents of its connotation. Eventually, you will get to those elements of the connotations of your concepts that are not analyzable. Two types of concepts have this feature. First, atomic *thin* concepts, such as type demonstratives, which pick out their denotations without connoting any properties; second, atomic *narrow* concepts, which connote the very properties they denote. If an unanalyzable part of a concept has no intrinsic content to it—if it is opaque to us, a mere mental placeholder—then we have probably landed on a thin concept. But not all the unanalyzable parts of our concepts are going to be like this; if they were, our mental states would not present us with any substantive content, from the first-person perspective.<sup>14</sup> Some of the unanalyzable parts of our concepts are going to be narrow concepts—viz., those that afford us a direct mental grip on their referents.

Now, performing such self-analysis is no small cognitive feat; there is no reason to expect that we will be very good at it. That does not mean that the procedure would be a waste of time. Still, it would be nice if we could supplement it with a third-personal method. In fact, there are at least three third-personal methods that, like the first-personal method just described, attempt to identify the unanalyzable components of our thoughts. First, we can investigate the *neural-functional primitives*: we can use the tools of cognitive neuroscience to discover the intentional contents that correspond to

<sup>13</sup> This is not conceptual analysis in the traditional sense. The concept of a dog is not a narrow concept, and thus unpacking its connotation does not reveal the essence of its referent. We do not learn what dogs are by unpacking the connotation of our concept—we learn, rather, about our means of denoting dogs.

<sup>14</sup> Schroer (2010) seems to deny this. Our color concepts, he suggests, are complex: they can be analyzed into hue, saturation, and brightness concepts, and these concepts are thin type-demonstratives. But it seems to me that a complex of thin concepts will inherit the “radical opacity” of its constituents. A concatenation of thin concepts is itself thin.

the earliest stages of perceptual processing that correlate with conscious perceptual states.<sup>15</sup> Second, we can investigate the *developmental primitives*: we can use the tools of cognitive psychology to discover which intentional contents show up very early in human development and very widely across human cultures.<sup>16</sup> Third, we can investigate the *cognitive-linguistic primitives*: we can use the tools of cognitive linguistics to determine which set of basic intentional contents best explain the peculiarities of human speech and behavior.<sup>17</sup> A conservative list of our most primitive thought constituents, compiled on the basis of convergence among these three methodological routes, includes the following: concepts of objecthood, agency, causation, motion, color, egocentric location, and some set of shape and size properties. Given that these concepts are not analyzable, and given that they are not thin (we have a positive, substantive grasp of their referents), they are narrow.

It is obvious, but no less interesting for being obvious, that we can discriminate and co-identify a priori these primitive intentional contents. First of all, we know a priori that all of these determinable properties are distinct from one another—causation is not the same as location, size is not the same as color, etc.—and we know that phenomenal properties are distinct from each of the other primitive intentional contents in our list. Second, we know a priori that simple determinates of these properties are distinct from one another—red and orange are distinct, moving rapidly and moving slowly are distinct, etc.—and we can likewise discriminate among determinate phenomenal properties. Third, we know a priori when complexes of these primitive intentional contents are the same or different, though a bit of reasoning may be required. For example, we know a priori that triangularity is not the same as hexagonality, and we know a priori that triangularity *is* the same as trilaterality (in the sense in which necessarily co-extensive properties are the same). Note, too, that we can tell a priori when complexes of narrow concepts have *overlapping* contents. For example, we can co-identify the feature in common between headaches and backaches or even between *intense* visual stimuli and *intense* auditory stimuli.

Are there any counterexamples to PP? One might argue that some of our narrow concepts are vague, and that we are not able always to tell whether pairs of vague concepts co-denote. I do not think this is quite right, however. Vague concepts have indeterminate extensions; we may not be able to tell whether pairs of vague concepts are co-extensive. But we will be able to tell whether they have the same vague connotation. For example, I may not be able to tell whether a reddish object is marginally crimson, marginally scarlet, both, or neither. But if I really have narrow concepts of the two colors (as opposed to *deferential* concepts of them, along the lines of *shade of red called 'crimson'/'scarlet' by color experts*), then I will know a priori that they pick out different, indeterminately bounded regions of the color-quality space.

In short, we have a robust ability to tell a priori whether our narrow concepts pick out the same contents or distinct contents. This ability manifests itself with respect to a wide range of content types, as well as with respect to both atomic and compound concepts. Further, there do not seem to be any uncontroversial examples of the breakdown of this ability. We thus have strong inductive evidence for PP. Probably,

<sup>15</sup> Cf., e.g., Marr (1982) and Martin (2007).

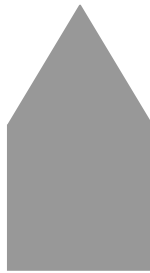
<sup>16</sup> Cf., e.g., Kinzler and Spelke (2007) and Carey (2009).

<sup>17</sup> Cf., e.g., Lakoff and Johnson (1980) and Langacker (1991).

if we do not know a priori that two of our narrow concepts co-refer, then they do not co-refer, and if we do not know a priori that two of our narrow concepts distinctly refer, then do not distinctly refer—because PP is true.

