

## 6

# An Externalist's Guide to Inner Experience

*Benj Hellie*

The aim of John M. Hinton's *Experiences: An Inquiry into Some Ambiguities* is set out in its introduction as follows:

Someone who has more sympathy with traditional empiricism than with much of present-day philosophy may ask himself, 'How do my experiences give rise to my beliefs about an external world, and to what extent do they justify them?' He wants to refer, among other things, to unremarkable experiences, of a sort which he cannot help believing to be so extremely common that it would be ridiculous to call them common experiences. Drawing a breath of fresh air is still a very common experience in many parts of the globe, but he does not mean that kind of an experience. Like the common experience of doing a monotonous job of work amid exhausting noise, it would be too much of an event in what he calls the external world. He mainly has in mind sense-experiences, and he thinks of them in a particular way. His way of thinking of them, roughly speaking as something 'inner,' is one on which recent logico-linguistic philosophy has thrown a good deal of light. I still hope to throw a little more, mainly or wholly reflected, light on it, . . . where the visual case is concerned. (Hinton 1973, 1)

Thanks to Stephen Biggs, Bill Brewer, Mike Caie, John Campbell, Tim Crane, Terry Cullen, Imogen Dickie, Kati Farkas, Dominic ffytche, Chris Hill, Jim John, Matt Kennedy, Michael Lachelt, Geoff Lee, Fiona MacPherson, Chris Mole, John Morrison, Jennifer Nagel, Adam Pautz, Ian Phillips, Diana Raffman, Gurpreet Rattan, Howard Robinson, Susanna Siegel, Charles Siewert, Nico Silins, Declan Smithies, Paul Snowdon, Scott Sturgeon, Brad Thompson, Charles Travis, Jonathan Weisberg, the Toronto M&E Working Papers Group, and audiences at the 2008 Syracuse Philosophy Annual Workshop and Network, 'Consciousness and Thought' (Dubrovnik, 2008), 'Hallucination on Crete' (Rethymno, Crete, 2008), 'The Philosophical Significance of Attention' (Dubrovnik, 2009), and the University of Barcelona; and especially to Alex Byrne, David Chalmers, Bence Nanay, and Jessica Wilson. Thanks to Ranpal Dosanjh for LaTeX-to-RTF conversion.

I follow Delia Graff Fara's liberating convention of using single-quotes for all sorts of quotation devices.

In June 2009, after this chapter was largely complete, I came to regard as false the doctrine that anything deserving the name 'consciousness' or 'experience' has the passive character of the visual and sensational states under discussion in the chapter. I came to this opinion in part as a result of the anxieties gestured at in the parenthetical remark at the close of section 3. Unfortunately, this doctrine is a fundamental presupposition of the chapter, and of the literature to which it attempts to contribute. And yet I continue to regard the ideas in the chapter as worth publishing, because I believe both that its conciliatory project leads inexorably to these anxieties about the doctrine of the passivity of consciousness and that if this doctrine were to be dislodged from its current position as a starting point for all discussion in the philosophy of mind, that would be a good thing.

My aim here can be described in much the same way; the central difference is that the light I hope to reflect on the subject matter emits from philosophy that is still more recent, and of a rather more epistemo-semantic than logico-linguistic character.

In a nutshell, working from the perspective of the ‘externalist’—a philosopher who thinks that, in good circumstances, one’s visual consciousness is in part ‘an event in the external world’—I provide a conceptual and metaphysical interpretation of the view that consciousness is ‘something ‘inner’.”

More expansively, I take for granted a direct realist position on which, when one is seeing, one’s perceptual consciousness is a relation to particular objects in one’s environment and their intrinsic features, and also the doctrine of the ‘transparency of experience.’ My aim then is to strike a balance between keeping externalism vital—by refraining as far as possible both from watering it down with qualifications and from larding it up with finicky theoretical commitments—and clawing back as much as possible of the notion of *phenomenality*: roughly, a notion of ‘what an experience is like,’ bound to a range of internalist intuitions such as the possibility of perfect hallucination and the possibility of ‘spectral inversion’ without error.

This interpretive task is important to the future of externalism for two reasons. First, we have an evident curiosity about the nature of phenomenality. This is reflected in the fact that a great deal of theorizing about consciousness is, whether explicitly or implicitly, in fact theorizing about phenomenality. Most significant, it is phenomenality that is believed to give rise to the ‘explanatory gap’ (Levine 1983): on one way of understanding it, the failure of *a priori* entailment of certain beliefs distinctively concerning phenomenality by physical or biological theory. And the explanatory gap, in turn, gives rise to the widely discussed threat of dualism (Chalmers 2009). Accordingly, on pain of failure to address much of what motivated us to concern ourselves with consciousness in the first place, the externalist needs to explain what is going on in all this research.

Second, phenomenality is widely seen as a paradigm aspect of consciousness. Accordingly, if the externalist posits aspects of consciousness beyond the phenomenal while failing to explain how they relate back to this paradigm, the externalist will be at risk either of being accused of changing the subject (as Fiona MacPherson has put it, of engaging in mere ‘terminological hoodwinking’), or—if, making appeal to the conceptual fundamentality of the concept of consciousness, in reply the externalist insists on a shared subject matter—of saying something obviously false. Moreover, I argue, the externalist needs to provide a story of this relationship from which it can be argued that the phenomenal properties are not the only properties that deserve to be regarded as genuine aspects of consciousness. Given the strength of the internalist tradition, a theory that cannot on its own terms show phenomenal properties to be less than fully fundamental is unlikely to make much dialectical progress when allowed out into the wild.

Fortunately, if my arguments are correct, the externalist is in an excellent position. Externalism can be taken neat. The relation of *perceiving*,

understood as a sort of primitively conscious openness to one's environment, generates an explanatory gap all on its own. Almost all internalist intuitions can be chalked up to the structure of our reasoning about the world in response to experience; the remainder can perhaps be ignored. Although my story inevitably links such reasoning to intrinsic features of some sort or other, we can remain entirely neutral on the nature of these features. Indeed, we must remain neutral on this for the time being, because philosophers have yet to produce any arguments which constrain this matter. Putting it roughly, phenomenality can be thought of as a sort of 'projection' by thought of these mysterious features onto experience.

Section 1 provides an explicit characterization of the externalist view under discussion, contrasting its notion of consciousness with phenomenality. Sections 2 and 3 provide the analytical part of my interpretation of phenomenality. I argue that if there is a coherent notion of phenomenality at odds with the externalist's notion of consciousness, it is rooted in a kind of introspective or reflective indiscriminability. Section 2 makes some framework points about what it would take to analyze phenomenality and discusses the notion of indiscriminability in general. Section 3 develops a 'transparency'-based view of reflective knowledge of perceptual experience according to which we extract such knowledge from logical operations on the Fregean senses of perceptual demonstrative judgments. This view yields patterns of reflective indiscriminability tracking the notion of phenomenality, and requires only that these senses somehow track intrinsic aspects of one's condition. This discussion yields a range of compatibilist conclusions: for instance, transparency and direct realism are compatible with 'qualia.' Section 4 turns at last to the metaphysics of phenomenality, developing my claim that phenomenality is a projection onto experience of the intrinsic aspects tracked by the senses of perceptual demonstrative concepts.

A recurrent theme in the chapter will be that a pair of epistemic limitations through which the distinct may fail to be discriminable—ignorance and error—have significantly different 'valences.' Internalists have tended to regard ignorance as their touchstone, assimilating the distinction between seeing and hallucination to spectral inversion without illusion; externalists, by contrast, have focused exhaustively on error, treating hallucination differently from seeing but, when not ignoring inversion without illusion, implausibly dismissing the possibility. But by fully exploiting the resources inherent in each limitation, a theory can respect both one's conscious openness to the environment and one's subjective contribution to one's sense of the environment.

## 1. EXTERNALIST PERCEPTUAL CONSCIOUSNESS

In this section I characterize a certain strong version of an externalist theory of perceptual consciousness. Rather than arguing for the view, I assume an audience antecedently sympathetic toward externalism about

consciousness. After stating the view I provide four arguments that the externalist view of consciousness recognizes aspects of consciousness beyond the purely phenomenal. I then explain where the externalist should locate the explanatory gap.

### 1.1. Theories of Consciousness

To begin with, I briefly explain what I mean by ‘theory of consciousness.’ By ‘consciousness’ I mean to express the concept ‘consciousness’ or ‘experience’ and thereby to refer to that property, *consciousness*, that all and only the experiences share (here and throughout quantifiers and expressions such as ‘extension’ are understood as concerning actual and possible entities). Consciousness generates an explanatory gap: we can conceive of a zombie, a being like us physically but lacking in consciousness. Our concept of consciousness is therefore irreducible to any physical or functional concepts (even if consciousness itself is identical to some physical or functional property), so it will be therefore be a convenience to treat consciousness as an extremely natural property, at least within our image of the world.<sup>1</sup>

The notion of a *theory* of consciousness can be sharpened with appeal to two pieces of metaphysical apparatus. The first is the notion of a *real definition* of a property, the familiar notion applying to the answer to the question of *which* property the property is, or of what *kind* of property the property is, and is answered by providing a ‘definition’ or ‘canonical conception’ of the property, or a conceptually most fundamental description of what it is to have that property (e.g., to have the property *water* is to have the property  $H_2O$ ; accordingly,  $H_2O$  is the real definition of water).

