



Phenomenal States

Brian Loar

On a natural view of ourselves, we introspectively discriminate our own experiences and thereby form conceptions of their qualities, both salient and subtle. These discriminations are of various degrees of generality, from small differences in tactual and color experience to broad differences of sensory modality, for example, those among smell, hearing, and pain. What we apparently discern are ways experiences differ and resemble each other with respect to *what it is like to have them*. Following common usage, I will call these experiential resemblances *phenomenal qualities*, and the conceptions we have of them, *phenomenal concepts*. Phenomenal concepts are formed “from one’s own case.” They are *type-demonstratives* that derive their reference from a first-person perspective: ‘that type of sensation,’ ‘that feature of visual experience.’ And so third-person ascriptions of phenomenal qualities are projective ascriptions of what one has grasped in one’s own case: ‘she has an experience of that type.’

‘Phenomenal quality’ can have a different sense, namely, how the *object* of a perceptual experience appears. In this sense, a phenomenal quality is ascribed to an object and not directly to an experience. Some have argued that all we discern phenomenologically are phenomenal qualities in this sense; they deny that experiences themselves have introspectible qualities that are not ascribed primarily to their objects (Harman 1990; Block 1990). I will not pursue the issue here, but will assume a certain view of it. For the present objective is to engage antiphysicalist arguments and entrenched intuitions to the effect that conscious mental qualities cannot be identical with ordinary physical properties, or at least that it is problematic to suppose that they are so. Antiphysicalists typically suppose that such mental properties are not relational—that is, that they present themselves as not intrinsically involving relations to things outside the mind. They may allow that, say, visual experiences are in some sense intrinsically representational.

That is hard to deny because, as regards ordinary visual experiences, we cannot apparently conceive them phenomenally in a way that abstracts from their *purporting* to represent things in a certain way. The antiphysicalist intuition is compatible with visual experiences’ having (some sort of) internally determined intentional structure, so that it is an introspectable and nonrelational feature of a visual experience that it represents things visually as being thus and so. Antiphysicalists suppose that we have conceptions of how visual experiences differ and resemble each other with respect to what it is like to have those experiences. These conceptions then are of qualities of experiences, whatever allowances one may also make for the apparent qualities of the intrinsic objects of those experiences. I will assume that the antiphysicalists’ phenomenological and internalist intuitions are correct. The idea is to engage them over the central point, that is, whether those aspects of the mental that we both count as phenomenologically compelling raise substantive difficulties for the thesis that phenomenal qualities (thus understood) are physical properties of the brain that lie within the scope of current science.

We have to distinguish between *concepts* and *properties*, and this chapter turns on that distinction. Antiphysicalist arguments and intuitions take off from a sound intuition about concepts. Phenomenal concepts are conceptually irreducible in this sense: they neither a priori imply, nor are implied by, physical-functional concepts. Although that is denied by analytical functionalists (Levin 1983, 1986), many other physicalists, including me, find it intuitively appealing. The antiphysicalist takes this conceptual intuition a good deal further, to the conclusion that phenomenal qualities are themselves irreducible, are not physical-functional properties, at least not of the ordinary sort. The upshot is a range of antireductionist views: that consciousness and phenomenal qualities are unreal because irreducible;¹ that they are irreducibly

Original version published in J. Tomberlin, ed., *Philosophical Perspectives 4, Action Theory and the Philosophy of Mind*, pp. 81–108, 1990. Copyright © Ridgeview Publishing Company 1990. Revised version published in N. Block, O. Flanagan, & G. Güzelde, *The Nature of Consciousness* (MIT Press), 1997. Reprinted with permission of Ridgeview Publishing Company.

non-physical-functional facts;² that they are forever mysterious, or pose an intellectual problem different from other empirical problems, or require new conceptions of the physical.³

It is my view that we can have it both ways. We may take the phenomenological intuition at face value, accepting introspective concepts and their conceptual irreducibility, and at the same time take phenomenal qualities to be identical with physical-functional properties of the sort envisaged by contemporary brain science. As I see it, there is no persuasive philosophically articulated argument to the contrary.

This is not to deny the power of raw metaphysical intuition. Thoughtful people compare phenomenal qualities and kinds of physical-functional property, say the activation of neural assemblies. It appears to them to be an evident and unmediated truth, independent of further premises, that phenomenal qualities cannot be identical with properties of those types or perhaps of any physical-functional type. This intuition is so compelling that it is tempting to regard antiphysicalist arguments as rationalizations of an intuition whose independent force masks their tendentiousness. It is the point of this chapter to consider the arguments. But I will also present a positive account of the relation between phenomenal concepts and physical properties that may provide some relief, or at least some distance, from the illusory metaphysical intuition.

In recent years the central problem with physicalism has been thought by many to be “the explanatory gap.” This is the idea that we cannot *explain*, in terms of physical-functional properties, what makes a certain experience ‘feel like this,’ in the way we can explain what makes a certain substance a liquid, say. It is concluded that physicalism is defective in some respect, that there cannot be a (proper) reduction of the mental to the physical. Before we consider this explanatory gap, we must first examine, in some detail, a more basic antiphysicalist line of reasoning that goes back to Leibniz and beyond, a leading version of which is now called the knowledge argument. Answering this argument will generate a framework in which to address antiphysicalist concerns in general.

1. The Knowledge Argument and Its Semantic Premise

The knowledge argument is straightforward on the face of it. Consider any phenomenal quality

and any physical property however complex. We can know that a person has the physical property without knowing that she experiences the phenomenal quality. And no amount of a priori reasoning or construction can bridge this conceptual gap. That is the intuitive premise. The conclusion is drawn that the phenomenal quality cannot be identical with the physical property. The argument is equivalent to this: since physical and phenomenal conceptions can be connected only a posteriori, physical properties must be distinct from phenomenal properties.

The best known and liveliest version of the knowledge argument is Frank Jackson’s, which features the physiologically omniscient Mary, who has never seen color and so does not know what it is like for us to see red, despite her knowing all the physical-functional facts about us.⁴ She later sees colors, and thus learns what it has been like all along for us to see red. She learns a new fact about us. Jackson concludes that this fact is not among the physical facts, since Mary already knew them. It is not difficult to see that this argument depends on a more or less technical premise.

In my view, the physicalist should accept Jackson’s intuitive description of Mary: she fails to know that we have certain color experiences even though she knows all relevant physical facts about us. And when she acquires color experience, she does learn something new about us—if you like, learns a new fact or truth. But this is to be granted, of course, only on an *opaque* reading of ‘Mary learns that we have such and such color experiences,’ and on corresponding readings of ‘learns a new fact or truth about us.’ For as regards the transparent versions of those ascriptions of what Mary did not know and then learned, they would beg the question, amounting to this: ‘as for the property of having such and such color experiences, Mary did not know, but then learned, *of* that property that we have it.’ Physicalists reject this, for according to us those experiential properties are physical properties, and Mary already knew of all our physical properties that we have them—under their physical descriptions. What she lacked and then acquired, rather, was knowledge of certain such properties couched in experiential terms.

Drawing metaphysical conclusions from opaque contexts is risky. And in fact inferences of Jackson’s form, without additional premises, are open to straightforward counterexamples of a familiar sort. Let me describe two cases.

(1) Max learns that the bottle before him con-

tains $\text{CH}_3\text{CH}_2\text{OH}$. But he does not know that the bottle contains alcohol. This holds on an opaque reading: he would not assert that there's stuff called alcohol in the bottle, or that the bottle contains the intoxicating component of beer and wine. Let sheltered Max even lack the ordinary concept 'alcohol.' After he acquires that ordinary concept, he learns something new—that the bottle contained alcohol. If the knowledge argument has a generally valid form, we could then infer from Max's epistemic situation that alcohol is not identical with $\text{CH}_3\text{CH}_2\text{OH}$. Evidently this does not follow.