## 5.2 The Abductive Argument

We can imagine performing the following experiment, as a way of learning about our acquisition of narrow concepts. Suppose you introduced subjects to two new terms, “fisk” and “squib.” (Perhaps you tell your subjects that these are terms used in architecture.) You tell them that they are geometrical terms, but you do not tell them that the two terms are co-referential, as in fact, they are: they both pick out the same two-dimensional Euclidean shape—a square topped by an equilateral triangle whose sides are the same length as the sides of the square:



You introduce the concept of a fisk to subjects by presenting examples of varying sizes and colors and by presenting examples of other shapes and telling subjects that these are *not* fisks. You then do the same with the concept of a squib (which is of course the same concept), making sure to use different examples. Next, you test their concept acquisition by way of simple recognition tasks: you present an array of shapes, some of which are fisks/squibs, and some of which are not, and you ask subjects to pick out the fisks; then you show them a different array and ask the same question regarding squibs.

Suppose a subject performs well on the recognition tasks. You then ask the subject the following two questions: “Elena’s favorite shape is a fisk. Is Elena’s favorite shape a squib?” and “Are fisks the same as squibs or different?” Let us suppose that the subject answers these questions wrongly or does not know how to answer. What could explain this situation? Here are six possible explanations, in order of decreasing intrinsic plausibility:

- (1) The subject does not possess the concepts of a fisk and of a squib; instead, she possesses narrow concepts that are in the neighborhood—concepts of closely related determinates or determinables of fiskhood/squibhood (e.g., *squarishly bottomed, pointy thing*).
- (2) The subject does not possess the concepts of a fisk and of a squib, though she does possess a blind, non-conceptual recognitional capacity—a “sub-personal” mechanism, we might say, for picking out fisks/squibs.
- (3) The subject possesses adequate non-narrow concepts of a fisk and of a squib. Perhaps she has memorized the test arrays, for example, in which case she associates “fisk” with the description *shape found on array 1 in positions X, Y, and Z*. Perhaps she is reading cues from the experimentors, in which case she

- associates “squib” with the description *shape that, when pointed at, causes raised eyebrows in experimenter 1*. (Another way to put it: she is cheating.)
- (4) The subject possesses narrow concepts of a fisk and a squib but does not understand the questions being put to her. (Perhaps she lacks the concepts of a favorite thing and of sameness and difference. Or perhaps she cannot understand the question because she is certain the experimentors must mean to be asking something interesting rather than something stupid.)
  - (5) The subject possesses narrow concepts of a fisk and a squib but lacks the attentional capacities or the cognitive endurance to think through the question—i.e., she is not suitably astute.
  - (6) The subject possesses narrow concepts of a fisk and a squib and has all of the cognitive capacities necessary to understand the question but lacks the ability to tell that fisks are identical to squibs—i.e., PP is false.

Only as the subject’s success in the recognition task proves increasingly robust are we likely to appeal to explanations further down the list. If the subject initially succeeds at the recognition task but cannot answer our questions, our first thought is likely to be, per (1), “She hasn’t quite homed in on the concept yet.” As her success continues, we’ll think, per (2), “Huh—she can respond behaviorally to the distinction between fisks-and-squibs and other shapes, but she hasn’t consciously grasped what fisks/squibs are yet.” After a bit longer, we’ll think, per (3), “How can she not know the answer to our questions? She’s *got* to be cheating somehow.”

Only once we have ruled out cheating would we turn to (4) and (5), which imply that there is some breakdown in the subject’s ability to understand or respond to the question posed. The reason we would not initially reach for these explanations is that we have a sense for which cognitive capacities tend to bundle and which tend to dissociate, and it would be pretty strange for a subject to be able to acquire new shape concepts and to be able to deploy them in recognition tasks, yet not to be able to understand or focus on simple questions about how those concepts relate. Nevertheless, we know that cognitive capacities have been found to dissociate in surprising ways, so we could be persuaded that our subject exemplifies a heretofore undiscovered mix of capacity and deficit.

But it is hard to imagine what would push us to embrace (6). To embrace (6) is to maintain that the subject could be in possession of co-referential, narrow shape concepts, fully understand simple questions, be a fully astute cognizer, and yet have no idea—or persistently deny—that the concepts pick out the same shape. That sounds really implausible. And whereas each of the explanations on the list has a lower intrinsic plausibility than those above it, the difference in plausibility between (5) and (6) seems wider than the difference between the plausibility of earlier items.

The pressing question for our purposes is *why* we would be so reluctant to reach for (6). Notice that (6) is certainly not the most complicated of the explanations: (2) requires positing a blind recognitional capacity; (3) entails that the subject is engaging in an elaborate deception; (4) and (5) attribute peculiar cognitive deficits to the subject. (6) is rather elegant in comparison.

(6) entails the falsehood of PP. If PP were false, or if it were merely contingently true, we would not find (6) so implausible. The best explanation for why we find (6) so implausible is that the denial of PP is not an intelligible proposal, but rather a

conceptual impossibility. Discriminatory abilities just are constitutive of narrow-concept possession: suitably astute cognizers possess narrow concepts only of those properties they can “keep straight.” PP is a necessary truth after all.

### 5.3 The Transcendental Argument

The abductive argument is third personal: it turns on those conditions under which it would make sense to attribute the possession of a narrow concept. The transcendental argument, like the inductive argument, is first personal: it turns on the epistemic capacities we take ourselves to have. But it is more complicated than the inductive argument. The heart of the transcendental argument is this: if we believed that PP were false, the consequence would be a debilitating form of skepticism. Insofar as it is right for us not to allow ourselves to become epistemically debilitated, we ought to embrace PP.