The second is the notion of a *determinate* of a property. For example, red is a determinate of color, scarlet is a determinate of red, and the *exact* shade of scarlet instantiated in this thread of the people’s flag is a maximal determinate of scarlet, red, and color. Sometimes the determinate–determinable relation is cashed out so that *F* is a determinate of *G* if *F* metaphysically necessitates *G*. On my usage, determinate–determinable is more restrictive than metaphysical necessitation: The determinates of a property somehow form a natural family of properties. So, for instance, on my use,

1. I am inclined to agree with Chalmers (unpublished manuscript) that the concept of consciousness is conceptually fundamental and cannot be elucidated or defined in more basic terms. Still, when using any public language term to express a concept, there is a risk that just *which* concept the term is used to express might be misunderstood. To forestall such misunderstanding, let me say that I do not mean ‘access consciousness’ (Block 1995) or any etiolated substitute such as the representability or representation of some state of a system by another state of that system, at least if representation is understood in terms of *ceteris paribus* nomic covariation. Rather, by ‘consciousness’ I mean to express the concept that those of us who take zombies to be conceivable apply to the most general respect in which zombies differ from their experiencing twins.

*iodine* is a determinate of the determinable *chemical element*, while *iodine in Chicago* is not.<sup>2</sup>

A bit of terminology will improve readability: I will say that a determinate of *consciousness* is an *experiential property*. Then, a 'theory of consciousness' in my sense provides an at least partial answer to the question 'what are the *real definitions* of the *experiential properties*?'

Many familiar views can be seen as providing (somewhat abstract) answers to this question. For instance, *representationalism*: Every experiential property is, definitionally, a representational property. Or *qualia theory*: Every experiential property is, definitionally, a monadic, intrinsic, normatively inert feature. Or *sense-datum theory*: Every experiential property among the class of those distinctively instantiated in *perceptual* experience is, definitionally, a case of bearing a primitive relation of awareness ('acquaintance') to a mental particular (the 'sense-datum'). Or *direct realist* theory: Every experiential property among the class of those distinctively instantiated in *veridical* perceptual experience is, definitionally, a case of bearing acquaintance to an entity *in the subject's environment*. (Henceforth in this context I will leave off the qualifier 'definitionally.') In each case, the theory aims to provide a conceptually maximally fundamental characterization of the different varieties of conscious experience.

Note a contrast between the representational and qualia theories, on the one hand, and the sense-datum theory, on the other: The former pair are *general* theories of the natures of experiential properties; by contrast, the latter is more modest in its aim to provide only a theory of *perceptual* consciousness, of the experiential properties distinctive of *perceptual* experiences (where I mean these to include not just cases of genuine perception but also cases of perception-like dreams and other 'perceptual hallucinations').

## 1.2. Externalism and Its Antecedents

The externalist view I discuss here is rooted in the direct realist view just mentioned. Explicitly, the view is as follows:

In a typical veridical perceptual experience, a great many of the experiential properties will be of the form *being a case of seeing o's F-ness*, where *o* is a material particular in the subject's near environment and *F* is a color actually instantiated in visible parts of *o*.

2. Perhaps my commitments here can be weakened slightly: All I insist on is that my subject matter concerns a *natural family* of properties that in one sense or another are *varieties* of *consciousness* and that come in more-or-less specific levels. So it would not trouble me if it turned out that my subject is more aptly treated in terms of some such relation as *genus-species*. The aim is to capture the Williamsonesque sense that the wide properties I claim to be determinates of *consciousness* are not merely aggregates that involve some purely narrow variety of *consciousness* in an entirely consciousness-free environment (Williamson 2000, ch. 1).

Henceforth this is the view I refer to as ‘externalism.’ (I assume throughout that material particulars and their colors are entirely mind-independent.)<sup>3</sup>

Externalism can be compared with several more familiar antecedents. A first antecedent is a singularist representational theory of visual consciousness. On such a theory, some experiential properties of visual experiences are representational properties with *singular content*. For instance, if one sees or hallucinates one’s mother as wielding a tire iron, then, it is alleged, among the experiential properties of one’s visual experience is a singular representational property concerning one’s mother to the effect that she is wielding a tire iron (for discussion, see Siegel 2005, §5.1). On this view, experiential properties are, as I shall say, *haecceitistic*: They reflect the particular ‘thisness’ of objects in the environment. This contrasts with a nonhaecceitistic ‘abstract’ version of representationalism on which the represented conditions are purely general (for discussion of abstract representationalism, see Chalmers 2006; Byrne and Logue 2008).

A second antecedent is a ‘Russellian’ representational theory of visual consciousness. On such a theory, some experiential properties of visual experiences are representational properties with content that is, as it were, *Russellian in its predicate position*. For instance, if one sees or hallucinates something as red, then, it is alleged, among the experiential properties of one’s visual experience is a representational property concerning redness, to the effect that something before one instantiates it (for discussion, see Chalmers 2004, §6). On this view, experiential properties are, as I shall say, *quidditistic*: They reflect the particular ‘thusness’ of properties in the environment. This contrasts with a nonquidditistic ‘Fregean’ version of representationalism on which the represented conditions quantify over properties in the environment (for discussion of Fregean representationalism, see Chalmers 2004; Thompson 2009).

A third antecedent is a *factive/relational* theory of veridical visual consciousness. On such a theory, some experiential properties of *veridical* visual experiences are *factive* properties concerning facts about the subject’s here-and-now environment, or *relational* properties relating the subject to entities in the here-and-now environment. For instance, if one sees a black shoe on a brown chair (as such), then, it is alleged, among the experiential properties of one’s visual experience is a property of bearing a certain factive attitude toward the fact that a black shoe is on a brown chair, or a property of being related to a black shoe’s state of being on a brown chair (for discussion, see Campbell 2002;

3. I do not want to deny the existence of such ‘qualified’ experiential properties as *blurrily seeing a computer screen* and *seeing a white wall with extensive phenomenal noise*, but I ignore these properties here.



Hellie 2007a, §§1–2). This contrasts with representationalist views on which experiential properties interact with the environment not directly, by containing the environment, but merely by imposing conditions of correctness that the environment is in the business of rendering met or unmet.

Externalism is factivist/relational, and, like the singularist and Russellian approaches, it builds the haecceitistic and quidditistic aspects of the environment into experiential properties.

To foreshadow a bit, I treat the factivist/relational aspect of externalism rather differently from its haecceitistic and quidditistic aspects. The crucial epistemic limitation that generates internalistic intuitions about the latter is our propensity to *ignorance*; about the former, our propensity to *error*.

### 1.3. Externalism and the Phenomenal

It may perhaps be clear at this point that, according to externalism, at least some experiential properties are not phenomenal properties. To ensure that it is, I will now advance four arguments for this thesis. For now, I will assume that the reader shares an intuitive sense for the notion of phenomenality as used in the literature and sketched in the introduction: roughly, 'what an experience is like' understood as a purely internal matter; the next section takes a more analytical approach.

The first argument concerns *twins*. Suppose that Bill sees a red tomato, *t*, while Tina (whose brain and environment are qualitatively identical to Bill's) sees a numerically distinct tomato, *t'*. Going by the common conception of phenomenality, Bill's and Tina's experiences are paradigms of experiences that are just alike phenomenally. But externalism predicts that they are distinct experientially. On externalism, among the experiential properties of Bill's visual experience is *being a case of seeing the redness of t*; it also predicts that while this is not among the experiential properties of Tina's visual experience, *being a case of seeing the redness of t'* is. (Henceforth, I ignore all experiential properties not explicitly mentioned in this and the following arguments.) Our two subjects thus have two experiences that are the same phenomenally but, according to externalism, differ experientially.

The second argument concerns *spectral inversion*. Suppose that Inez is spectrally inverted with respect to Bill: Her visual system is so wired that, in a situation that causes Bill's visual system to go into the intrinsic state it typically goes into when he sees a red thing in the position he sees *t*—let us call this state '*R*'—Inez's goes into the typical state Bill's visual system goes into when he sees a green thing in that position (call this '*G*'), and so forth. And suppose that Inez sees a green tomato *t''* (in shape and other relevant properties, qualitatively identical to *t*). Many authors have had a certain pair of intuitions about this sort of case: first, that Bill's and Inez's experiences are the same phenomenally; second, that while Bill sees the

redness of  $t$ , Inez sees the *greenness* of  $t''$ .<sup>4</sup> If so, then plausibly externalism predicts that among the experiential properties of Inez's experience is *being a case of seeing the greenness of  $t''$*  (and that *being a case of seeing the redness of  $t$*  is *not* among those experiential properties). Accordingly, these two subjects are phenomenally the same but (according to the externalist) experientially different.

The third argument concerns *illusion*. Suppose that Ilya is a normal subject, seeing a tomato  $t'''$  of a rare, white heirloom variety under a carefully tuned red spotlight; his environment is (in all other respects) qualitatively identical to Bill's, as is his visual system (both are in the state  $R$ ). Going by the common conception of phenomenality, Bill's and Ilya's experiences are paradigms of experiences that are just alike phenomenally. But externalism predicts that they are distinct experientially. After all, there is no redness for Ilya to see: at best, he sees the whiteness of his tomato (as well, plausibly, as its property of *being illuminated in red*). Accordingly, these two subjects are phenomenally the same but (according to the externalist) experientially different. While it is not a strict consequence of externalism, I assume throughout that Ilya's experiential properties include *being a case of seeing the whiteness of  $t'''$* .