(2) Margot learns about the element Au and reads that people decorate themselves with alloys of Au. But she has never seen gold and cannot visually identify it: she lacks an adequate visual conception. She later is shown some gold and forms a visual conception of it, "that stuff," and she acquires a new piece of information—individuated opaquely—to the effect that those previously read about embellishments are made of that stuff. Again, if the knowledge argument were unrestrictedly valid, it would follow that that stuff is not identical with Au. This case differs from the case of Max by involving not a descriptive mode of presentation but (as we might say) a perceptual mode of presentation.

It is not difficult to find a difference between both these cases and the case of Mary. Max lacks knowledge of the bottle's contents under a contingent description of it—"ingredient of wine and beer that makes you intoxicated." What Margot lacks is a certain visual conception of Au, which is to say gold. This typically would not be a descriptive conception; it would not self-consciously take the form "the stuff that occasions this type of visual experience." Still on the face of it such a concept implicates a visual-experience type. For it picks out the kind it picks out by virtue of that kind's occasioning experiences of that type. And that is a crucial *contingency* in how the concept that Margot lacks is related to its reference. I hope I will be understood, then, if I say that the visual take on Au that Margot lacks would have conceived Au 'under a contingent mode of presentation.'

This brings us back to Mary, whose acquired conception of what it is like to see red does not conceive it under a contingent mode of presentation. She is not conceiving of a property that presents itself *contingently* thus: it is like such and such to experience *P*. Being experienced like that is essential to the property Mary conceives. She conceives it directly. When Mary

later acquires new information about us (construed opaquely), the novelty of this information cannot be explained—as in the case of Margot—as her acquiring a new contingent mode of presentation of something she has otherwise known of all along. She has a *direct* grasp of the property involved in the new information; she conceives of it somehow, but not under a contingent mode of presentation. Proponents of the knowledge argument will say that is why it is valid on an opaque reading: there is no contingency in Mary's conception of the new phenomenal information that explains it as a novel take on old facts. She learns new facts simpliciter and not new conceptions of old facts.

Notice how close this comes to Saul Kripke's well-known antiphysicalist argument (1980). Kripke assumes that a phenomenal concept such as 'pain' cannot be a priori linked with a physical concept such as that of the stimulation of C-fibers. The case of Mary is a vivid way of making the same point. Kripke points out that property identities can be true even if not a priori, for example, 'heat = such and such molecular property.' It seems fair to represent the next step in his argument as follows. 'Heat' has a contingent higher-order mode of presentation that connotes the property 'feeling like this.' That is what accounts for the a posteriori status of the identity. But, as Kripke points out, this cannot be how 'pain' works: the phenomenal concept 'pain' does not pick out its referent via a contingent mode of presentation; it conceives pain directly and essentially. Kripke concludes that pain is not identical with a physical property.

The two arguments then turn on the same implicit assumption. The only way to account for the a posteriori status of a true property identity is this: one of the terms expresses a contingent mode of presentation. This ought to be given a place of prominence.

(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.

The knowledge argument and Kripke's argument then depend on two assumptions: the conceptual independence of phenomenal concepts and physical-functional concepts, which I accept, and the semantic premise, which I deny.

The antiphysicalist intuition that links concept-individuation and property-individuation (more closely than is in my view correct) is per-

haps this. 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.

But as against this, if a phenomenal concept can pick out a physical property directly or essentially, not via a contingent mode of presentation, and yet be *conceptually independent* of all physical-functional concepts, so that Mary's history is coherent, then Jackson's and Kripke's arguments are ineffectual. We could have two conceptually independent conceptions of a property, neither of which connote contingent modes of presentation, such that substituting one for the other in an opaquely interpreted epistemic context does not preserve truth. Even granting that our conception of phenomenal qualities is direct, physicalism would not entail that knowing the physical-functional facts implies knowing, on an opaque construal, the phenomenal facts; and so the failure of this implication would be quite compatible with physicalism. The next few sections give an account of phenomenal concepts and properties that would justify this claim.

2. Recognitional Concepts

Phenomenal concepts belong to a wide class of concepts that I will call recognitional concepts. They have the form 'x is one of *that* kind'; they are type-demonstratives. These type-demonstratives are grounded in dispositions to classify, by way of perceptual discriminations, certain objects, events, situations. Suppose you go into the California desert and spot a succulent never seen before. You become adept at recognizing instances, and gain a recognitional command of their kind, without a name for it; you are disposed to identify positive and negative instances and thereby pick out a kind. These dispositions are typically linked with capacities to form images, whose conceptual role seems to be to focus thoughts about an identifiable kind in the absence of currently perceived instances. An image is presumably 'of' a given kind by virtue of both past recognitions and current dispositions.

Recognitional concepts are generally formed against a further conceptual background. In

identifying a thing as of a recognized kind, we almost always presuppose a more general type to which the kind belongs: four-legged animal, plant, physical thing, perceptible event. A recognitional concept will then have the form 'physical thing of that (perceived) kind' or 'internal state of that kind,' and so forth.⁵

Here are some basic features of recognitional concepts that it will help to have in mind in connection with the account of phenomenal concepts that follows.

1. You can understand 'porcelain' from a technical description and only later learn visually, tactually, and aurally to recognize instances. By contrast, in the phenomenon I mean the concept is recognitional at its core; the original concept is recognitional.
2. A recognitional concept need involve no reference to a past instance, or have the form 'is of the same type as that (remembered) one.' You can forget particular instances and still judge 'another one of those.'
3. Recognitional abilities depend on no consciously accessible analysis into component features; they can be irreducibly gestalt.
4. Recognitional concepts are perspectival. Suppose you see certain creatures up close and form a recognitional concept—'those creatures₁'; and suppose you see others at a distance, not being able to tell that they are of the same kind (even when they are), and form another recognitional concept—'those creatures₂'. These concepts will be a priori independent. Now the respect in which they differ is *perspectival*, in some intuitive sense. A recognitional concept is in part individuated by its constitutive perspective. Here is the important point: a recognitional concept can be ascribed outside its constitutive perspective; 'that thing (seen at distance) is one of those creatures₁ (seen up close)' makes perfectly good sense. This plays a key role below in the account of third-person ascriptions of phenomenal concepts.

(This casual invoking of reference-determining dispositions will be a red flag for many who are aware of the vexing foundations of the theory of reference. Problems about referential scrutability, rule-following, naturalizing intentionality—however one wishes to put it—are as frustrating as any in contemporary philosophy. I do not propose to address them here. The idea rather is to appeal to unanalyzed common sense concerning a natural group of concepts and ap-

parent conceptual abilities. The apparent irreducibility of phenomenal qualities itself arises from appeal to intuitions independent of the theory of reference; and it seems reasonable that we should, in resolving that issue, appeal to notions that arise at the same intuitive level. That we *appear* to have recognitional concepts and identifying dispositions that are more or less determinate in their reference is hard to deny. My conception of 'those hedges' [seen around the neighborhood] may unambiguously pick out a variety of eugenia. An example closer to the present topic is this. We can imagine an experiment in which the experimenter tries to determine which internal property is the focus of her subject's identifications: 'again,' . . . 'there it is again.' There seems no commonsensical implausibility—putting aside foundational worries about the inscrutability of reference—in the idea that there is a best possible answer to the experimenter's question, in the scientific long run.⁶⁾

3. Phenomenal Concepts as Recognitional Concepts

Here is the view to be defended. Phenomenal concepts are recognitional concepts that pick out certain internal properties; these are physical-functional properties of the brain. They are the concepts we deploy in our phenomenological reflections; and there is no good philosophical reason to deny that, odd though it may sound, the properties these conceptions *phenomenologically reveal* are physical-functional properties—but not of course under physical-functional descriptions. Granted that brain research might discover that (what we take to be) our phenomenal concepts do not in fact discriminate unified physical-functional properties. Failing that, it is quite coherent for a physicalist to take the phenomenology at face value: the property of *its being like this* to have a certain experience is nothing over and above a certain physical-functional property of the brain.