The regress argument above turned on the observation that, in order to acquire new discriminatory knowledge, we need to have some discriminatory knowledge in the background already, and in order to acquire new identificatory knowledge, we need to have some identificatory knowledge in the background already. Recall Sid. Sid says he believes that the stuff in Grandpa’s pipe is identical to tobacco and that the stuff in the adjacent field is identical to tobacco. We would be tempted to formalize these belief as “ $A = F$ ” and “ $B = F$ .” But if that were an accurate representation of his beliefs, Sid would be in a position to infer that  $A = B$  (that the stuff in Grandpa’s pipe is identical to the stuff in the adjacent field). But he cannot do that, because he takes his two uses of “tobacco” as merely homophonic rather than synonymous. So we ought to formalize his two beliefs as “ $A = F$ ” and “ $B = G$ ,” beliefs that are inferentially isolated from one another. Let’s call Sid’s cognitive malady “cognitive fragmentation.”<sup>18</sup>

Cognitive fragmentation can vary in degree and scope. If Sid cannot co-identify *any* properties, then none of his predicates is ever synonymous with any others; it is as though every one of his judgments involves the predication of a unique property. We could call such an extreme case “complete fragmentation.” If Sid’s fragmentation is only *partial*, then he will be able to co-identify *some* properties. Anyone who possesses two co-denoting concepts but does not know that they co-denote suffers from some degree of fragmentation. (This is true, for example, of someone who does not know that mountain lions are cougars. Whenever such a person makes assertions about mountain lions, she is not thereby in a position to assert the same things about cougars; her set of beliefs about *Puma concolor* fragments into two inferentially isolated clusters.) What PP rules out is that the scope of fragmentation extends to *narrow* concepts, i.e., to the basic conceptual elements of an astute cognizer’s thoughts. Call the type of fragmentation ruled out by PP “basic

<sup>18</sup> Here, I finesse a difference between (a) possessing two co-denoting concepts and failing to know that they co-denote and (b) tokening the same concept twice and failing to know that the tokens are of the same type. I do this because, on some theories of concepts, concepts are individuated by their connotation (or by their conceptual role, if conceptual role determines connotation) such that it would be impossible to possess two distinct narrow concepts that co-denote. If your theory of concepts allows that it is possible to possess multiple concepts with the same connotation, then read what follows in accordance with (a). If not, read what follows in accordance with (b).

fragmentation” (in contrast with the more familiar *non-basic* kind, which frequently occurs in connection with wide concepts).<sup>19</sup>

Now, recall DJ. She believes that amplitude varies only with respect to size and that frequency varies only with respect to speed. We would be tempted to formalize these beliefs as “A is F” and “B is G,” and we would expect, further, that DJ believes that  $F \neq G$  (since size obviously is not speed). But if that were an accurate representation of her beliefs, DJ would be in a position to infer that A is not B (that amplitude is distinct from frequency). But she cannot do that, because she does not know that ‘size’ and ‘speed’ are *not* synonymous; for all she knows, they co-denote. We could better capture her epistemic situation if we formalized her two beliefs as sharing a disjunctive predicate, viz., “A is F-or-G” and “B is F-or-G.” Because she cannot tell size from speed, her two beliefs have the same inferential role. Let’s call DJ’s cognitive malady “cognitive collapse.”

If DJ cannot discriminate between *any* properties, all of her property attributions have the same inferential power; it is as though she is just predicating the property of *having some property or other* over and over again. We can call such a state “complete collapse,” contrasted with *partial* collapse. Anyone who possesses two distinctly denoting concepts but does not know that they distinctly denote suffers from at least partial collapse. (This is true, for example, of someone who does not know that elms are not beeches. Whenever such a person makes assertions about elms, the proposition she expresses is really about elms or beeches; her elm beliefs and her beech beliefs collapse into one another.) But again, PP rules out the possibility of *basic* collapse—the inability, on the part of an astute cognizer, to tell when the basic conceptual elements of her thoughts are different (in contrast with the more familiar *non-basic* kind).<sup>20</sup>

Armed with these distinctions, I now present the transcendental argument in brief, and then proceed to defend its central premises:

- (1) If I have a defeater for all of my basic property identifications or discriminations, I become epistemically paralyzed.
- (2) If I believe I am subject to partial basic collapse or fragmentation, I have a defeater for all of my basic property identifications or discriminations.
- (3) I should not let myself become epistemically paralyzed.
- (4) So, I should not believe I am subject to partial basic collapse or fragmentation.
- (5) So, I am not subject to partial basic collapse or fragmentation.

The heart of the argument is found in the first two premises.

<sup>19</sup> A minor case of cognitive fragmentation can be found in Kripke’s subject who knows that Paderewski is a pianist and that Paderewski is a statesman but is unable to infer, from an empirical belief that Paderewski the pianist is present, that a statesman is present, owing to the fact that he does not know that Paderewski is Paderewski (see Kripke 1979).

<sup>20</sup> A few years ago, the satirical newspaper *The Onion* published a ridiculous article on the financial recovery from the perspective of a writer who lacks the ability to discriminate among various financial notions. The comical confusion of the piece illustrates how debilitating cognitive collapse in a domain can be. The article can be found at <http://www.theonion.com/articles/something-about-tax-cuts-or-earnings-or-money-or-s,18169/>.