The fourth argument concerns *hallucination*. Suppose that Dean is dreaming of a red tomato (without seeing a red tomato, of course) and that his visual system is in exactly the same condition  $R$  as Bill's. Externalism as stated makes no predictions about what the experiential property of Dean's dream *is*.<sup>5</sup> But since *being a case of seeing the redness of  $t$*  is *not* among the experiential properties of Dean's experience, externalism predicts that his experience differs in experiential properties from Bill's. Still, going by the common understanding of the notion of phenomenality,

4. In the literature, both intuitions are contested. Granting both makes life *harder* for the friend of externalism, so I should not be accused of begging any important questions if I operate under the assumption that both are legitimate. Moreover, the latter intuition is often expressed in more theory-laden terms that make certain assumptions about representation. It seems to me that my way of putting the point is on an equally good footing. For discussion of the intuitions, see Chalmers 2004, §§5–9.

5. In my view, it may be unknowable what the property is that replaces the external properties; perhaps there are none. Considering the failure of an attempt to constrain the nature of this property will be instructive.

Johnston (2004) advances a certain phenomenological judgment on behalf of the view that a hallucination is a relation to a complex uninstantiated property: namely that Dean's experience puts him in a position to refer to redness (Hawthorne and Kovakovich (2006) and Pautz (chapter 10 this volume) advance refinements of Johnston's argument). It is not credible to me that this could be the brute content of a first-person judgment, so Johnston's judgment must be justified inferentially.

On the present approach, we can diagnose the reasoning to which Johnston must be implicitly appealing as running something like this: (i) Dean's experience has *seeing this color* as an experiential property; (ii) *this color* is redness; (iii) when one's experience has *seeing F-ness* as an experiential property, this enables one to refer to *F-ness*; hence Dean's experience enables him to refer to redness. While step (iii) seems fine, step (ii) concerns the quiddity of a property one sees, and hence, in light of considerations from spectral inversion, one could not hope to establish this claim by first-person reflection. But more relevant in the present context is that (i) is false, and is based on the false presupposition, perhaps requisite for reflection on perceptual experience, that one sees: for more on this see section 3.1.



Bill's and Dean's experiences are paradigms of experiences that are the same phenomenally.

We want a final example, a subject phenomenally *unlike* the other four. This subject will be Greg: a normal subject, seeing green tomato Inez's *t* under normal conditions.

A quick classificatory comment (using the terminology of Byrne and Logue 2008): Since it provides a distinct theory of Bill's and Dean's experiential properties, externalism is in one sense a 'disjunctivist' theory; since, as extended, it asserts that Ilya's experience is more similar to Bill's than Dean's, externalism counts as a VIVH disjunctive view rather than a VVIH disjunctive view.

#### 1.4. Externalism and the Explanatory Gap

Where should the externalist locate the explanatory gap? The externalist's relational properties are asserted to be experiential, or forms of consciousness, so we should expect that they generate an explanatory gap: after all, *consciousness* does. In which aspect of these relational properties is the gap rooted?

To begin with, it would not be plausible for the externalist to locate the explanatory gap in *objects* or in their *colors*. Our judgments about Bill, Tina, and Inez seem to prevent doing so. After all, to commit to any of them being distinctively privileged with concepts of the intrinsic nature of their objects or colors (of the sort that would raise an explanatory gap) would be to commit to the others not being so privileged, and the friend of inversion without illusion should reject any such asymmetry.

So the externalist must say that some quantity of explanatory gap comes from the relation of *seeing*: There is nowhere else to locate it. Fortunately, the symmetry considerations just discussed cannot be invoked to prevent doing so. Dean's condition is *worse* than that of our other subjects. He is *out of touch* with the world, the rest are *in touch*. Since all of our subjects *seem* to be in touch with the world, Dean's condition is delusive, misleading, deceptive. This breaks the symmetry. (Note the ignorance/error alternation: while there is no polarity in the many manifestations of ignorance, error is inferior to truth. Representationalist theories overlook this maneuver, assimilating the relationship between Bill and Dean to the relationship between Bill and Inez, as revealing ignorance by both rather than error by Dean.)

So the externalist needs to claim that a physical duplicate zombie of a person who is seeing would not also be seeing. Fortunately, this position is a natural one for the externalist to adopt: it is the relation of seeing, after all, that all veridical experiences have in common, and which is therefore characteristic, according to the externalist, of the distinctive sort of consciousness involved in visual experience: *consciousness of one's surroundings*. How is the externalist conceiving of *seeing* here? Evidently not as a sort of right-causal stimulation by reflected light of visual

organs: this characterization raises no explanatory gap. Rather, to put the point in terms that may ring somewhat poetical, the externalist is conceiving of seeing as a sort of *consciousness of the environment*, as a sort of *experiential visual openness* to the entities in one's surroundings, through which those surroundings become *subjectively manifest*. While this sort of openness may require a certain causal or informational substructure in order to be present in the world, the externalist should deny that this substructure suffices to generate it.

Conceiving of *seeing* as conscious visual openness to the world provides an externalist-friendly source of an explanatory gap. But this clearly does not suffice to generate all the explanatory gaps there are. After all, Greg and Bill generate distinct explanatory gaps: knowing a full physical story about the world and also knowing what it is like for Bill would not suffice to know what it is like for Greg. What generates these remaining explanatory gaps? There is nowhere else to look than *R*, *G*, and the like. One might think this establishes that these features are also experiential properties, a doctrine the externalist could accommodate by thinking of them as *qualifying* relations of seeing, so that the experiential properties are features like *R-ly seeing o's redness*. I return to the question of whether this doctrine is compulsory in section 3.5.

## 2. THE CONCEPT OF PHENOMENALITY

In this section and the next I carry out the conceptual part of my project of interpreting the notion of phenomenality. I begin with a closer look at the common notion of the phenomenal property. Two styles of explanation of this notion predominate among contemporary philosophers: One appeals to the notion of 'what it's like,' while the other has a more ostensive flavor.<sup>6</sup> In the course of discussing these notions, I explain the connection I see between phenomenality and reflective indiscriminability. I then begin the explication of my understanding of the notion of reflective indiscriminability, starting with some general framework discussion and then addressing the issue of taking up the perspective of the other. Section 3 completes the explication of reflective indiscriminability, with a discussion of the nature of reflective knowledge of experience.

Before moving into this discussion, I digress briefly into the use of 'look'-statements to characterize a notion of phenomenality (limitations of space prohibit a full treatment of this important issue). It is sometimes

6. A conception of phenomenality that few advance officially, but that often seems to be running in the background, is that for a property to be phenomenal is, by definition, for it to be a *quale*, in the sense of section 1.1. The difficulty with this conception is that it rules out as analytically false certain widely discussed theories that are advanced as theories of the phenomenal: *Sense-datum* theories are ruled out because their relations of acquaintance to sense-data are not intrinsic; *representational* theories are ruled out because their intentional properties are normatively active. If we are looking for a reasonably ecumenical conception of the phenomenal, we will have to look elsewhere.

suggested that we can nail down a notion of phenomenal character via a 'phenomenal use' of such perceptual copular verbs as 'look' and 'sound' (Jackson 1977; compare Byrne 2009). So, for example, we would say that their respective tomatoes phenomenally look red to Bill, Tina, Inez, and Ilya (or, to bring Dean in, that it phenomenally looks to each of them as if a red tomato is before them), while Greg's phenomenally looks green to him.

What might this phenomenal use amount to? It is plausible that '*o* looks *F* to *S*' has a function in reasoning or conversation roughly along the lines of 'taking a certain body of information for granted: taking up *S*'s perceptual situation: going by looking: *o* is *F*.' The reference to a given body of information (genuinely endorsed, or merely supposed) is not trivial: does the fish in the tank look to be over here, or over there (Siegel, this volume)? The answer changes based on whether the differing refractive indices of water and air are included in the body of information presupposed in the act of looking. So presumably the phenomenal use involves a certain privileged body of information.

What information would this be? Is it the information available to the subject of the 'look' attribution? It had better not be. After all, Inez will judge when looking that her tomato is green, because she is subject to no illusion. So plausibly Inez implicitly assumes that going *R* is a sign of seeing something green. So, taking Inez's beliefs and implicit assumptions for granted, her tomato looks *green* to her, rather than red, as was desired.

So we had better rule out Inez's beliefs and implicit assumptions as a candidate background to the phenomenal use. Bill's information would be a better choice: he plausibly implicitly assumes that going *R* is a sign of seeing something red; so taking Bill's information for granted, going by looking, the tomato of each subject is red, as desired. But what is special about Bill's information, which lets it set the standard for the nature of phenomenal properties? The friend of inversion without illusion should say: nothing.

So it would be a mistake to look to the phenomenal use for positive characterizations of phenomenal properties. At best we could make an arbitrary selection of a single body of information and use it to establish equivalence classes. This approach would predict that their tomatoes *phenomenally look the same* to Bill and Inez in the sense that taking Bill's (Inez's) beliefs and implicit assumptions for granted: taking up either of their perceptual situations: going by looking: the tomato is red (green). And it would predict that Greg's tomato phenomenally looks different: taking Bill's (Inez's) beliefs and implicit assumptions for granted: taking up Greg's perceptual situation: the tomato is green (red). (Or we could even take *our own* information as setting the standard of phenomenal looks.)