Phenomenal concepts are conceptually independent of physical-functional descriptions, and yet pairs of such concepts may converge on, pick out, the same properties. Rebutting the semantic premise of the knowledge argument requires making sense of the idea that phenomenal concepts conceive physical-functional properties 'directly,' that is, not by way of contingent modes of presentation. The objective is

to show that the knowledge argument fails for the same reason in the case of Mary as in the case of Max: both arguments require substitution in opaque contexts of terms that are conceptually independent. In the case of Max, the conceptual independence appears to derive from 'alcohol's connoting a contingent mode of presentation that is metaphysically independent of the property referred to by the chemical concept. In the case of Mary it has a different source.

What then accounts for the conceptual independence of phenomenal and physical-functional concepts? The simple answer is that recognitional concepts and theoretical concepts are in general conceptually independent. It is true that recognitional concepts other than phenomenal concepts connote contingent modes of presentation that are metaphysically independent of the natural kinds they pick out, and hence independent of the kind referred to by the theoretical term of the pair. But we need not count this metaphysical independence as essential to the conceptual independence of corefering recognitional and theoretical concepts. Concepts of the two sorts have quite different conceptual roles. It is hardly surprising that a recognitional conception of a physical property should discriminate it without analyzing it in scientific terms. Nor should it be surprising that, if there are recognitional concepts that pick out physical properties *not* via contingent modes of presentation, they do not discriminate their references by analyzing them (even implicitly) in scientific terms. Basic recognitional abilities do not depend on or get triggered by conscious scientific analysis. If phenomenal concepts reflect basic recognitions of internal physical-functional states, they *should* be conceptually independent of theoretical physical-functional descriptions. That is what you expect quite apart from issues concerning physicalism.

An antireductionist may reply that the physicalist view depends on an ad hoc assumption and that it is tendentious to suppose that phenomenal concepts differ from all other recognitional concepts in not having contingent modes of presentation.

But this is not fair. Even on the antiphysicalist view, phenomenal concepts are recognitional concepts, and we have 'direct' recognitional conceptions of phenomenal qualities, that is, conceptions unmediated by contingent modes of presentation. Evidently it would be absurd to insist that the antiphysicalist hold that we con-

ceive of a phenomenal quality of one kind via a phenomenal mode of presentation of a distinct kind. And why should the physicalist not agree that phenomenal recognitional concepts are structured in whatever simple way the antiphysicalist requires? That is after all the intuitive situation, and the physicalist simply claims that the intuitive facts about phenomenal qualities are compatible with physicalism. The physicalist makes the additional claim that the phenomenal quality thus directly conceived is a physical-functional property. On both metaphysical views, phenomenal concepts differ from other recognitional concepts; phenomenal concepts are a peculiar sort of recognitional concept on any account, and that can hardly count against physicalism. The two views agree about conceptual structure and disagree about the nature of phenomenal qualities. To insist that physicalism implies, absurdly, that phenomenal concepts could pick out physical properties only via metaphysically distinct phenomenal modes of presentation is unmotivated. There is, though, still more to be said about whether phenomenal concepts should be regarded as having modes of presentation of some sort, and we continue the account in section 5.

Suppose this account of how phenomenal concepts refer is true. Here is a semantic consequence. The physicalist thesis implies that the judgments "the state *a* feels like that" and "the state *a* has physical-functional property *P*" can have the same truth condition even though their joint truth or falsity can be known only a posteriori. I mean, same condition of truth in a possible world. For truth conditions are determined in part by the possible world satisfaction conditions of predicates; and if a phenomenal predicate directly refers to a physical property, that property constitutes its satisfaction condition.

On this account, a phenomenal concept rigidly designates the property it picks out. But then it rigidly designates the same property that some theoretical physical concept rigidly designates. This could seem problematic, for if a concept rigidly designates a property not via a contingent mode of presentation, must that concept not capture the *essence* of the designated property? And if two concepts capture the essence of the same property, must we not be able to know this a priori? These are equivocating uses of 'capture the essence of.' On one use, it expresses a referential notion that comes to no more than 'directly rigidly designate.' On the other, it means something like 'be conceptually inter-

derivable with some theoretical predicate that reveals the internal structure of' the designated property. But the first does not imply the second. What is correct in the observation about rigid designation has no tendency to imply that the two concepts must be a priori interderivable.

4. The Concept 'Phenomenal Concept'

Not all self-directed recognitional concepts are phenomenal concepts, as may be seen in these two cases.

(1) Cramps have a characteristic feel, but they are not feelings. Cramps are certain muscle contractions, while feelings of cramp are, if physical, brain states. (Witness phantom-limb sufferers.) One has a recognitional concept that picks out certain muscle contractions in the having of them. This is not a phenomenal concept, for it does not purport to pick out a phenomenal quality. But of course, in exercising this concept, one often conceives its reference by way of a phenomenal mode of presentation, a cramp feeling or a cramp-feeling image.

(2) A more fanciful self-directed nonphenomenal concept can be conceived. To begin with, consider blindsight. Some cortically damaged people are phenomenally blind in restricted retinal regions; and yet when a vertical or horizontal line (say) is presented to those regions, they can, when prompted, guess what is there with a somewhat high degree of correctness. We can extend the example by imagining a blindsight that is exercised spontaneously and accurately. At this point we shift the focus to internal properties and conceive of a self-directed recognitional ability, which is like the previous ability in being phenomenally blank and spontaneous but which discriminates an internal property of one's own. If this recognitional ability were suitably governed by the concept 'that state,' the resulting concept would be a self-directed recognitional concept that is phenomenally blank.

The two examples show that 'phenomenal concept' cannot mean 'self-directed recognitional concept.' This is compatible with my proposal. For it implies neither (a) that we can reductively explicate the concept 'phenomenal quality' as 'property picked out by a self-directed discriminative ability,' or (b) that we can reductively explicate the concept 'phenomenal concept' as 'self-directed recognitional concept.' Phenomenal concepts are certain self-

directed recognitional concepts. Our higher-order concept 'phenomenal concept' cannot be reductively explicated, any more than can our concept 'phenomenal quality.' The higher-order concept 'phenomenal concept' is as irreducibly demonstrative as phenomenal concepts themselves.

5. Phenomenal Modes of Presentation

Self-directed recognitional concepts of the blindsight type might appear to raise a problem for the claim that phenomenal concepts pick out physical-functional properties directly. Here is a way to put the point.

The difference between a self-directed blindsight recognitional concept and a phenomenal concept appears to be that the latter involves a phenomenal mode of presentation while the former conceives its referent in some other, odd, way. So, if the phenomenal concept is taken to discriminate some physical property, it then does so via a phenomenal mode of presentation. But that conflicts with your assertion that phenomenal concepts refer directly, with no contingent mode of presentation. A similar point arises concerning recognitional concepts of cramps and of cramp feelings. Both concepts must presumably have modes of presentation. It is far-fetched to suppose that one of them has and the other lacks a mode of presentation; the phenomenal concept does not pick out a physical state *nakedly*. The 'cramp' concept connotes a mode of presentation of the form 'the physical state that causes such and such phenomenal state.' If we attempt to capture the phenomenal concept analogously, its mode of presentation would have the form 'the state that has such and such phenomenal aspect.' But then, contrary to what the physicalist must say, phenomenal concepts point to physical states only by way of phenomenal modes of presentation.

What might an antiphysicalist say about these various self-directed recognitional concepts? Let me make a good-faith attempt to present a reasonable version.