*Premise 1: If I have a defeater for all of my basic property identifications or discriminations, I become epistemically paralyzed.* Suppose you came to believe that, for any intuitive judgment you're inclined to make to the effect that properties P1 and P2 are different, you are as likely to be wrong about it as right about it—even when it comes to those properties you seem to “grasp”, to conceive essentially. You believe you are becoming like Sid. How bad would this be for you? Well, it would be at least as bad for you as it is for Sid, viz., it would prevent you from being able to infer, from claims about the sharing of properties by two items A and B, that those two items are identical. If you found yourself thinking, “A has P1, P2, and P3, and B has P1, P2, and P3—properties not likely to be had in common by two distinct things—so probably A is B,” you would be psychologically prevented from making the inference. Instead, you would be forced to think, “For all I know, ‘P1’ is a name for two different properties, and likewise for ‘P2’ and ‘P3’. For all I know, I would be committing the fallacy of equivocation by making the inference. So I had better withhold judgment.” You would find yourself epistemically paralyzed.

Now suppose you came to believe that, for any intuitive judgment you are inclined to make to the effect that properties P1 and P2 are *identical*, you are as likely to be wrong about it as right about it: you believe you are becoming like DJ. If you found yourself poised to infer that A is distinct from B on the grounds that P1 and P2 are properties located in the same quality-space,<sup>21</sup> and A has P1 whereas B has P2, you would be psychologically prevented from completing the inference. Instead, you would be forced to think, “For all I know, ‘P1’ and ‘P2’ name the same region of quality space. So I had better withhold judgment.”

But your epistemic paralysis would not be restricted to Sid- and DJ-style inferences. Paralysis would set in *wherever the same property is connoted in distinct steps in an inference process*. Consider a much more quotidian example. Suppose my wife asks me to retrieve the lilac dress from her closet, and suppose further that she owns both a lilac dress and a violet dress. My ability to retrieve the right dress is a matter of my being able to form accurate perceptual beliefs about the colors of the dresses—*that dress is lilac* and *that dress is violet*—and then to hook those beliefs up inferentially with my prior belief that *the dress to be retrieved is lilac*. Armed with these beliefs, I am able to infer that *\*that\** [perceptually demonstrating a dress] *is the dress to be retrieved* and that *\*that\** [perceptually demonstrating another] *is not the dress that is to be retrieved*, as follows:

- (1) *If the dress is P1, then it is the dress that is to be retrieved.*
- (2) *X is P2.*
- (3) *Y is P3.*
- (4) *Hence, X is the dress that is to be retrieved.*
- (5) *Hence, Y is not the dress that is to be retrieved.*

(X and Y are stand-ins for perceptual demonstratives. P1 is lilac; P2 is lilac; P3 is violet.) It is obvious that (4) and (5) do not follow *formally* from (1) to (3). I can only

<sup>21</sup> The significance of the point about a shared quality-space is this: if P1 and P2 are located in the same quality space and P1 is distinct from P2, then P1 and P2 are *incompatible* properties: a thing cannot have both (in the same way, at the same time, etc.).



make these inferences on the strength of my ability to discriminate lilac from violet and my ability to identify lilac with lilac—to know that P1 is identical to P2 and P1 is distinct from P3.<sup>22</sup> But if I believe that my capacities to discriminate and co-identify properties are unreliable, I will withhold judgment about the identity or non-identity of P1, P2, and P3. I will be prevented from making the most common and obvious of inferences.

*Premise 2: If I believe I am subject to partial basic collapse or fragmentation, I have a defeater for all of my basic property identifications or discriminations.* What could possibly serve as a defeater for my basic discriminatory and identificatory knowledge? My claim is that the belief that I am subject to partial basic collapse or fragmentation would serve as just such a defeater. Here is the idea. When it comes to basic discriminations and identifications (judgments of sameness and difference among my narrow concepts), my judgements ultimately rest on their *seeming obviousness*. (Not that there is a further claim—*this judgment seems obvious*—that justifies them, but rather that such judgments are *self-evident* if any judgments are.) But once this epistemic grounding is called into question—once I am told that I am just brutally wrong about the very most obvious judgments I am apt to make—I can no longer trust any of the judgments I make that are of the same type, in the absence of a principled reason to think that fragmentation or collapse is *quarantined* (a proposal to which I return in the next section).

It may help to consider an analogous scenario. Suppose you were told that Disjunctive Syllogism is, after all, an invalid argument form, a form you are hard-wired to accept as valid. Your first reaction, I suspect, would be bemused curiosity—which is to say, you would not take the idea seriously. If you began to take the idea seriously, you would request an explanation in terms of more basic logical concepts that allows you to understand how you could make the error. But if you came to believe that (a) you suffer from this cognitive defect, and (b) the defect is located *at the level of your most basic logical concepts*, your natural reaction would be to lower your confidence in your grasp of *any* logical concepts. And at that point you would find yourself undergoing a sort of cognitive vertigo, lacking the degree of confidence in your own faculties required to make conscious inferences. Analogously: the allegation that one is subject to partial basic collapse or fragmentation is not at all the familiar point that we are sub-optimal cognizers, that we employ heuristics and biases and are subject to illusions, etc. No: this is the claim that we are doomed to err with respect to the most central, general, common, obvious judgments in our cognitive lives, and that there is no cognitive vantage-point from which to observe the mechanics of our folly. Such is a damning indictment of our cognitive capacities. It is not clear to me that a rational cognizer could proceed with her inquiries after taking the indictment to heart.<sup>23</sup>

In sum: if a subject denies PP, and comes to believe instead that she is subject to (even partial) basic collapse or fragmentation, she will find herself in possession of a defeater for her basic discriminatory/identificatory capacities generally, and the

<sup>22</sup> I also have to know that if anything is P1 then it is not P3. Such knowledge seems to come along with the possession of color concepts, as it does with the possession of any two concepts that connote properties in a shared quality space. See the previous footnote.