I have no objection to this approach; indeed, it can be thought of as a compressed and operationalized version of the account of phenomenal sameness to be developed beginning in section 2.2. However, it is important to avoid overreaching, in two ways.

First, one might think that if *o* phenomenally looks *F* to *S* that *F*-ness is, as it were, 'represented in the phenomenal content' of perception. This would be a mistake. Is redness represented in the phenomenal content of Bill's experience? Is redness represented in the phenomenal content of Inez's experience? There is no absolute answer to these questions: taking Bill's beliefs (or ours) as the standard, both are answered affirmatively; taking Inez's as the standard, both are answered negatively. Nor would it be promising to use the approach to assess whether these or those *general families* of properties are represented in the phenomenal content of perception. Are artifactual kinds represented in the phenomenal content of perception? This object is a blender: does it phenomenally look that way to me, or to a technologically unsophisticated person? That object is a really fancy scientific instrument of a certain sort: does it phenomenally look that way to me, or to the experimentalist who built it? Well, whose information is setting the standard here: the scientist's, mine, or the unsophisticated person? It is hard to see how there could be a correct answer.

Second, Chalmers advances the following case against a theory of the nature of phenomenal properties according to which they are Fregean representational features (example modified to accord with mine):

Even if Bill's and Inez's experiences are associated with distinct properties (redness and greenness), there is a strong intuitive sense in which the tomatoes look to be the same to Bill and Inez. That is, the phenomenal similarity suggests that there is a common property (intuitively, a sort of redness) such that the tomatoes look to have that property both to Bill and to Inez.

This intuitive point stands in tension with the Fregean view. . . . The Fregean view does not entail that the experiences represent a common property. In fact, it suggests that Bill's and Inez's experiences represent distinct properties, redness and greenness. (Chalmers 2006, 62)

But it does not follow from the claim that there is a property that the tomatoes phenomenally look to have both to Bill and to Inez that their experiences represent a common property. The claim that Bill's and Inez's tomatoes phenomenally look *F* to them is equivalent to the claim that taking certain information for granted: taking up either of their perceptual situations: going by looking: the tomato is *F*. Nothing in that claim involves any distinctive commitment to what the nature of a perceptual situation is, or what it is to go by looking. It requires only that what one judges going by looking is determined by one's background assumptions and the intrinsic aspects of one's perceptual condition. Since nothing about the nature of these intrinsic aspects is revealed in the surface structure of discourse about phenomenal 'looks'-statements, this is entirely compatible with nearly any claim about the relevant intrinsic aspect of Bill's and Inez's perceptual condition, including the claim that it is a Fregean representational property.

## 2.1. 'What It's Like'

Consider the claim that for *F*-ness to be among the phenomenal properties of an experience is, by definition, for it to be the case that part of what the experience is like is: *F*. For instance, for unpleasantness to be among the phenomenal properties of an experience is for it to be the case that part of what the experience is like is: unpleasant (for representative examples, see Block 1995; Chalmers 1996, 2003, 2004, 2006; Byrne 2001; Siewert 2003; Siegel 2006; Gendler and Hawthorne 2006; Thompson 2006; Bayne 2007; Speaks 2009).

The externalist should deny the success of this characterization, on the grounds that it fails to yield a notion that both is comprehensible and obviously circumscribes a space that is narrower than the externalist's experiential properties. The concern takes the form of a dilemma, depending on whether the phenomenalist intends the notion of 'what it's like' to be understood in line with the ordinary use of that notion.

If not, then a technical notion of phenomenality is being defined in terms of a technical notion of what an experience is like, so the externalist could reasonably protest a lack of comprehension.

If the ordinary use *is* intended, then the externalist should insist that no clear distinction with externalism has been drawn. After all, it is a part of ordinary practice that we answer questions about what experiences are like by appealing to external properties. For instance, if asked what my experience of writing this chapter at this moment is like, I would say things like 'as I write this chapter, I see black text appearing on my computer's screen.' One very straightforward hypothesis about the semantic function of my answer is that it serves to say that the property *being a case of seeing black text appear on screen s* is part of what my experience of writing this chapter is like. This property is an external property, so on this interpretation of my answer's semantic function, if what I say is true, external properties are part of what the experience of writing this chapter is like. So if so, the external properties are phenomenal properties (against the common understanding of phenomenality).

Is this the correct interpretation of the ordinary discourse? I don't know; in my view this discourse is not well understood (compare Hellie 2004, 2007b). Still, in order to establish an obviously nonexternalist notion of phenomenality, the phenomenalist needs to provide some reason to rule this interpretation out, and the work needed to provide such a reason has not (yet) been done.

## 2.2. Definition by Ostension

We will therefore need to look elsewhere for an elucidation of the common notion of phenomenality; the latter, ostensive, approach has much more

promise (for representative examples of this style of definition, see Chalmers 1996; Byrne and Hilbert 1997a, 1997b; Byrne 2001, 2004; Tye 2000; Lycan 2001; Thompson 2006; see also Pautz, chapter 10 this volume).

An especially judicious instance of the relevant ostension is performed by Byrne and Hilbert:

We may classify [experiences] by their phenomenology. Experiences as of red objects resemble one another in a salient phenomenological respect. In that respect, they resemble experiences as of orange objects more than they resemble experiences as of green objects. Let us say that a *red-feeling experience* is an experience of the phenomenological kind picked out by the following examples: the typical visual experiences of ripe tomatoes, rubies, blood, and so forth. (1997b, xii)

When people with normal vision look at grass, shamrocks, and jade, in daylight, . . . assuming, as we shall, that 'spectrum inversion' does not actually occur, such experiences are also phenomenologically alike: there is something obviously similar in respect of what it is like to undergo them. Let a *green-feeling experience* be an experience with this phenomenological character. (1997a, 264)

Let us consider the structure of the definition of a phenomenal property that is partly implicit in these passages.

A certain relation is ostended by consideration of pairs of experiences that are paradigms of instances and counterinstances of that relation. For Byrne and Hilbert, the relation is phenomenal sameness in respect of color phenomenology: Visual experiences of tomatoes, rubies, and blood stand in this relation, as do visual experiences of grass, shamrocks, and jade; by contrast, experiences of grass and tomatoes do not bear it to one another. But for purposes of technical convenience, I depart from the approach that Byrne and Hilbert adopt in these passages in a way that is inessential to the philosophical issues I am treating: Instead of focusing on phenomenal sameness and difference in a respect, I focus on a relation of *phenomenal exact sameness*. Experiences of tomatoes do *not* bear this relation to experiences of rubies. Rather, a famous example of a pair experiences that stand in our phenomenal exact sameness relation would be Descartes's actual experience of writing in his dressing gown and the (merely possible) dream he was calling on his reader to imagine; less famous examples would be each pair selected from Bill's, Tina's, Inez's, and Dean's experiences, discussed in section 1.4—henceforth, I will group these four experiences together as *our paradigms of phenomenal sameness*. (Strictly speaking, it is the *pairs* of these experiences that are the paradigms.) An example of a pair of experiences that do *not* stand in this relation would be that consisting my current experience as I write this chapter and the experience of a Canal Street shopkeeper haggling over the price of a belt, as well as Bill's experience and an experience that differs only in that the color of the tomato seen is a bit more orange. Consider also Greg, who sees the same green tomato as Inez, but who is normal rather than inverted: We judge that Greg's



experience differs phenomenally from those of our paradigms of phenomenal sameness.

The relation of phenomenal exact sameness (henceforth, sometimes just 'phenomenal sameness') induces a partition on experiences, which in turn projects onto the maximally specific phenomenal properties: A property *F* is a *maximally specific phenomenal property* just if for some (possible) experience *e*, all and only the experiences that are phenomenally exactly the same as *e* have *F*. Finally, the *phenomenal properties* are the (proper or improper) determinables of the maximally specific phenomenal properties, and *consciousness* is the maximally determinable phenomenal property.

Note the importance of the appeal to the determinate–determinable structure with *consciousness* as a maximally determinable element. Without this appeal, the initial ostension leaves it massively underdetermined how the extension and antiextension of 'phenomenal sameness' are to spread out beyond the initial sample of paradigms and foils.

Summarizing, we can put the definition as a Ramsification, with the following presupposition: There is a relation *S* over pairs of experiences such that

- A. (i) *S* is an equivalence relation on experiences, and (ii) the equivalence classes of *S* correspond to experiential properties; and
- B. *S* holds between *these* experiences (in particular, holds among our paradigms of phenomenal sameness, Bill's, Tina's, Inez's, and Dean's experiences) and does *not* hold between *those* experiences (in particular, fails to hold between Bill's experience and my current tomato-free experience).

Phenomenal sameness, then, is the relation *S*, if there is a unique such relation.

But we have still not clearly nailed down a reasonably determinate extension. Grant that maximally determinate phenomenal properties are determinates of *consciousness* that bear to one another the ostensibly picked out relation of phenomenal sameness. This does not yet *settle* which relation phenomenal sameness is, or which determinates of *consciousness* the maximally determinate phenomenal properties are. After all, the determinates of *consciousness* form a huge and complex structure, and a vast number of relations on this structure are such that the initial sample of paradigms instantiate them and the initial sample of foils do not. What further condition should we add to the presupposition to secure uniqueness?