(1) A cramp concept picks out a muscular property indirectly, by way of a causal chain that is mediated by the phenomenal quality associated with the concept. In addition to this mode of presentation type—the phenomenal quality—we can also note the role of, as we might say, "token modes of presentation." One and the same cramp concept (type) can on dif-

ferent occasions be focussed differently: by an actual cramp feeling, by a cramp-feeling image, or by an imageless inclination to identify cramp feelings when they occur (with a cramp-feeling image on the tip of one's imagination.)

(2) We turn from cramp concepts to cramp-feeling concepts. These do not refer (i.e., to cramp feelings) by way of contingent modes of presentation. But they can mimic the working of cramp concepts as regards "token modes of presentation." If one can focus attention on the bodily property of cramp by way of a token cramp feeling, surely one can focus attention on the phenomenal quality cramp-feeling by way of a token cramp feeling. The same goes for cramp-feeling images and those gossamer identifying inclinations. Should antiphysicalists say that cramp-feeling concepts have 'noncontingent' modes of presentation? We might say that a phenomenal concept has as its mode of presentation the very phenomenal quality that it picks out. We might also say that phenomenal concepts have "token modes of presentation" that are noncontingently tied to the phenomenal qualities to which those concepts point: particular cramp feelings and images can focus one's conception of the phenomenal quality of cramp feeling.

(3) As for self-directed blindsight concepts, the antiphysicalist then ought to say, they differ from phenomenal concepts in the obvious way, whether one puts it by saying that they lack the noncontingent phenomenal modes of presentation (types) that phenomenal qualities have, or that they lack their phenomenal "token modes of presentation."

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

A phenomenal concept exercised in the absence of the phenomenal quality it stands for often involves not merely a recognitional disposition but also an image. And so, as a psychological state in its own right, a phenomenal concept—given its intimate connection with imaging—bears a phenomenological affinity to

a phenomenal state that neither state bears to the entertaining of a physical-theoretical concept. When we then bring phenomenal and physical-theoretical concepts together in our philosophical ruminations, those cognitive states are phenomenologically so different that the illusion may be created that their references must be different. It is as though antiphysicalist intuitions rest on a resemblance theory of mental representation, as though we conclude from the lack of resemblance in our phenomenal and physical-functional conceptions a lack of sameness in the properties to which they refer.

6. Third-Person Ascriptions

Ascriptions of phenomenal qualities to others ostensibly refer to properties that others may have independently of our ascribing them.⁷ we have realist conceptions of the phenomenal states of others. But at the same time they are projections from one's own case; they have the form 'x has a state of this sort,' where the demonstrative gets its reference from an actual or possible state of one's own.

Can phenomenal concepts as we predicate them of others be identified with the recognitional concepts we have characterized? A question naturally arises how essentially self-directed recognitional concepts can be applied in cases where it makes no sense to say that one can directly apply these concepts. This is a question that exercised Wittgensteinians.

As we have already pointed out, recognitional concepts are perspectival, in the sense that their reference is determined from a certain constitutive perspective (depending on the concept). The above concept 'those creatures_i' (seen up close) picks out a creature-kind that one discriminates on close sightings. But nothing prevents ascribing the recognitional concept 'one of those creatures_i' to something observed from a different perspective, seen in the distance or heard in the dark. We have to distinguish the perspective from which reference is determined and the far broader range of contexts in which the referentially fixed concept can be ascribed. The former perspective hardly restricts the latter contexts. This holds also for phenomenal concepts. We acquire them from a first-person perspective, by discriminating a property in the having of it. Assuming that we successfully pick out a more or less determinate physical property, the extraperspectival ascription 'she is in a

state of *this* kind' makes complete sense. And so it is not easy to see that Wittgensteinians succeeded in raising a philosophical problem that survives the observation that we can discriminate physical properties and so fix the reference of phenomenal concepts from a first-person perspective, and then go on to ascribe those concepts third-personally.

There is though a more up-to-date worry about the interpersonal ascribability of first-person concepts, however physical we suppose their references to be. Evidently there will be vagueness, and indeterminacy, concerning whether another person—whose neural assemblies will presumably always differ from mine in various respects—has a certain physical property that I discriminate phenomenally. And this on the face of it poses a problem, which may be framed as follows:

The question whether another person's phenomenal states resemble yours can hardly consist in their neural assemblies' resembling yours. Any physical similarity you choose will be arbitrarily related to a given phenomenal similarity. Suppose there is a small physical difference between a neural state of yours and another person's state. What makes it the case that this small neural difference constitutes a small phenomenal difference or a large one or no phenomenal difference at all? It appears that there cannot be a fact of the matter.

But this objection appears to me to overlook a crucial element of the physicalist view we have presented—that phenomenal concepts are (type) demonstrative concepts that pick out physical properties and relations. A first step in answering it is to consider the connection between interpersonal and intrapersonal phenomenal similarity. It appears that one's phenomenological conception of how others' phenomenal states resemble one's own has to be drawn from one's idea of how one's own phenomenal states resemble each other. A person's quality space of interpersonal similarity must derive from her quality space of intrapersonal similarity. How else is one to get a conceptual grip on interpersonal phenomenal similarity? This seems inevitable on any account—physicalist or antiphysicalist—on which phenomenal concepts are formed from one's own case.

But conceptions of phenomenal similarity relations are as much type-demonstrative concepts as those of phenomenal qualities. All one can apparently mean by "that spectrum of phenomenal similarity" is "*that ordering* among my phe-

nominal states.” Physicalism implies that if such a type-demonstrative refers, it picks out a physical ordering. And there is no obvious philosophical difficulty (if we put aside scepticism in the theory of reference) in the idea that discriminations of resemblances and differences among one’s own phenomenal properties pick out reasonably well defined physical relations.

Now I have to confess some uneasiness about extending this to interpersonal similarity without qualification; but the implications of the foregoing remarks are clear enough. If they are correct, whatever physical ordering relations are picked out by one’s personal notions of phenomenal similarity must also constitute (what one thinks of as) interpersonal phenomenal similarity. It is easy to see that there still is room here for further trouble. But the difficulty the objection raises seems considerably diminished if one insists on the demonstrative nature of all phenomenal concepts, however relational and of whatever order. For the objection then becomes, “Suppose there is a small physical difference between a neural state of yours and another person’s state. What makes it the case that this small neural difference constitutes a small difference of *that* type, or a large one, or no difference of *that* type at all?” If “that type” picks out a physical relation, then the question answers itself, and there seems no gloomy philosophical threat of phenomenal incommensurability.

Naturally there is the risk that physical investigation will not deliver the right physical properties and relations. Even if the risk is increased by bringing in interpersonal similarities, the nature of the risk is the same as in one’s own case: the phenomenal might turn out to be not adequately embodied.

It goes without saying that one can coherently conceive that another person has *P*, conceived in physical-functional terms, and doubt that she has any given phenomenal quality; that has been central to this chapter. But one cannot coherently wonder whether another person in a *P* state has a state with *this* phenomenal quality if one acknowledges that one’s concept ‘this quality’ refers to the property the concept discriminates in oneself (what else?) and that moreover it discriminates *P*.

Why then is there an apparent problem of other minds? It is as if one wishes to do to others as one does to oneself—namely, apply phenomenal concepts directly, apply phenomenal recognitional capacities to others from a first-person perspective. The impossibility of this

can present itself as an epistemological barrier, as something that makes it impossible to know certain facts. Doubtless more can be said in explanation of the naturalness of the conflation of the innocuous conceptual fact with a severe epistemological disability. It is not easy to shake the grip of that conflation or therefore easy to dispel the problem of other minds. The cognitive remedy, the fortification against the illusion, is the idea of recognitional concepts that can be ascribed beyond their constitutive perspective, coupled with the reflection that there is no reason to doubt that it is physical-functional properties that those recognitional concepts discriminate.