<sup>23</sup> An objection: “A posteriori physicalists believe themselves to be subject to basic cognitive fragmentation, yet they are not epistemically paralyzed in the way you describe.” Reply: nor do Humean skeptics about nomic necessity fail to plan for the future. I am arguing that a posteriori physicalism has unacceptable skeptical implications, whether or not a posteriori physicalists live accordingly.

consequence will be a debilitating form of skepticism. Endorsing PP allows us to avoid such epistemic paralysis, and that is a very good reason to endorse it.<sup>24</sup>

## 6 Restricted Perspicuity: the Phenomenal Concept Strategy

According to the inductive argument, a mighty host of narrow concepts are perspicuous, so probably all are. According to the abductive argument, we are very hesitant to attribute inferential ineptitudes to the non-perspicuity of the relevant narrow concepts; the best explanation of our hesitation is that it is nonsensical to treat narrow concepts as non-perspicuous. According to the transcendental argument, if we believed that some of our narrow concepts were non-perspicuous (in other words, that we suffer from basic cognitive fragmentation or collapse), we would find ourselves epistemically paralyzed; but we should avoid such paralysis, so we should believe that all narrow concepts are perspicuous: so probably all narrow concepts are perspicuous.

The a posteriori physicalist maintains, of course, that phenomenal concepts are *not* perspicuous. She can maintain this while granting that the three arguments have some force, simply by asserting that phenomenal concepts are peculiar among narrow concepts. She can then resist each argument by locating the step at which this possibility (that phenomenal concepts are inferentially anomalous) is ruled out. Contra the inductive argument, we cannot generalize from normally behaved narrow concepts to the whole lot, since some narrow concepts do not behave normally. Contra the abductive argument, even if we would never attribute inferential ineptitude to the non-perspicuity of normally-behaved narrow concepts, we *should* attribute it to the non-perspicuity of phenomenal concepts, when phenomenal concepts are involved. Contra the transcendental argument, so long as basic fragmentation is “quarantined” to a specific subset of our narrow concepts, our confidence in our general cognitive capacities remains undefeated.

I grant that such a strategy for resisting the three arguments is available to the a posteriori physicalist; I deny that it should be attractive to the a posteriori physicalist. It involves the ascription of anomalous inferential profiles to narrow concepts of a specific type, with no grounds for so ascribing other than that a posteriori physicalism requires it.<sup>25</sup> In order to avoid making an ad hoc move, the a posteriori physicalist needs to give us an independent reason to think that phenomenal concepts are not perspicuous, despite the fact that most (all other?) narrow concepts are. In other words, we need a reason to believe that the basic fragmentation we suffer from is quarantined to the domain of the phenomenal.

This is one way of understanding what Loar (1997) is up to. That is, he provides a theory of phenomenal concepts such that if the theory is true, then psychophysical identifications will not be a priori for us, whether or not dualism is true. As we saw above, the details of Loar’s proposal turn out to be untenable. But the general

<sup>24</sup> I suspect Goff (2015, p. 139) has something similar in mind when he says that “the idea that there could be brute necessary connections even between descriptions which render transparent real natures is to suppose that the world as it is in and of itself is in certain respects unintelligible.”

<sup>25</sup> Let us keep in mind that physicalism does not itself entail that phenomenal concepts are non-perspicuous. Physicalism entails that the a priori discriminations we are inclined to make between phenomenal and physical properties are not all true. A posteriori physicalists add that we have no a priori method for correcting our mistake.

strategy—dubbed by Stoljar (2006) the “Phenomenal Concept Strategy” (or PCS)—is not vitiated by one failed attempt to implement it.<sup>26</sup>

It has turned out that the PCS is very hard to implement. Some proposals have failed because they make implausible predictions about our conceptual capacities. As we saw above, Loar’s version fails because it gets the phenomenology of the deployment of phenomenal concepts wrong: certainly there is a difference between a blind introspective recognitional capacity and a phenomenologically *guided* recognitional capacity. According to another proposal in the literature,<sup>27</sup> phenomenal concepts are lexical primitives in the language of thought. Just as two distinct logically atomic expressions are never synonymous with each other (in the sense that their formal properties do not reveal whether they co-refer), so no two lexical primitives in the language of thought are synonymous with each other. The trouble with this proposal is that if it is sufficient to explain the non-perspicuity of phenomenal concepts, it will also predict that *all* narrow concepts are non-perspicuous, a possibility ruled out by the Regress Argument above. (Further, it predicts that phenomenal concepts will be non-perspicuous *with respect to each other*. But of course we have no trouble determining a priori whether our concept of a tactile sensation and our concept of an auditory sensation co-denote.)