Candidate approaches come in two varieties: *Metaphysical* approaches aim to spread the extension and antiextension beyond the initial sample by appeal to conditions concerning further natural structure among the experiential properties, while *epistemic* approaches aim to do so with conditions that concern the cognitive perspective taken in the course of ostending the initial sample.

### Achieving Determinacy through Metaphysics

The most straightforward added condition contributed by the metaphysical approach would look like this:

(M) *S* is an extremely natural relation.

One way to make (M) a bit more specific would assert that the maximally determinate phenomenal properties are maximally determinate experiential properties, so that the phenomenal sameness relation is the extremely natural relation of exact sameness in experiential properties. Since Bill's and Dean's experiences instantiate the phenomenal sameness relation, it follows from (M) that they bear some extremely natural relation to one another: on the more specific version of (M) under consideration, that they share exactly their experiential properties.

The externalist's view regarding this approach is as follows: The phenomenal sameness relation is not extremely natural. Bill's and Dean's experiences do not have much of significance in common; in particular, they differ in their experiential properties. Accordingly, building (M) in with (A) and (B) yields a falsehood. Hence, any concept introduced via the metaphysical approach is semantically defective: overconstrained in a way that makes no possible entity a semantic value of the concept. Since the discussion of this chapter adopts the externalist's point of view, we reject the (M)-enhanced concept.

Let us say that a concept that incorporates some of the presuppositions of a defective, overconstrained concept, but not so many as to itself be overconstrained, is a *descendant* of that concept. While the metaphysical approach does not yield a nondefective notion of phenomenal sameness, it does point the way to a *metaphysical descendant* of that concept. This descendant is introduced by hanging on to (A) and (M) while dumping the initial ostension (B). In effect, then, the metaphysical descendant is the concept of that relation that two experiences share just in case they share all their experiential properties. By the lights of the externalist, this relation is in fact extremely natural, and the maximally determinate phenomenal properties it induces are the maximal determinates of *consciousness*. So on the metaphysical descendant, the phenomenal properties are just the experiential properties. However, by the lights of the externalist, Bill's and Dean's experiences will differ in their maximally determinate phenomenal properties. Accordingly, the metaphysical descendant does not raise any further distinctive interpretive or theoretical challenges for the externalist.

### Achieving Determinacy through Epistemology

The epistemic approach, by contrast, generalizes beyond the initial sample by appeal to a certain indefinite extensibility inherent in the the epistemic capacities exercised in the course of the ostension of the sample.

The most straightforward added condition contributed by the epistemic approach would look like this:

(E) A pair of experiences instantiates *S* just if they are *reflectively indiscriminable*.<sup>7</sup>

Condition (E) is in line with the idea that tying ostension to our own abilities goes with 'knowing it when we see it': in this case, knowing whether phenomenal sameness is there when we reflect.

The remainder of this section and the next constitute a sustained development of the notion of reflective indiscriminability in such a way as to establish the adequacy for the externalist of the (E)-enriched concept as an interpretation of phenomenal sameness (one that sticks closer to the initial ostensive sample than does the metaphysical descendant). As I discuss in section 2.5, the (E)-enriched concept may well not be equivalent to the common notion of phenomenality: there may well be pairs of experiences that phenomenologists would regard as in the antiextension of phenomenal sameness, which the (E)-enriched concept does not class in its antiextension.

Still, I do not see this as a roadblock for our interpretive project. The externalist does not seem to be under any obligation to provide a perfectly accurate account of the concept of phenomenality. It suffices for the interpretive project to provide a concept that makes sense on the externalist's terms, and that is close enough in spirit to the initial concept to avoid charges of obviously changing the subject. Accordingly, I regard the (E)-enriched concept as an *epistemic descendant* of 'phenomenal sameness,' rather than as an analysis. The criterion of adequacy I adopt for an epistemic descendant is *freedom from counterparadigmatic example*, a criterion suited to the assessment of nondefective descendants and less stringent than the familiar *freedom from counterexample* suited to the assessment of analyses.

A bit more specifically, what we want out of the epistemic descendant is that it groups together experiences we would be inclined to use as paradigms and separates experiences we would be inclined to use as foils, if we were in the business of explicating the old concept of phenomenal sameness. Such clear paradigms and foils count as the *counterparadigmatic*

7. Some readers might protest that condition (E) is opposed to condition (A(i)), to the effect that *S* is an equivalence relation on experiences: after all, isn't indiscriminability a nontransitive relation? Well, yes, but the nontransitivity of a relation does not entail the nontransitivity of every subrelation of that relation.

The more directly relevant issue is whether reflective indiscriminability is nontransitive over experiences. Some authors suggest that this is an obvious part of our everyday life (Byrne and Logue 2008, §5; Stalnaker 2008, 89), but this does not show that there is no legitimate way understanding the modality (e.g., by idealizing our ordinary capacities) that can eliminate nontransitivity. Nontransitivity does not obviously infect the notion I develop over this section and the next.

I discuss the bearing of the nontransitivity of visual indiscriminability of colors to the nontransitivity of reflective indiscriminability of color experience in Hellie 2005. I leave as an exercise making that article consistent with this chapter!

candidate counterexamples. I count our paradigms of phenomenal sameness (the experiences of Bill, Tina, Inez, and Dean) as clear cases that the descendant concept must group together. By contrast, a candidate counterexample that is not counterparadigmatic is one that derives its plausibility as a counterexample only from the assumption that the relation of phenomenal sameness is an extremely natural relation, as with the (M)-enriched concept.

Note the important difference between the purpose to which I put reflective indiscriminability and that to which Martin (2004, 2006) puts it. In Martin's system, 'there is no more to the phenomenal character' of a 'causally matching' hallucination as of a white picket fence 'than that of being indiscriminable from corresponding visual perceptions of a white picket fence as what it is' (Martin 2006, 367): apparently, Martin regards indiscriminability properties as being *identical* to certain phenomenal properties. By contrast, I regard indiscriminability as merely providing a condition that any property must meet to be regarded as a phenomenal property: Phenomenal properties may well have natures that go beyond those of indiscriminability properties. The distinction is roughly analogous to the famous Kripkean distinction between 'sense-giving' (Martin's approach) and 'reference-fixing' (my approach).

### 2.3. Indiscriminability and the Phenomenal

To preview my notion of reflective indiscriminability, when 'reflective indiscriminability' is understood as in (E), I intend the claim that two experiences are reflectively indiscriminable to mean that

the following is not medically possible: one of us makes a knowledgeable reflective judgment about each of those experiences, where those judgments are *a priori* inconsistent.

Some comments about the various ingredients in this elucidation: First, following the orthodox approach to indiscriminability due to Williamson (1990), I understand indiscriminability to be unknowability of distinctness of some pair of things; for two experiences to be reflectively indiscriminable in this sense is for it to be in some sense *impossible* for them to be known to be distinct *by reflection*.<sup>8</sup> (Indiscriminability despite distinctness is an

8. Williamson's argument that discrimination is (activation of) knowledge of distinctness runs as follows: 'Discrimination is a cognitive act. . . . If we can characterize discrimination as knowledge, we shall be in a position to explain both why discrimination cannot be in error, and why the alternatives to it are ignorance and error. . . . One cannot discriminate between *a* and *a* because there can be no knowledge that *a* and *a* are distinct to be activated' (Williamson 1990, 5–7).

Heterodox theorists of indiscriminability include Fish (2009), for whom discrimination need not be a cognitive act; Raffman (2000), for whom discrimination need not be knowledgeable; and Martin (2006), for whom discrimination need not be recognition of numeric distinctness but can rather be predicative (one knows that this is not *F*). Martin's notion seems to be a legitimate extension of Williamson's; I suspect that the others may not be accounts of indiscriminability.

epistemic limitation, which can result from either ignorance or error; over the course of this chapter, both limitations will make it onto the stage.)

Second, I understand the modality in the widest sense that is compatible with assessment of the epistemic capacities of creatures recognizably like *us*, subject to certain idealizations discussed at length below; I sometimes refer to this modality as 'medical possibility.'

Third, I assume that two experiences are reflectively discriminable to someone just when reflective knowledge about each is medically available to that person, such that those pieces of knowledge are *a priori* inconsistent. My approach to reflective discrimination is therefore built on a skeleton of reflective knowledge about single experiences. I know of no developed body of theory concerning the nature of reflective discrimination of experiences; by contrast, the literature on self-knowledge of the features of single experiences is rich and extensive. These phenomena are surely not unrelated, so it would contribute to theoretical neatness if the former could be reduced to the latter.

Single-experience self-knowledge has at least the following relation to reflective discrimination: If it is *a priori*—on the basis of medically available (single-experience) reflective knowledge about each of two experiences—that those experiences are distinct, then the experiences are reflectively discriminable. Does the converse hold? A counterexample would be a case in which one just has a 'gestalt' reflective sense that two experiences are distinct, despite there being no features of those experiences that could be arrived at on the basis of reflection that one could point to which would ground this sense. But finding myself in this position, I would begin to question my assessment that the experiences were in fact distinct: The in-principle elusiveness of the distinctness-making factors would undermine my justification, and therefore the knowledgeability, of my judgment of distinctness, so there would not have been discrimination after all. At any rate, I assume that the converse *does* hold.

The condition of apriority reflects the nature of the procedure for ostending phenomenal sameness, as I understand it, on which we simply ignore all other nonreflective evidence about the experiences at issue when introducing them as paradigms or foils of phenomenal sameness. We know, for instance, that Bill's and Dean's experiences take place in different people and have different veridicality statuses, but this does not bear at all on whether they count as paradigms of phenomenal sameness. All we have to go on is the evidence of reflection, plus general-purpose rationality.