7. Knowing How versus Knowing That

Consider a different physicalist reply, to an antiphysicalist argument posed in this form: “knowledge of physical-functional facts does not yield knowledge of the phenomenal facts; therefore phenomenal facts are not physical-functional.” Laurence Nemirow and David Lewis have replied in effect that the premise is true only if you equivocate on “knowledge.”⁸ The first occurrence means theoretical knowledge, the second the ability to discriminate introspectively or to imagine certain properties. But theoretical knowledge of physical-functional properties that are identical with phenomenal qualities does not yield the other sort of knowledge of the same properties, that is, the ability to discriminate them in introspection or to imagine them. There are two epistemic relations to one class of properties.

Now this suggests something significantly different from my account. On the Nemirow-Lewis proposal, the only knowledge “that such and such” is knowledge couched in physical-functional terms, while what corresponds to (what we have been calling) phenomenal concepts is knowing how to identify or to imagine certain states. What I have proposed is evidently different. Knowing that a state feels a certain way is having distinctive information about it, couched in phenomenal conceptions. There is of course a central role for recognitional abilities, but that is in the constitution of phenomenal concepts. Antiphysicalists are right to count phenomenal knowledge as the possession of distinctive information, for it involves genuinely predicative components of judgment, whose

association with physical-functional concepts is straightforwardly a posteriori.

Physicalists are forced into the Nemirow-Lewis reply if they individuate pieces of knowledge or cognitive information in terms of possible-world truth-conditions, that is, hold that 'knowing that *p*' and 'knowing that *q*' ascribe distinct pieces of knowledge just in case 'that-*p*' and 'that-*q*' denote distinct sets of possible worlds. Then knowing that *x*'s phenomenal qualities are such and such will be distinct from knowing that *x*'s physical properties are so and so only if the former qualities are distinct from the latter properties. So then a physicalist who counts the basic antiphysicalist premise as true on some interpretation must deny either that knowledge, cognitive information, is individuated in terms of possible-world truth-conditions or deny that knowing the phenomenal facts (in the sense that makes the basic antiphysicalist premise true) is knowing that such and such or having distinctive information about it. Nemirow and Lewis deny the latter. Of course I deny the former; there are ample independent reasons to deny it, and it seems otherwise unmotivated to deny the latter.

There are straightforward reasons to prefer the phenomenal concept view.

1. A person can have thoughts not only of the form "coconuts have *this* taste" but also of the form "if coconuts did not have *this* taste, then *Q*." You may get away with saying that the former expresses (not a genuine judgment but) the mere possession of recognitional know-how. But there is no comparable way to account for the embedded occurrence of "coconuts have this taste"; it occurs as a predicate with a distinctive content.
2. We entertain thoughts about the phenomenal states of other people—"she has a state of that type"; this clearly calls for a predicative concept. It does of course involve a recognitional ability, but one that contributes to the formation of a distinctive concept.
3. For many conceptions of phenomenal qualities, there is no candidate for an independently mastered term that one then learns how to apply: thinking of a peculiar way my left knee feels when I run (a conception that occurs predicatively in various judgments) is not knowing how to apply an independently understood term. I suppose a functionalist might say that, in such cases, one implicitly individuates the state in terms of some functional description that is

fashioned on the spot, but this appears psychologically implausible.

8. The Explanatory Gap

Can we *explain* how a certain phenomenal property might be identical with a certain physical-functional property? The answer is no, and then again, yes.

First, the no. When we explain, say, liquidity in physical-functional terms, the explanation is in crucial part a priori. You may find this surprising; but what we in effect do is analyze liquidity (or more precisely those aspects of liquidity that we count as explained⁹) in terms of a functional description, and then show that the physical theory of water implies, a priori, that the functional description is realized. But given the conceptual independence of phenomenal concepts and physical-functional concepts, we cannot have such an a priori explanation of phenomenal qualities in physical-functional terms.

Does this matter? The explanatory gap, as it appears to me, is an epistemic or conceptual phenomenon, without metaphysical consequences,¹⁰ and it is predictable from the physicalist account we have proposed. But this may seem somewhat glib. As Georges Rey points out (Rey 1995), the mere fact of conceptual inequivalence for recognitional type-demonstratives and descriptive terms does not generate an explanatory gap. Many examples would make the point. We do not find a troubling explanatory gap in judgments of the form "that stuff is CH₃CH₂OH," even though this does not hold a priori.

Now what is it that needs accounting for? This seems to me to be it: how identity statements that connect phenomenal concepts and physical-functional concepts can be true despite our sense that, if true, they *ought to be* explanatory and yet are not. We can explain how such identity statements fail to be both explanatory (conceptual independence) and true; but this does not account for the thought that something that ought to be there is missing. We have to explain away the intuition that such identity statements ought to be explanatory.

There must be something special about phenomenal concepts that creates the expectation and the consequent puzzle. We have already seen a significant difference between phenomenal concepts and all other phenomenally mediated recognitional concepts. Might this make

the difference here as well? That is what I will try to show.

Perhaps this is why we think that true phenomenal-physical identity judgments ought to be explanatory. It is natural to regard our conceptions of phenomenal qualities as conceiving them as they are in themselves, that is, to suppose we have a direct grasp of their essence. So in this respect there is a parallel with liquidity: the phenomenal concept and the concept 'liquid' both pick out properties directly, that is, not via contingent modes of presentation. And of course the physical-functional theoretical term of the identity, couched in fundamental theoretical terms, also reveals the essence of the property it picks out. Since both conceptions reveal this essence, then, if the psychophysical identity judgment is true, the sameness of that property, it might seem, ought to be evident from those conceptions, as in the liquidity case. The physical-functional concept structurally analyzes the property, and so we expect *it* to explain, asymmetrically, the phenomenal quality, much as physics explains liquidity, on the basis of an a priori analysis. The fact that this is not so makes it then difficult to understand how there can be just one property here.

If this is what makes the explanatory gap troubling, then the idea that phenomenal concepts are recognitional concepts of a certain sort does account for the explanatory gap in a way compatible with physicalism. Phenomenal concepts, as we have seen, do not conceive their reference via contingent modes of presentation. And so they can be counted as conceiving phenomenal qualities directly. Calling this a grasp of essence seems to me all right, for phenomenal concepts do not conceive their references by way of their accidental properties. But this is quite a different grasp of essence than we have in the term "liquid": for that term (or what there is in it that we count as functionally explained) is conceptually equivalent to some functional description that is entailed by the theoretical term of the identity.

The problem of the explanatory gap stems then from an illusion. What generates the problem is not appreciating that there can be two conceptually independent "direct grasps" of a single essence, that is, grasping it demonstratively by experiencing it, and grasping it in theoretical terms. The illusion is of *expected transparency*: a direct grasp of a property ought to reveal how it is internally constituted, and if it is not revealed as physically constituted, then it is

not so. The mistake is the thought that a direct grasp of essence ought to be a transparent grasp, and it is a natural enough expectation.

The explanatory gap has led many philosophers of mind seriously astray into mistaken arguments for epiphenomenalism, for mystery, for eliminativism. At the root of almost all weird positions in the philosophy of mind lies this rather elementary and unremarkable conceptual fact, blown up into a metaphysical problem that appears to require an extreme solution. But it is a mistake to think that, if physicalism is true, consciousness as we conceive it at first hand needs explaining in the way that liquidity as we ordinarily conceive it gets explained.

There is another interpretation of "can we understand how physicalism might be true?," for which the answer is clearly yes. For we can explain, and indeed we have explained, how a given phenomenal concept can manage to pick out a particular physical-functional property without remainder: the concept discriminates the property but not via a contingent mode of presentation. This in its way closes the explanatory gap between the phenomenal and the physical. We understand how "such and such phenomenal quality" could pick out physical property *P*, even though "such and such phenomenal quality = *P*" does not provide an (a priori) explanation in physical terms of why a given phenomenal quality feels as it does. Since the former, when generalized, would entail that physicalism about phenomenal qualities is true, and since we understand both of these things, we thereby understand how physicalism can be true.