Other proposals do a better job saving the phenomenological and epistemological appearances, but these proposals have failed to imply that phenomenal concepts are not perspicuous. For example, a number of philosophers have observed that phenomenal concepts are canonically deployed introspectively, and no other concepts have this feature.<sup>28</sup> Writes Christopher Hill: “when the use of one concept is controlled by experiences that are altogether disjoint from the experiences that control the use of a second concept, it is in general impossible for us to determine, simply in virtue of having mastered the use of the concepts, whether they are co-referential.”<sup>29</sup> We can begin to see the explanatory inadequacy of this approach by noting that phenomenal concepts themselves are justifiedly deployed *both* in introspective contexts *and* in perceptual contexts, viz. when we attribute phenomenal states to other people on the basis of their observable behavior. Despite the fact that the two types of use are “controlled by experiences that are altogether disjoint”—my feeling of pain vs. my perceptual experience of your wince, say—we have no temptation to think that we must be attributing distinct properties in each case. The canonical context of deployment of a concept appears to be orthogonal to whether it is perspicuous.

Perhaps the most carefully-developed proposal is the “constitution-view” of phenomenal concepts, according to which phenomenal concepts are partly constituted by instances of (or by imaginative facsimiles of) the very properties they connote.<sup>30</sup> According to Katalin Balog’s version of the view, we are *acquainted* with phenomenal properties in virtue of their instances’ being embedded in our concepts of them; “this

<sup>26</sup> The PCS is really a family of strategies aimed at subtly different objectives. It has often been construed as a strategy for showing that psychophysical identifications inevitably *seem false* (whether or not they are). Showing that phenomenal concepts are non-perspicuous is a weaker goal, since phenomenal-physical identifications can fail to be a priori without seeming false.

<sup>27</sup> Lycan (1990), Tye (1999), and Papineau (2002) appear to have sympathies with this view.

<sup>28</sup> Sturgeon (1994), Hill (1997), and Hill and McLaughlin (1999).

<sup>29</sup> Hill (1997, p. 81).

<sup>30</sup> Advocates include Balog (2012), Papineau (2002, 2007), and Block (2007). This model has the virtue of theoretical neutrality, endorsed as it is by anti-physicalists such as Chalmers (2003).

kind of direct insight...does not reveal anything about the metaphysical nature of phenomenality.”<sup>31</sup> According to Papineau’s version, because phenomenal properties *use* an instance of the property they *mention* (as a result of their being partly constituted by the instances of the properties they connote), their deployment is always accompanied by an experience (or an imaginative copy of an experience), whereas co-referring non-phenomenal concepts need not be accompanied by an experience: “We can think about our conscious experiences using concepts to which they bear a phenomenal resemblance. And this then creates the fallacious impression that other, material ways of thinking about those experiences fail to refer to the felt experiences themselves.”<sup>32</sup>

Notwithstanding their insistence to the contrary, Balog and Papineau simply have not shown us why the partial constitution of phenomenal concepts by phenomenal properties entails that phenomenal concepts are non-perspicuous. Depending on what Balog means by “acquaintance”—depending, for example, on whether acquaintance - ascriptions are transparent or opaque—her view might very well entail that phenomenal concepts *are* perspicuous.<sup>33</sup> Papineau’s characterization of phenomenal concepts appears consistent with non-perspicuity but does not *entail* non-perspicuity. That one of two co-denoting concepts resembles its referent does not mean that a subject will not be able to tell a priori that the concepts co-denote. Suppose, for example, that I have two triangle concepts: one is an appearance-based concept that is partly constituted by an imaginative facsimile of the visual appearance of a triangle; the other consists in a geometrical description. This difference between the two concepts does not mean that I am unable to determine a priori whether they co-denote.

So far as I can tell, none of the versions of the PCS developed over the last 20 years has succeeded in giving us a reason, independent of any prior commitment to physicalism, to think that phenomenal concepts are epistemically anomalous among narrow concepts. Ought we to attribute such repeated failure to the impossibility of the task, rather than to the limitations of the particular proposals? I think we should.<sup>34</sup> The PCS is an attempt to explain our cognitive fragmentation without providing any remedy for it. But it is probably impossible for a subject to understand such an explanation while still suffering from the very cognitive deficit the explanation purports to explain. Consider a parallel case. Kahnemann and Tversky (1982) have shown that subjects judge “Linda is a bank teller and a feminist” as more probable than “Linda is a bank teller,” when told facts about Linda that fit their stereotype of a feminist—despite the fact that conjunctions are never more probable than their individual conjuncts. With a little bit of redescription, we can understand the case as an instance of cognitive collapse: viz., as a conflation of the property *resembling an F-stereotype* with the property *probably being an F*. Now, subjects have no trouble understanding how their error can be explained in terms of a subconscious heuristic—but only if they have some

<sup>31</sup> Balog (2012a, p. 31).

<sup>32</sup> Papineau (2002, p. 171), Tye (1999), and Block (2007) tell similar stories, though Tye points to the cognitive effects of phenomenal concepts (they tend to generate concomitant images when deployed) rather than their constitutive structure. Papineau himself seems not to believe that concepts have connotations, or narrow intentional features of any kind. My purpose here is not to evaluate Papineau’s view from within his own philosophical framework, but to determine whether it can help the a posteriori physicalist provide a principled restriction of PP.

<sup>33</sup> See Goff (2015) for a helpful discussion of notions of acquaintance vis-a-vis the PCS.