Fourth, my claim that *we* are the ones whose discriminations are relevant is in reaction to intensive recent investigation of reflective indiscriminability (see, e.g., Siegel 2004, 2008; Martin 2004, 2006; Hawthorne and Kovakovich 2006; Sturgeon 2008; Fish 2009).<sup>9</sup> This discussion points the way to a cluster of complaints about the adequacy of the (E)-enriched concept as a

9. For a very early discussion of the relevance of varying powers of discrimination of different subjects to disjunctive theories of perception, see Hinckfuss 1970, 279.

descendant of ‘phenomenal sameness’: Surely a dog (or baby or inattentive person) can have experiences with phenomenal characters more finely grained than their capacities to knowingly judge experiences distinct; conversely, perhaps an extremely attentive person would be able to discriminate experiences we would regard as phenomenally the same. The notion of *reflective* discrimination of *two distinct* experiences doesn’t even make sense, since one can only reflect on a single experience at once, namely, the experience one is undergoing (compare Byrne and Logue 2008, §5).

I argue that all of these concerns can be met with a single maneuver, once it is observed that ostension of phenomenal sameness is always from our own case, taking a sort of ‘pseudo-first-person perspective’ on someone else’s experience, a perspective from within which one’s own reflective capacities govern the limits of discriminability. Developing these ideas is the work of the next subsection.

Fifth, even granting the success of this approach, one might wonder whether the phenomenal properties of one’s own experience can outrun one’s capacities to know them reflectively. Assessing this concern requires a theory of reflection and of the dimensions along which it can be idealized. Providing such a theory is the work of section 3.

At the end of this discussion, quite a fair bit of apparatus will have accreted: idealizations of our psychology as it pertains to self-knowledge that the skeptical reader might greet with incredulity. My response to this reader is that my aim here is to articulate the most charitable possible interpretation the externalist could give of the phenomenalist. It is not clear how to make sense of the phenomenalist position without this apparatus, so the skeptical reader should reject the coherence of the notion of phenomenality. But this would leave the externalist conception of consciousness as the last position standing; accordingly, I count this skeptical reader as an ally on the deeper questions.

#### 2.4. The Pseudo-First-Person Perspective

I now address the question of what sense can be made of someone discriminating a pair of distinct experiences—one of which is, *a fortiori*, not present to the subject during the act of discrimination—*reflectively*. My take on this question also turns out to dissolve the concern about subjects more and less in a position to acquire reflective knowledge than we are.

Assuming that we have in hand some sense in which *we* can reflect on *Bill’s* experience, the motivation for taking *us* to be the discriminating subjects referred to in condition (E) is clear. After all, *we* perform the ostension involved in the definition of ‘phenomenal sameness.’ Other subjects (e.g., the subjects involved in the paradigms of phenomenal sameness, or some children or dogs, or some zombies, or some idealized version of such subjects, or God)<sup>10</sup> might be more or less accurate and/or

10. References to God should not be understood as presupposing theistic commitments by the author.



complete in their reflective judgments than we are. If so, if such a subject were to set up for itself a concept that functions similarly to our concept of phenomenal sameness, it would correspondingly start out with a smaller or larger set of paradigms. So keying our epistemic descendant of 'phenomenal sameness' to the discriminative capacities of other subjects would generate a risk of counterparadigmatic example; by contrast, keying it to our capacities removes this risk.

Now to the question of what it is for one to reflect on the experience of another. It seems that there is a sort of 'pseudo-first-person perspective' one can take toward an experience one is not currently undergoing that is relevantly like actually undergoing the experience, such that one may then acquire reflective knowledge about that experience. This perspective is the attitude of 'taking up the point of view' of or, as I will say, *projecting* that experience. With respect to our paradigms of phenomenal sameness, one takes up the attitude of projection toward *merely possible* experiences, but one may also project the experiences of actual other subjects or of one's own past experiences.

Projection is surely an ability we have and regularly use whenever we recollect our own past experiences, 'put ourselves in the shoes of' other people, or imagine merely possible experiences (in each case, from the inside). And certainly the acts of ostension that get the concept of phenomenal sameness off the ground in the first place rely on some sort of projection: When we hear the stories about Bill, Tina, Inez, and Dean, we understand these stories by putting ourselves in their shoes and simulating their experiences from the inside. The phenomenalist would undermine the foundation of their position by objecting too strongly to the notion of projection.

A few remarks about the psychology of projection should forestall some confusions and highlight a few points at which I idealize. After that, I argue that there is no relevant difference for present purposes (given these idealizations) between projecting an experience and undergoing it.

Although one might tend to think of projecting an experience as 'imagining' the experience, thinking of projection as a kind of imagining can be misleading. Since projection is targeted at knowledge of some condition that is independent of the act of projection (the experiential nature of some other experience), projecting can be done correctly or incorrectly: The internal representation might be faithful or unfaithful to the character of the experience one was trying to project. For instance, George sees a red thing but someone has told Francine he sees a green thing: She projects George's experience but does so incorrectly (and, accordingly, she does not know what George's experience is like/which experiential properties it has). The experiential properties George's experience is represented as having by Francine's internal representation diverge from the properties George's experience actually has, so in this sense Francine's act of projection is not accurate to George's actual experience. Accordingly, projection is more closely analogous to perception than to the free, rationally unconstrained faculty of visual imaging.

Of course, perception and projection are not alike in all respects. One does not literally see the mind of the other, in the way one sees a bucket of pig food before one, in at least the following sense: Causally implicated in any act of projection, there is always another conscious mental representation (this representation could well be what people are thinking of when assimilating projection to ‘imagination’). The ‘input’ to the projective faculty is (or can be) a description of someone’s situation in entirely objective terms of the sorts we offered in characterizing our various paradigms;<sup>11</sup> the ‘output’ is a representation of that person’s experience, ‘from the inside,’ as having certain experiential properties. It is the tokening of this representation that I am thinking of as the act of projection. Plausibly, only experiential properties are attributed to an experience in an act of projection: Although one might attribute nonexperiential properties to the experience during the act of projection, this should not be regarded as *part* of the act of projection.

(Despite this disanalogy between perception and projection, there may still be room for raising a debate about theories of projection analogous to the traditional debates in philosophy of perception. Can projection intelligibly be regarded as a *sui generis* form of ‘mind reading,’ in which in the ideal case one comes into a sort of acquaintance with the mind of the other? Such a view would be analogous to a direct realist theory of visual perception. Although my sympathies are with this acquaintance-like story about projection, present purposes do not require taking a stand on this issue.)

Let us say that *ideal* projection of *e* is projection done with total relevant knowledge about *e* as input, and applying whatever method is involved in projection perfectly on the basis of this knowledge. On pain of skepticism about our knowledge of other minds, we should deny that ideal projection can lead us into error; accordingly, if we focus only on ideal projection, we can assume that an act of projection does not misrepresent the projected experience.

The converse of accuracy is determinacy: I idealize here as well. It is clear that the representation of an experience generated in any act of projection is, in all realistic cases, far less determinate than any actual experience. Still, when setting up cases for the purpose of ostensibly defining phenomenal sameness, the context is such as to implicitly treat points of explicit indeterminacy as points of experiential sameness. (It would not be to the point to complain that Bill’s and Tina’s experiences may not in fact be phenomenally the same, since we have not been told whether Tina’s tomato is sitting on a plate while Bill’s is in a dish.) Accordingly, it will be harmless to assume that an act of projection does not fail to represent any experiential features of its target.<sup>12</sup> Putting this pair of

11. Compare Siegel (2008, 212) on the direct realist’s need to characterize the ‘situation’ of the other in nonphenomenal terms.

12. This decision is a result of an earlier choice, namely, to regard the explanandum to be *exact* phenomenal sameness, rather than phenomenal sameness in some respect. The rejected approach would have allowed us to treat projection in a respect at this stage but would have also considerably complicated the relationship between the relation on experiences and the determinate–determinable structure of *consciousness*.

idealizations together, I henceforth ignore cases in which a projected experience has the experiential property *F* but is not projected as *F*, as well as cases in which an experience is projected as *F* but is not *F*.

To project an experience is not yet to reflect on it. To conjure up a sympathetic sense for what it is like to be Bill is not yet to make any judgment about what it is like to be Bill, any more than to see a red thing is to judge that this is what it is like to see a red thing. Still, it seems that when one projects an experience, one can reflect on it and judge how it is, in much the same way as one does when one reflects and judges how an occurrent experience is. Absent this assumption, it is difficult to see how the method of hypothetical example used in the philosophy of perceptual consciousness could get any purchase with the heart of the matter, namely, the natures of our occurrent perceptual experiences.

(It is important to distinguish between reflecting on *an experience of projecting* and reflecting on *a projected experience*: The difference is akin to that between mention and use. E.g., Bill sees a red thing, and so do I. Our experiences share a large number of their experiential properties. Accordingly, I can apply reflective concepts to either my *undergone* experience or Bill's *projected* experience, and this range of concepts will be very similar. But my experience of projecting Bill's experience is also an experience. This experience is very different in its experiential properties from my experience of seeing a red thing. It follows that the reflective concepts I apply to Bill's projected experience, and to my experience of projecting Bill's experience, will also be very different. I will always intend the former, never the latter.)