9. Subjective Concepts and Subjective Properties

You can ascribe an objective property—one completely expressible in the objective terms of natural science—under a subjective conception: '*x*'s state has *this* quality.' Thomas Nagel writes that mental facts are "accessible only from one point of view."¹¹ This does reflect something about phenomenal concepts; they are in some intuitive sense "from a point of view" and moreover subjective. Phenomenal concepts are subjective because they are essentially self-directed, involving capacities to discriminate certain states in the having of them and also involve imaginative capacities anchored in such recognitional capacities. If that is it, then Nagel takes

a correct observation about concepts and draws a wrong conclusion about facts and properties. For concepts can in that sense be “from a point of view” and subjective, and still introduce properties that are exhaustively captured in objective science.

But we can go further. Let us grant even that the *property* of experiencing such and such is aptly counted as subjective, as intrinsically involving a point of view. Why should this subjectivity not itself be identical with a physical-functional property, and therefore completely objectively conceivable under its physical-functional description? There is no contradiction in supposing that a property that is subjective—in the sense of being individuated in a way that invokes a relation to a mind—is also conceivable under an objective mode of presentation. There is no incoherence in the thought that the “subjectivity” of a phenomenal quality is identical with an objective physical-functional aspect of that property.

Does a fully objective description of reality not still leave something out, viz. the subjective conceptions? This is a play on ‘leave something out.’ A complete objective description leaves out subjective conceptions, not because it cannot fully characterize the properties they discriminate or fully account for the concepts themselves as psychological states but simply because it does not employ them.

10. Phenomenal Structure, and Exotic Others

Some functionalists might think this account ignores a major feature of our conceptions of the mental, namely, their systematic structure. We have conceptions of different sensory modalities, and of intramodality comparisons along various spectra, of pitch, timbre, hue, brightness, shape, size, texture, acidity, acridity, and so on. These could be seen as subsidiary functional organizations within a theory of the mental. Antiphysicalists may share something of the point, wanting to speak of phenomenological structures. My account could seem to imply that phenomenal concepts are atomistic, unstructured, unsystematic, for are these recognitional dispositions not in principle independent of each other?

We have phenomenal recognitional concepts of various degrees of generality. Some are of highly determinate qualities, and others are of

phenomenal determinables: crimson, dark red, red, warm colored, colored, visual. The last is the recognitional conception of a whole sensory modality. And there is the most general of all, the recognitional concept *phenomenal* (state, quality), the highest ranking phenomenal determinable. (This is a recognitional concept. One discriminates phenomenal states from nonphenomenal states, feeling a twinge from having a bruise, hearing a chirp from jerking a knee, and that highly general discriminative capacity is the basis of the concept of a phenomenal quality.)

There are also relational concepts: quality *x* is a determinate of quality *y*; quality *x* is more like quality *y* than like quality *z*; quality *x* is of a different modality from quality *y*. These are also recognitional concepts: dispositions to classify together, on phenomenal grounds, certain pairs and triples of phenomenal qualities. Combining them yields complex conceptions of abstract phenomenal structures, for example, of a structured sensory modality. One’s general conception of such a structure is in effect one’s ability to exercise in concert a group of such general phenomenal concepts.

Now it is important that our conceptions of such phenomenal structures, while abstract, are yet phenomenal conceptions. No purely functional conception of a complex structure, however isomorphic to a phenomenal-structure conception it may be, will be cognitively equivalent to it; purely functional conceptions ignore that the structures are of phenomenal similarity relations, of phenomenal determinateness, and so on.

But given the falsity of the semantic minor premise, that is no impediment to holding that those abstract phenomenal conceptions can have purely functional or physical-functional structures as their references. For such structures may well be what these abstract phenomenal recognitional capacities in fact discriminate. Indeed we may go on to say that, if our phenomenal conceptions are to be fully vindicated by brain science, then the brain must have a certain functional structure; any possible totality of (as it were) semantic values for our phenomenal conceptions must have certain functional structures. This perhaps explains the strong intuition of some commonsense functionalists that phenomenal concepts are functional concepts, without our having to accept that counterintuitive view.

“Can your projection analysis accommodate the thought that a bat has highly specific, deter-

minate, phenomenal states that are not like anything I can experience or imagine? It seems to me that your program will require you to bring in the bat's own recognitional-imaginative capacities, such as they are."¹²

When one thinks about a bat's sonar phenomenal states, one thinks about them as phenomenal, that is, as having in common with my phenomenal states what I discriminate them all as having in common, and that may be something physical-functional. One also thinks of them as of a distinctive phenomenal kind or modality, different from one's own states, of roughly that order of determinateness at which one's visual states are marked off from one's auditory states. One has such a general concept from one's own case, and one can project it. Again, that concept—'distinctive phenomenal modality'—may denote a physical-functional property of sets of phenomenal states. And one thinks of the bat's sonar states as exhibiting phenomenal variation of different degrees of specificity. These conceptions of general phenomenal structure, determinable-determinate relations, resemblance relations, and so on, we have, as I have said, from our own case.

Now nothing in the foregoing requires that a necessary condition of having certain phenomenal qualities is having the capacity to discriminate them. (See, however, the discussion below of transparency.) We ascribe to bats not phenomenal concepts but phenomenal states; and we do that by projection, in the manner characterized above. Other-directed phenomenal conceptions are of others' states, and not as such of their conceptions.

Nagel proposes that we can achieve objectivity about the mental by abstracting from subjective conceptions of our own psychology, fashioning objective mental conceptions that are neither physical nor functional.¹³ This would enable us to conceive abstractly of mental lives of which we have no subjective, projective, understanding whatever. Now that is evidentially at odds with my proposal. It appears to me that all mental concepts that are not functional concepts (where the latter include concepts of theoretical psychology) are subjective-projective concepts, however general and abstract they may be. The reason is simple: as far as I can determine, I have no objective nonfunctional mental concepts. If I try to conceive an alien mind in nonfunctional mental terms, I rely on concepts like 'sensory modality' and other general conceptions of phenomenological structure of the

sort mentioned above, and I understand them from my own case. They are abstract conceptions; but, it appears to me, they are still recognitional concepts and hence as subjective as the highly specific phenomenal concept of having an itch in the left ankle.

11. Transparency

The following could appear possible on my account: another person is in the state that in me amounts to feeling such and such but sincerely denies feeling anything relevant. It apparently has been left open that others have phenomenal states that are not introspectable at will, for no requirement of transparency has been mentioned. Then the property that is the referent of my concept of feeling like *that* could, even if it occurs transparently in me, occur nontransparently in you. But (the objection continues) denying transparency is tantamount to allowing unconscious experiences; and it would not be unreasonable to say that the topic of phenomenal states is the topic of certain conscious states.

There really is no issue here. Suppose that any phenomenal quality must be essentially transparent, and that no property I correctly identify as phenomenal can be realized in another nontransparently. If cognitive integration is essential to the intuitive property of transparency, so be it; there is no reason to think that such integration itself is not a physical-functional property, as it were implicated by each phenomenal property.

But it is not obvious that phenomenal properties must be transparent in such a reflexive cognitive sense. What about infants and bats? There has always been a philosophical puzzle about how subtracting reflexive cognitive awareness from phenomenal or conscious states leaves something that is still phenomenal or conscious. But that puzzle is independent of the present account. All that is implied here is that if I have a conception of a phenomenal quality that is shared by me and an infant, my conception of it involves a recognitional concept, and there is no reason why that phenomenal quality itself should not be a physical-functional property. Whatever indefinable, elusive aspect of phenomenal qualities might constitute their being conscious—transparent in some appropriately minimal sense—without requiring reflexive conceptualizability, there would be no reason to doubt it is a physical-functional property.