<sup>34</sup> Chalmers (2007) also argues that the PCS faces in-principle obstacles.

other way of teasing apart the concepts of probability and stereotype-resemblance. For suppose that a subject were really unable to see that *resembling an F-stereotype* and *probably being an F* were distinct properties. When a researcher tells her, “You mistakenly conflate these two concepts, because you have a subconscious disposition to deploy the heuristic, *determine probability of being an F by determining degree of resemblance to an F-stereotype*,” her understanding of the heuristic will be equivalent to, *determine probability of being an F by determining probability of being an F*. Hence, the subject can only understand the researcher’s explanation for her conceptual error if she can be shown independently how it could be an error. And this will not be possible unless she can see a priori that resembling an F-stereotype is distinct from probably being an F. So the subject is only in a position to understand an explanation of her cognitive collapse if she can overcome it.

If the case of cognitive fragmentation is analogous, then so long as we are in a state of cognitive fragmentation with respect to our phenomenal concepts, we cannot understand an explanation of that fragmentation. A purported psychological explanation of the non-perspicuity of phenomenal concepts will take the form: *if phenomenal concepts have feature X, phenomenal concepts are not perspicuous, and hence they are cognitively isolated from physical concepts*. What does it take for me to understand this purported explanation? Presumably, I deploy my own concepts *phenomenal concept* and *physical concept*, or perhaps I use specific phenomenal and physical concepts as surrogates (perhaps, that is, I have narrow concepts only of determinate phenomenal and physical properties, not of their determinables). That is, I bring to mind the phenomenal and physical properties that serve as the connotations of these concepts. And once I do so, it seems obvious to me that these properties are distinct (even if I believe that appearances are misleading and physicalism is true after all). That is, in the very act of trying to understand how feature X explains my intuitive judgment that phenomenal properties are distinct from physical properties, I am presented with a *rival* explanation of that judgment: their apparently obvious distinctness. But unless I can set aside that intuitive judgment and see how physical and phenomenal properties could be identical, I will not be able to see how feature X has its alleged effect.

To see this, let us return to an analogy used in the previous section: suppose you were offered a psychological explanation for your intuitive judgment that disjunctive syllogisms are valid, in terms of some feature X of your concepts of the logical operators. Understanding such an explanation would require that you understand the following conditional: *were disjunctive syllogisms invalid, feature X would cause them to seem valid*. I cannot imagine what it would be like for anything other than the obvious validity of disjunctive syllogism to be responsible for my persistent judgment that it is valid, and because I cannot imagine what it would be like for disjunctive syllogism to be *invalid*, I cannot imagine the conditions necessary for feature X to kick in. Likewise, I am only able to understand a proposal made by an advocate of the PCS if I am able to comprehend the subjunctive conditional, *if physical and phenomenal properties were identical, feature X of phenomenal concepts would cause them to seem distinct*. But for all that my deficient cognitive faculties are able to grasp, the antecedent of

that conditional is a counterpossible.<sup>35</sup> Feature X is, as it were, a failsafe: if phenomenal properties turn out to be physical properties, it kicks in to prevent my being able to see that they are. But I cannot imagine the obtaining of the circumstances under which it *would* kick in, so I cannot decompose the relevant psychological mechanisms so as to understand what its kicking in would look like—even if my judgment is in fact false and the proposed psychological explanation is in fact correct.

This is dialectical bad news for the a posteriori physicalist, because it means that the PCS cannot do what it was invoked to do. Because we will not be able to understand how feature X renders phenomenal concepts non-perspicuous, the PCS cannot supply us with an independent reason to think that phenomenal concepts are anomalous among narrow concepts. The best the advocate of the PCS can offer is the bare possibility that some feature X renders phenomenal concepts non-perspicuous. But so saying is no less ad hoc than the mere insistence that phenomenal concepts are non-perspicuous.

## 7 Conclusion

Loar observed that influential anti-physicalist arguments make a leap from conceptual dualism to metaphysical dualism on the strength of a controversial principle he called “the semantic premise.” He was right that the semantic premise is false, but wrong that dualists must rest their argument on it. Rather, they ought to rest their argument on the perspicuity principle (in connection with the claim that phenomenal concepts are narrow). Identifying this principle as the current locus of controversy marks real progress in the contemporary debate between physicalists and dualists.

It is common among those sympathetic to a posteriori physicalism to have an attitude of skepticism toward PP and of optimism toward the PCS. PP, they will say, is a relic of a more naively rationalistic era, twice over: it is a rationalistic principle defended rationalistically (if defended at all). Perhaps something approximating PP is true, but if so, what makes it true are the contingent facts about our cognitive architecture. And anyway, the PCS has shown, or promises to show, that phenomenal concepts occupy a peculiar corner of our cognitive architecture, so we should not expect them to behave the way other narrow concepts do. Thus, we should put very little stock in our a priori discriminatory judgments about physical and phenomenal properties.

The purpose of this paper has been to argue that such an assessment of the situation cannot be sustained. The case for PP is quite strong and does not rest on a priori conceptual considerations. Further, not only have versions of the PCS repeatedly failed; likely, all versions will, given the strange combination of lucidity and opacity that the PCS has to juggle (it requires that we come up with a lucid

<sup>35</sup> This is not to say that I cannot come to believe that apparent counterpossibles are true. It is just to say that I will not be able to understand why I am inclined to deem them counterpossibilities until I am able to see how they are possible after all. For example, I could come to believe that light is both a particle and a wave, despite the apparent non-sensicality of the belief; but I am not going to understand why I am prone to wrongly judge particlehood and wavehood incompatible properties unless I can first see how they could be compatible.



explanation of an irremediable cognitive blindspot). In short, a posteriori physicalists currently lack a principled objection to anti-physicalist arguments of the sort analyzed by Loar.