The upshot of this discussion, then, is the following: One's epistemic position with respect to acquiring reflective knowledge about an experience with certain experiential properties that one *ideally projects* is identical to one's epistemic position with respect to acquiring reflective knowledge about an experience with certain experiential properties that one *undergoes*. If this is correct, then we are free, henceforth, to suppress consideration of whether an experience is ideally projected or undergone.

## 2.5. Aliens

A straightforward objection to the view that if two experiences—sequential experiences, let us suppose—are reflectively indiscriminable *by their subject*, they share their phenomenal properties, concerns subjects not in a position to reflect on their experiences (dogs or children, perhaps): All their experiences are indiscriminable to them, but, intuitively, they might nonetheless differ in phenomenal properties. This specific difficulty does not affect our analysis, but it is an instance of a more general issue, namely, the status of experiences that cannot be reflected upon by the subjects at issue in the analysis, whoever they might be: An indiscriminability analysis

of phenomenal sameness would class any pair of such experiences as phenomenally the same.

The form of this issue for our analysis concerns experiences that we are not in a position to project ourselves into: If there are such experiences, we cannot discriminate them; hence, the analysis classes them as phenomenally the same. Could there be such ‘alien’ experiences? Nagel (1974) famously suggested that bat experiences might be unprojectable.

Still, even supposing that Nagel is correct here and that there are experiences we cannot project, nothing follows about the adequacy of the proposed epistemic descendant of ‘phenomenal sameness.’ After all, no such experience could be a member of a *paradigm* of phenomenal sameness. Presenting a paradigm requires reflection on the experiences from the inside, and if the discussion of the previous subsection is correct, this requires projection. Accordingly, no unprojectable experience could be a member of a *counterparadigmatic* counterexample to the epistemic descendant; hence, this issue does not threaten my aim.

The prediction that any pair of experiences including an unprojectable experience is phenomenally the same grates on the ear. Accordingly, I henceforth regard the domain to which the descendant concept is intended to apply as including only pairs of projectable experiences: A pair of experiences containing an alien is neither phenomenally the same nor phenomenally distinct.

### 3. TRANSPARENCY-BASED ACCOUNTS OF REFLECTIVE KNOWLEDGE

In this section, I elaborate two accounts of reflective knowledge of the nature of one’s perceptual state, in the context of the assumption of ‘the transparency of experience.’ In section 3.1, I expand upon the notion of transparency and introduce the first account. This is a sort of ‘displaced perception’ account, according to which, roughly, one comes to know about one’s perceptual state by using the state in making a judgment about what one sees while also self-ascribing seeing these things.<sup>13</sup> Sections 3.2 and 3.3 develop technical background for expanding further upon this account: after reviewing the ‘two-dimensionalist’ account of the senses of concepts, I characterize senses of perception-based color concepts, and then explain how one’s self-ascriptive judgment is entailed *a priori* by an ordinary perceptual judgment about the world. It follows that if one can have perceptual knowledge, one can have reflective knowledge by this route. Section 3.4 employs this apparatus to explain how Bill’s,

13. Precursors to the first account are developed by Evans (1982), Dretske (1995), Martin (1998), Byrne (2001, 2005), and especially Kennedy (2009, 2010). Special thanks to Matt Kennedy for discussion. Precursors to the second account are developed by Shoemaker (1991) and Loar (2003a).

Tina's, and Inez's experiences are indiscriminable from one another, but discriminable from Greg's experience.

Reflection of this sort cannot, however, provide knowledge if one *cannot* have perceptual knowledge: if one is suffering illusion or hallucination, for example. So this account cannot explain how Ilya's and Dean's experiences are discriminable from Greg's experience. In section 3.5, I develop a second theory of reflection to remedy this gap, involving a style of thinking I call 'becoming alienated.' When one becomes alienated, one is in a position to directly refer to properties such as *R* and *G*. Section 3.6 draws a range of revisionary morals from this discussion: for example, that transparency is compatible with reflective knowledge of intrinsic qualities of experience and that phenomenological study cannot constrain theories of the nature of these qualities.

### 3.1. Transparency and Displaced Judgments

The core of the displaced perception theory is that reflective knowledge of experiential properties (of visual experiences) involves concepts that themselves incorporate *demonstrative concepts* based in visual experiences. So, for instance, Bill's visual experience makes available the object-demonstrative concept 'that object,' referring to tomato *t*, and the color-demonstrative concept 'that color,' referring to the property of redness. These concepts have their referents only thanks to acts of attention, of course: If Bill were attending to some other object he saw, or not attending at all but blindly thinking 'that object,' the concept would not refer to *t* (but rather to some other object or to no object at all).

The concepts employed in reflective knowledge of experience are concepts such as 'seeing that object' and 'seeing that color' (or, if you will, 'having *seeing that object/that color* as an experiential property'). The apparent logical syntax of these *displaced concepts* is genuine, in that the demonstrative concepts that appear to be constituents are in fact constituents; moreover, the meaning of a displaced concept reflects the meaning of the perceptual demonstrative embedded within it.

A *displaced judgment*, then, is one that predicates a displaced concept of a demonstrative term referring to an experience (e.g., 'this experience'), where both the subject and predicate concepts are deployed 'from the same perspective' (in particular, so that the displaced concept contains demonstrative concepts made available by the demonstrated experience). In accord with the conclusion of section 2.4, I do not distinguish between cases in which the perspective is the true first-person perspective on a current experience, and those in which the perspective is the pseudo-first-person perspective on an ideally projected experience. I use expressions like 'Bill's judgment' to encompass our judgment, ideally projecting Bill's experience. An example of a displaced judgment is 'this experience [Bill's experience] has *seeing this color* [redness, which is of course seen in Bill's experience] as an experiential property.'



The key thesis of my displaced perception approach, then, is that *any medically possible source of reflective knowledge of which experiential properties one's experience has is based on the displaced judgment*. (And, conversely, for any property a projected subject sees, a displaced judgment concerning that property is possible for us.) For the moment, I ignore the involvement of 'basing,' proceeding as if displaced judgments are the only medically possible source of such reflective knowledge; the notion of basing becomes relevant starting in section 3.5.

This key thesis is attractive for several reasons. First, it is in line with the so-called transparency of experience. This phenomenon has been explicated in scores of ways, but a central take on the phenomenon is this: Most reflective knowledge of which experiential property one's experience has requires attention to entities—objects or property tokens—in one's environment.<sup>14</sup> The connection should be obvious: Perceptual attention to entities in one's environment is required for deployment of displaced concepts. Moreover, little else is: only conceptualization of this attention and conceptual material required for self-attribution of visual openness to the entities. This requirement of minimal added cognitive exercise seems to be implicit in much discussion of transparency: The mere requirement of *some element* of exterior attention seems compatible with a great range of follow-on exercises that would violate the spirit of transparency.

Chalmers (2003) advances a view of reflective knowledge that contrasts with mine in rejecting the key thesis. According to Chalmers, we are able to grasp concepts that fully reveal the natures of phenomenal properties by directing attention to these properties; doing so does not involve any attention to entities in one's environment. This view is not compatible with transparency as I have characterized it. (Chalmers's

14. This transparency thesis is sometimes criticized on the grounds that while we can acquire reflective knowledge of hallucinatory perceptual consciousness, in hallucination no attention to external entities is possible (Siewert 2004; Crane 2006). The obvious reply is to deny that total reflective knowledge of hallucinatory perceptual consciousness is possible (Soteriou 2005). As I discuss below, knowledge of the phenomenal properties of hallucinatory consciousness is possible, but the nonphenomenal properties may be unknowable: hence 'most.' No requirement stronger than 'most' reflective knowledge would be plausible anyway, due to the direct knowability of blurriness, noisiness, double vision, and other visual effects. Finally, we shall see that reflective knowledge of hallucination still involves a sort of apparent attention to external entities, which may be sufficient to undergrid the allure of the unqualified transparency thesis.

There are of course alternative theories of reflective judgment that remain firmly in the displaced perception camp: for instance, one might suggest that certain medically possible reflective judgments are along the lines of 'this experience has *visually representing that something has this color* as an experiential property.' If the aim of this is to provide support for a representationalist view on which veridical and hallucinatory experience can share experiential properties, then the demonstrative cannot (in the hallucinatory case) refer to a property-instance in the environment, but must rather refer to a property-type in Platonic Heaven. This seems to run afoul of the transparency of experience. In my view, we do not understand this notion of visual representation: the notion of seeing, by contrast, is part of our common psychological lexicon. Finally, representationalism implausibly assimilates both error and truth to varieties of mutual ignorance.



later remarks (2004, 176) suggest that this view has been modified, while leaving less than fully explicit what the view is to be replaced with.)

Second, the key thesis is in line with a famous Humean observation about the psychological untenability of external-world skepticism:

It seems evident, that men are carried, by a natural instinct or prepossession, to repose faith in their senses; and that, without any reasoning, or even almost before the use of reason, we always suppose an external universe, which depends not on our perception. . . . Even the animal creation are governed by a like opinion, and preserve this belief of external objects, in all their thoughts, designs, and actions. (Hume 1777, XII/I)

A displaced judgment about one's experience presupposes that one sees. If we are medically unable to reflect on the nature of our experience without entertaining displaced judgments at some level, then reflection on our perceptual situation in regard to the external world begins in an antiskeptical stance. I argue in section 3.5 that we can reflectively conceive of ourselves as inhabiting skeptical scenarios; still, doing so is a highly theoretical cognitive achievement, and in any event requires a sort of double consciousness: we cannot entertain the sophisticated skeptical hypothesis without at some level simultaneously affirming an ordinary antiskeptical stance; in kicking the ladder away, we remain standing on it.