12. Incorrigeability

Physicalism, it may be said, cannot acknowledge the incorrigibility of phenomenal judgments of the form 'it feels like that.' For surely there is no guarantee that a capacity for recognizing a given physical property does not at times misfire; and perhaps even more to the point, there can be no guarantee that to a given recognitional disposition there corresponds a repeatable physical property. Perhaps an antiphysicist will grant that certain kinds of mistake about phenomenal qualities are possible;¹⁴ but the antiphysicist will insist that we cannot be wrong in thinking that *there are* phenomenal qualities.

Now suppose it turns out that no system of physical-functional properties corresponds to the system of our phenomenal concepts. Would a physicalist not then have to say there are no phenomenal qualities? And is the fact that physicalism leaves this open not a serious problem?

But that very possibility ought to make us dubious about the incorrigibility of the judgment that there are real phenomenal repeatables. What reason have we to think that our phenomenal judgments discriminate real properties? Memory, one might say, cannot be that mistaken: we can hardly deny that present inner states resemble past states in ways we would recognize again. Despite this conviction, however, if no system of physical-functional properties corresponded to one's putative phenomenal discriminations, an alternative to nonphysical qualities would be this: memory radically deceives us into thinking we discriminate internal features and nonrandomly classify our own states. Strong evidence that no suitable physical-functional properties exist might amaze and stagger one. It would then have emerged that we are subject to a powerful illusion, a cognitive rather than a phenomenal illusion; we would be judging falsely that we thereby discriminate real properties.

It does seem likely that we genuinely discriminate internal physical-functional states in introspection.¹⁵ But with that said, positing nonphysical properties to forestall the *possibility* of radical error, however theoretically adventurous (even reckless) this may be, would in something like a moral sense still be rather faint-hearted. The whole point about the phenomenal is how it appears. And that means there is no introspective guarantee of *anything* beyond mere appearance, even of discriminations of genuine repeatables. The dualist balks at the implications and invents a realm of properties to ensure that the

appearances are facts, but this does not respect the truly phenomenal nature of what is revealed by introspection at its least theoretical.

I have to grant that, if it were to turn out that no brain properties are suitably correlated with our ascriptions of phenomenal qualities, one might well feel some justification in questioning physicalism. But that does not imply that one now has such a justification. There is no good reason for prophylactic dualism.

13. Functionalism

There are two functionalist theses: that all concepts of mental states are functional concepts, and that all mental properties are functional properties. The first I rejected in accepting the antiphysicist intuition. I agree with the antiphysicist that phenomenal concepts cannot be captured in purely functional terms. But nothing in philosophy prevents phenomenal properties from being functional properties. There are two possibilities: they are commonsense-functional properties or they are psychofunctional, and I take the latter to be the interesting one.¹⁶ Might the phenomenal quality of seeing red be identical with a property captured by a detailed psychological theory? This would be so if the repeatable that triggers one's phenomenal concept 'seeing red' has psychofunctional rather than say biochemical identity conditions. That this is possible has been denied by antifunctionalist physicalists on the grounds of inverted qualia and absent qualia possibilities, but I do not find these arguments persuasive.

The inverted qualia argument is commonly advanced against identifying phenomenal qualities with commonsense functional properties and also against the psychofunctional identification. The position I espouse is agnostic: for all philosophers know, phenomenal qualities are psychofunctional, neurofunctional, or some other fine-grained functional properties. The opposing argument is that it is possible that the functional role that seeing red has in me is had in you by, as I would think of it, seeing green. If this is, as they say, metaphysically possible, then of course phenomenal qualities are not functional properties.

But it seems the only argument for the possibility is the coherent conceivability of inverted qualia. One cannot presuppose that inverted qualia are *nomologically* possible. There seems to be no philosophical reason to assert that, apart

from the coherent conceivability of inverted qualia. If there is empirical reason to assert that nomological possibility, then of course we should retreat from agnosticism. The present point is that nothing about the idea of inverted qualia provides philosophical reason to reject functionalism about qualia. For that would require another version of the antiphysicalist argument: it is conceivable that any given functional state can occur without the seeing of green and with the seeing of red, say; therefore the psychofunctional role and the phenomenal quality involve distinct properties. Clearly one cannot accept this argument against functionalism without also accepting the analogous argument against physicalism itself; the philosophical antifunctionalist argument requires a premise that implies antiphysicalism.

There is a well-known absent qualia argument against functionalism by Ned Block (1978). Suppose the Chinese nation were organized so as to realize the psychofunctional organization of a person seeing green. Evidently the Chinese nation would not collectively be seeing green or having any other sensation. Any psychofunctional property could in this way be realized without a given phenomenal quality and hence cannot be identical with one. Now this argument might appear dialectically more telling than the inverted qualia argument, for it apparently rests on more than a conceptual possibility. It seems a plain truth that the Chinese people would not thereby be having a collective sensation. Surely it is barmy to be agnostic about that. Block suggests a principle. "If a doctrine has an absurd conclusion which there is no independent reason to believe, and if there is no way of explaining away the absurdity or showing it to be misleading or irrelevant, and if there is no good reason to believe the doctrine that leads to the absurdity in the first place, then don't accept the doctrine."¹⁷

While we doubtless find an absurdity in ascribing phenomenal qualities to the Chinese nation as a whole, the matter is not so simple. It is hard to see how such a judgment of absurdity can be *justified* except by our having some intuitive

knowledge of the nature of phenomenal qualities whereby we can say that the Chinese nation cannot have them collectively. Have I a special insight into my physical states whereby I can say: the repeatable that I reidentify whenever I attend to my seeing green is not a functional property? One feels sceptical that introspection can yield such knowledge. If the argument is not 'they do not collectively have, by virtue of their functional organization, however fine-grained, what I have when *this* occurs,' then what is it? Is a further philosophical argument in the offing? It is difficult to see whence chest-beating to the contrary derives its credibility. Perhaps a dualist conception of Platonic insight into mental essences might help. But, on a naturalist view of human nature, one ought to find it puzzling that we have such a first-person insight into the nature of our mental properties. Perhaps there is reason to suppose that what one introspects and reidentifies is a categorical and not a dispositional property. That has an intuitive ring to it, but it is not that easy to produce a decent argument for it. We are left with this question: how might we know short of detailed brain research that what we reidentify in ourselves when we see green is not a fine-grained functional property? But if we cannot know this by sheer insight into the essence of our own properties, or by philosophical argument, then we cannot know that the Chinese nation lacks what we have. Our ignorance concerns the nature of our own properties, and that ignorance would appear to prevent drawing substantive conclusions from thought experiments of this type.

There is no question that ordinary intuition counts strongly against applying phenomenal concepts to things that are not single organisms, and one cannot deny that the reply just given makes one uncomfortable, at the very least. And yet the alternative appears to be Platonism about mental essences, and that sits awkwardly with naturalism. It is possible that phenomenal qualities are biochemical properties: and yet again it is difficult to see that philosophers know anything that implies that they are not fine-grained functional, or neurofunctional, properties.¹⁸

NOTES

1. Cf. Rey (1995) and Dennett (1991).
2. Jackson (1982, 1986).
3. Nagel (1974, 1986); McGinn (1993).
4. Jackson (1982, 1986).