### Compliance with Ethical Standards

**Conflict of Interest** The authors declare that they have no conflict of interest.

## References

- Balog, K. (2012). Acquaintance and the mind-body problem. In S. Gozzano & C. S. Hill (Eds.), *New perspectives on type identity: The mental and the physical* (pp. 16–42). New York: Cambridge UP.
- 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). New York: Oxford UP.
- Carey, S. (2009). *The origin of concepts*. New York: Oxford UP.
- Chalmers, D. (2003). The content and epistemology of phenomenal belief. In Q. Smith & A. Jokić (Eds.), *Consciousness: New philosophical perspectives* (pp. 220–272). New York: Oxford UP.
- Chalmers, D. (2007). Phenomenal concepts and the explanatory gap. In T. Alter & S. Walter (Eds.), *Phenomenal concepts and phenomenal knowledge* (pp. 167–194). New York: Oxford UP.
- Damjanovic, N. (2012). Revelation and physicalism. *Dialectica*, 66, 69–91.
- Demircioglu, E. (2013). Physicalism and phenomenal concepts. *Philosophical Studies*, 165(1), 257–277.
- Diaz-Leon, E. (2014). Do a posteriori physicalists get our phenomenal concepts wrong? *Ratio*, 27, 1–16.
- Diaz-Leon, E. (2016). Phenomenal concepts: neither circular nor opaque. *Philosophical Psychology*, 29, 1186–1199.
- Frege, G. (1892/1997). "On *Sinn* and *Bedeutung*." In M. Beaney (ed.), *The Frege Reader*. Blackwell. Trans. M. Black.
- Goff, P. (2011). A posteriori physicalists get our phenomenal concepts wrong. *Australasian Journal of Philosophy*, 89, 191–209.
- Goff, P. (2015). Real acquaintance and physicalism. In P. Coates & S. Coleman (Eds.), *Phenomenal qualities: sense, perception, and consciousness* (pp. 121–146). Oxford: Oxford UP.
- Hill, C. (1997). Imaginability, conceivability, possibility, and the mind-body problem. *Philosophical Studies*, 87, 61–85.
- Hill, C., & McLaughlin, B. (1999). There are fewer things in reality than are dreamt of in Chalmers' philosophy. *Philosophy and Phenomenological Research*, 59, 445–454.
- Holman, E. (2013). Phenomenal concepts as bare recognitional concepts: harder to debunk than you thought, ...but still possible. *Philosophical Studies*, 164(3), 807–827.
- Horgan, T., & Tienson, J. (2001). Deconstructing new wave materialism. In C. Gillett & B. Loewer (Eds.), *Physicalism and its discontents* (pp. 307–318). Cambridge: Cambridge UP.
- Kahneman, D., & Tversky, A. (1982). Judgments of and by representativeness. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), *Judgment under uncertainty: Heuristics and biases*. Cambridge: Cambridge UP.
- Kinzler, K., & Spelke, E. (2007). Core systems in human cognition. *Progress in Brain Research*, 164, 257–264.
- Kripke, S. (1979). A puzzle about belief. In A. Margalit (Ed.), *Meaning and use* (pp. 239–283). Dordrecht: Reidel.
- Lakoff, G., & Johnson, M. (1980). *Metaphors we live by*. University of Chicago Press.
- Langacker, R. (1991). *Foundations of cognitive grammar, descriptive applications* (Vol. 2). Stanford University Press.
- Levin, J. (2007). What is a phenomenal concept? In T. Alter & S. Walter (Eds.), *Phenomenal concepts and phenomenal knowledge* (pp. 87–110). New York: Oxford UP.
- Loar, B. (1997). Phenomenal states. In O. Flanagan, N. Block, & G. Guzeldere (Eds.), *The nature of consciousness*. MIT Press.
- Lycan, W. (1990). What is the 'subjectivity' of the mental? *Philosophical Perspectives*, 11, 229–238.



- Marr, David. (1982). *Vision*. Freeman.
- Martin, A. (2007). The representation of object concepts in the brain. *Annual Review of Psychology*, 58, 25–45.
- McLaughlin, B. (2001). In defense of new wave materialism: a response to Horgan and Tienson. In C. Gillett & B. M. Loewer (Eds.), *Physicalism and its discontents*. Cambridge: Cambridge UP.
- Papineau, D. (2002). *Thinking about consciousness*. New York: Oxford UP.
- Papineau, D. (2007). Phenomenal and perceptual concepts. In T. Alter & S. Walter (Eds.), *Phenomenal concepts and phenomenal knowledge* (pp. 111–144). New York: Oxford UP.
- Perry, J. (1979). The problem of the essential indexical. *Noûs*, 13, 3–21.
- Perry, J. (2001). *Knowledge, possibility and consciousness*. Cambridge: MIT Press.
- Schroer, R. (2010). Where's the beef? Phenomenal concepts as both demonstrative and substantial. *Australasian Journal of Philosophy*, 88, 505–522.
- Stoljar, D. (2006). *Ignorance and imagination: the epistemic origin of the problem of consciousness*. Oxford: Oxford University Press.
- Sturgeon, S. (1994). The epistemic basis of subjectivity. *Journal of Philosophy*, 91, 221–235.
- Tye, M. (1999). Phenomenal consciousness: the explanatory gap as a cognitive illusion. *Mind*, 108, 705–725.
- Tye, M. (2009). *Consciousness revisited: materialism without phenomenal concepts*. Cambridge: MIT Press.