5. How such background concepts themselves arise is not my topic; but we might think of them variously as deriving from more general recognitional capacities, or as functions of complex inferential roles, or

- as socially deferential; or they may be components of innate structures. Background concepts are not always presupposed. Someone may be extremely good at telling stars from other objects (e.g., lightning bugs, airplanes, comets, planets) without having any real idea of what they are.
6. For more on recognitional concepts and on the determinacy of reference, see Loar (1990; 1991, 1995).
 7. The earlier version of this chapter made heavy weather of third-person ascription of phenomenal concepts. General considerations about the perspectival nature of recognitional concepts permit a far neater account, which I here present.
 8. Nemirow (1980); Lewis (1983).
 9. This leaves open the possibility of twin-Earth cases in which the apparently defining properties of liquidity—those that are functionally explained—are kept constant across worlds even though the underlying kind changes. The defining properties then turn out to be merely reference-fixing.
 10. For illuminating accounts of the explanatory gap and its significance see Levine (1983, 1993). Levine's diagnosis of the significance of the explanatory gap is different from mine.
 11. Nagel (1974).
 12. Thomas Nagel, in a note commenting on an earlier draft.
 13. Nagel (1974).
 14. See Warner (1986).
 15. When I see a ripe lemon in daylight and attend to my visual experience, I form the memory belief that what I introspect is what I introspected (phenomenologically inclined as I am) the last time I saw a ripe lemon in daylight. It seems a reasonable empirical inference that probably ripe lemons in such circumstances cause in me states that my memory accurately records as the same. But this inference is, I take it, not reasonable on introspective grounds alone; it presupposes much about how the world works.
 16. It is empirically unlikely that phenomenal qualities are identical with commonsense functional properties. Here is one way to see this. We know sensations can be produced by nonstandard means, that is, by poking around in the brain; but this of course is no part of the commonsense functional role of the property of seeing red. Now suppose this property is produced in me by a brain probe. What constitutes its being a sensation of red? If it is its commonsense functional role, then that property would be the sensation of red by virtue of (something like) its *normally* having such and such causes and effects (it doesn't have them here). But this makes sense only if the property in question is itself a *distinct* lower-order property about which it is contingently true that normally it has such and such causes and effects although it lacks them here. That lower-order property would then be a far better candidate (than the commonsense functional property) for being the property one's phenomenal conception discriminates. For this reason, such brain probes turn out to be strong and perhaps even conclusive evidence that phenomenal qualities, the ones we discriminate in applying phenomenal concepts, are not identical with commonsense functional properties. There are other ways of reaching the same conclusion.
 17. Block 1978.
 18. (Original version) For pointing out a substantial error in an ancestor of the paper, I am indebted to George Myro, whose correction put me on the right track as I now see it. I have learned much from conversations on phenomenal qualities with Janet Levin and Richard Warner. Stephen Schiffer made several valuable suggestions about the structure of the paper and got me to clarify certain arguments. I am also grateful for comments on the mentioned ancestor to Kent Bach, Hartry Field, Andreas Kemmerling, Dugald Owen, Thomas Ricketts, Hans Sluga, Stephen Stich, and Bruce Vermazen.
- (Revised version) Many thanks to Ned Block for raising questions about modes of presentation and the blindsight case, to Georges Rey for making me see that more needed to be said about the explanatory gap, and to Kent Bach for helpful remarks on a number of points.

REFERENCES

- Block, N. (1978). "Troubles with Functionalism," in C. Wade Savage, ed., *Perception and Cognition: Issues in the Foundations of Psychology*. Vol. 9, *Minnesota Studies in the Philosophy of Science*. Minneapolis: University of Minnesota Press.
- . (1990). "Inverted Earth." *Philosophical Perspectives* 4, 53–79.
- Dennett, Daniel (1991). *Consciousness Explained*. Boston: Little, Brown.
- Harman, Gilbert (1990). "The Intrinsic Quality of Experience," *Philosophical Perspectives* 4, 31–52.
- Jackson, Frank (1982). "Epiphenomenal Qualia," *Philosophical Quarterly* 1982, 127–36.
- . (1986). "What Mary Didn't Know," *Journal of Philosophy* 83: 291–95.
- . (1994). "Armchair Metaphysics," in M. Michael ed., *Philosophy in Mind*. Norwell, MA: Kluwer.
- Kripke, Saul (1980). *Naming and Necessity*. Cambridge, MA: Harvard University Press.
- Levin, Janet (1983). "Functionalism and the Argument from Conceivability," *Canadian Journal of Philosophy*, Supplementary Volume 11.
- . (1986). "Could Love Be Like a Heatwave?" *Philosophical Studies*, 49: 245–61.
- Levine, Joseph (1983). "Materialism and Qualia: The Explanatory Gap," *Pacific Philosophical Quarterly*, 64: 354–61.
- . (1993). "On Leaving Out What It Is Like," in M. Davies and G. Humphreys, eds., *Consciousness*. Oxford: Blackwell.
- Lewis, David (1983a). "Mad Pain and Martian Pain," in *Philosophical Papers*, Vol. 1. Oxford: Oxford University Press.
- . (1983b). "Postscript" to the foregoing.
- Loar, Brian (1990). "Personal References," in E. Vil-

- lanueva, ed., *Information, Semantics and Epistemology*. Oxford: Blackwell.
- . (1991). "Can We Explain Intentionality?," in G. Rey and B. Loewer, eds., *Meaning in Mind*. Oxford: Blackwell.
- . (1995). "Reference from the First-Person Perspective," in *Philosophical Issues* vol. 5.
- McGinn, Colin (1993). "Consciousness and Cosmology: Hyperdualism Ventilated," in M. Davies and G. Humphreys, eds., *Consciousness*. Oxford: Blackwell.
- Nagel, Thomas (1974). "What Is It Like to Be a Bat?" *Philosophical Review*, 1974: 435–50.
- . (1986). *The View from Nowhere*. Oxford: Oxford University Press.
- Nemirow, Laurence (1980). Review of Nagel's *Mortal Questions*, *Philosophical Review*, July 1980.
- Rey, Georges (1995). "Towards a Projectivist Account of Conscious Experience," in T. Metzinger, ed., *Essays on Consciousness*.
- Warner, Richard (1986). "A Challenge to Physicalism," *Australasian Journal of Philosophy*, 64: 249–65.
- . (1993). "Incorrigibility," in H. Robinson, ed., *Objections to Physicalism*, Oxford: Oxford University Press.

Two Conceptions of the Physical

Daniel Stoljar

The debate over physicalism in philosophy of mind can be seen as concerning an inconsistent tetrad of theses: (1) if physicalism is true, a priori physicalism is true; (2) a priori physicalism is false; (3) if physicalism is false, epiphenomenalism is true; (4) epiphenomenalism is false. This paper argues that one may resolve the debate by distinguishing two conceptions of the physical: on the *theory-based conception*, it is plausible that (2) is true and (3) is false; on the *object-based conception*, it is plausible that (3) is true and (2) is false. The paper also defends and explores the version of physicalism that results from this strategy.

1

One way to view the contemporary debate in philosophy of mind over physicalism is to see it as being organized around an inconsistent tetrad of theses. These are:

- (1) If physicalism is true, a priori physicalism is true.
- (2) A priori physicalism is false.
- (3) If physicalism is false, epiphenomenalism is true.
- (4) Epiphenomenalism is false.

It is obvious of course that these theses *are* inconsistent: (1) and (2) entail that physicalism is false, while (3) and (4) entail that it is true. Barring ambiguity, therefore, one thing we know is that one of the theses is false.

On the other hand, each of the theses has powerful considerations, or at least what seem initially to be powerful considerations, in its favor.¹ In support of (1) are considerations of supervenience, articulated most clearly in recent times by Frank Jackson and David Chalmers. A priori physicalism is a thesis with two parts. The first part—the physicalist part—is that the mental supervenes with metaphysical necessity on the physical. The second part—the a priori part—is that mental truths are a priori entailed by physical truths. Many philosophers hold that supervenience stands in need of justification or explanation; Jackson and Chalmers argue that the project of justifying or explaining supervenience *just is* the project of making it plausible that there is an a priori entailment of the mental by the physical. This suggests that the first part of a priori physicalism inevitably involves the second. By considerations of supervenience, therefore, (1) is true: if physicalism is true, a priori physicalism is true.²

In support of (2) are considerations of the apparent epistemic distinctness of qualia from