Third, the theory is *austere*. It builds our capacity for first-person reflection entirely out of capacities that we obviously already have. Unlike Chalmers's view, mine does not require any further distinctive capacity of attention to experience. This prediction is attractive: as Byrne (2009, 434) remarks, 'cognitive scientists have distinguished many different kinds of attention, but have not yet seen the need to suppose that we can attend to our experiences.'

A final reason to accept the key thesis is especially attractive to friends of externalism. Support for that view and its relatives often takes a *phenomenological* form (compare Hellie 2007a; Kennedy 2009): reflection on experience makes manifest a perceptual connection to external entities, and it suffices for *F* to count as an experiential property of *e* that *e*'s *F*-ness be made manifest in reflection. The key thesis requires reflective judgment to begin in self-ascription of properties of the form *seeing this object/this color*, self-ascription which is accurate when we see without illusion. The key thesis therefore ratifies the phenomenological judgment brought in support of externalism.

### 3.2. Epistemic Two-Dimensional Semantics

From the key thesis, it follows that two experiences are reflectively indiscriminable just when there is a (knowledgeable) displaced judgment about one which is *a priori* inconsistent with a (knowledgeable) displaced judgment about the other. But to operationalize this thesis, we need a sense for when two judgments are *a priori* inconsistent. I propose to

exploit the technical resources of *epistemic two-dimensional semantics* (Chalmers 2005) for this purpose.

On the epistemic two-dimensional approach, the epistemic properties of a judgment are reflected in a set of *centered worlds* assigned to it (where a centered world is a pair of a possible world and a time slice of a subject in that world). A judgment is true *a priori* just when every centered world is assigned to it, false *a priori* just when no centered world is assigned to it, and *a posteriori* just when some but not all centered worlds are assigned to it. A conjunction is *a priori* false just when no centered world is assigned to both conjuncts; two judgments are *a priori* inconsistent just when no centered world is assigned to both.

Considerable artistry is involved in stating the assignment of centered worlds to judgments; limitations of space prevent addressing the many nuances here. Put crudely, on the approach I am exploiting, this assignment is supposed to reflect the *Fregean sense* of the judgment, where this is understood more-or-less as determined syntactically by *reference-fixing descriptive conditions* competent users implicitly associate with the constituents of the judgment.

Working through an example will be of pedagogical value. Consider the judgment that Hesperus is Phosphorus. Here the ‘constituents’ are ‘Hesperus,’ ‘Phosphorus,’ and the identity predicate. A plausible reference-fixing description for ‘Hesperus’ is *being the first heavenly body visible in the evening from locations near the center*, while one for ‘Phosphorus’ is *being the last heavenly body visible in the morning from locations near the center*; we may assume that the referent of the identity predicate is transparently understood without appeal to a reference-fixing description. Accordingly, a centered world is assigned to the judgment just when, from the perspective of the center, there is a unique heavenly body visible both last in the evening and first in the morning from locations near it. Some centered world does, in fact, satisfy this condition; for instance, the actual world centered on Hammurabi or Kepler, since the planet Venus is just such a heavenly body. But there are also centered worlds that fail to satisfy the condition. For example, it is plausible that there are planets in the actual world containing beings that see *distinct* heavenly bodies first in the evening and last in the morning. An example of a falsifying centered world, then, is the actual world centered on one of those beings. Since some worlds satisfy the condition and others do not, the judgment is *a posteriori*. (By contrast, a judgment explicitly identifying Hesperus with the unique satisfier of the reference-fixing description associated with ‘Hesperus’—such as ‘Hesperus is the first heavenly body visible in the evening’—will be true at all centered worlds, hence *a priori*.)

### 3.3. The Senses of Displaced Concepts

This provides enough of a grasp of the technical apparatus of two-dimensionalism for the purpose of explaining the notion of reflective

indiscriminability as amplified by the displaced perception theory. What is needed to apply the technical apparatus is a sense of the proper reference-fixing descriptions for displaced concepts like 'seeing that object' and 'seeing something with that color.'

Since the meaning of such a concept incorporates the meaning of the perceptual object- or color-demonstrative concept 'that object' or 'that color' embedded within it, it makes sense to proceed by stating senses for such object- and color-demonstrative concepts.

It is probably not too far off to assume that the sense of the object-demonstrative concept 'that object' is something like 'the first object in the direction of my gaze' or 'the object that is right-causing the current visual experience by being seen'; or, incorporating a bit more nuance to accommodate cases of seeing multiple objects, 'the object that is right-causing a certain privileged part of the current visual experience by being seen' (where the privilege in question results from an underlying act of selective attention, characterized perhaps in functional or neurological terms). The true theory of visual demonstration of objects may well be significantly more nuanced still (Dickie 2010). Still, as I discuss below, senses of the sort just described have the merit of generating the correct predictions about reflective discriminability of experiences (e.g., Bill's and Tina's) that differ merely haecceitistically, and are therefore at least close enough to correct for present purposes; it is doubtful that any more nuanced theory would differ in the respects that generate this merit.

As for color demonstratives, one option would be to straightforwardly roll through the senses just characterized for object demonstratives, along the following lines: 'the color in the direction of my gaze'/'the color that is right-causing the current visual experience.' However, such a 'blind' demonstrative does not seem sufficiently to track the subject's internal condition to count as appropriately perceptual. Suppose that a normal subject sees a red thing on the left and a green thing on the right, and judges 'that color [gazing left] = that color [gazing right],' where these are perceptual demonstrative concepts. Intuitively, assuming it to be known that a color is seen with both gazes, the judgment is *a priori* false, but the 'blind' demonstrative theory at issue cannot explain this: The judgment would be true at a centered world just if the color in the direction of the first gaze by the subject at the center = the color in the direction of the second gaze by the subject at the center, and there are such worlds. More important, blind demonstratives could not serve in a displaced perception theory of reflective discrimination, since the result would make the discriminability of Bill's and Greg's experiences inexplicable (Bill, recall, sees a red tomato, while Greg sees the same green tomato as Inez). I sketch the fairly straightforward reason (the discussion below provides resources that the reader may use to fill in the details here): Nothing would distinguish the senses of the displaced concepts incorporating them, so that there could be no *a priori* inconsistency in a single experience satisfying each.

A superior theory of perceptual color demonstratives builds in the subject's internal condition into the reference-fixing description (theories resembling the following are discussed in Peacocke 1983; Chalmers 2004; Thompson 2009). Recall the assumption from section 1.4 that when a normal subject sees a red (green) thing under normal conditions, there is some internal property  $R$  ( $G$ ) such that this puts the subject into condition  $R$  ( $G$ ). Consider, then, the following reference-fixing condition associated with perceptual demonstration by a normal subject of redness, 'the color instances of which when seen typically cause, and an instance of which is now thereby causing, me to instantiate  $R$ '; of greenness, 'the color instances of which when seen typically cause, and an instance of which is now causing, me to instantiate  $G$ .'

This theory of the senses of color demonstratives predicts the doctrine from section 1.4 that color concepts do not generate an explanatory gap. If a color concept refers only by description, then if we see that a certain property satisfies the description, we no longer ask how that property could be the color.

This theory of the senses of color demonstratives predicts the *a priori* falsity of the judgment 'this color [gazing left] = this color [gazing right]' described in the preceding paragraph. Ignoring the temporal difference between the two demonstrations, a centered world would satisfy the sense of the judgment just if, at the center,  $R$  and  $G$  are instantiated and caused to be so by colors that, when seen, typically have this effect, in the subject at the center. But this requires  $R$  and  $G$  to be simultaneously instantiated at the center. Let us suppose that  $R$  involves the opponency channels concerning color experience at a certain region of the visual field firing in one way, while  $G$  involves them firing in an incompatible way. Accordingly, it is impossible for  $R$  and  $G$  to be mutually and simultaneously instantiated. Therefore, there is no such centered world, so the judgment is false *a priori*.

Why build in the conditions *both* of typical causation and current causation? The need doesn't appear in the case of veridical perception because the current and typical causation requirements come apart only when the current case is atypical, and hence nonveridical. But the requirements come apart in cases of illusion. Consider Ilya's experience of seeing a white thing under a red spotlight. Ilya's instantiating  $R$  is currently being caused by seeing white, so red does not satisfy the description, and seeing white does not typically cause Ilya to instantiate  $R$ , so white does not satisfy the description. So there is no perfect satisfier of the description. White and red both satisfy the description better than any other property, however, so plausibly there is indeterminate reference between white and red. This strikes me as a credible prediction about such an illusory demonstrative: Intuitively, I do not see how to settle the question whether Ilya's 'this color' concerns white or red.

Completing our theory of the senses of displaced concepts like 'seeing this object/this color' requires a story about what 'seeing' contributes to the sense. As discussed in section 1.4, it is in the spirit of externalism to think of 'seeing' as an irreducible notion along the lines of 'conscious visual openness to the world' which could not apply to any 'zombie' or nonconscious physical duplicate of a conscious being, and which therefore generates an explanatory gap. If so, the nature of the property itself is revealed in its reference-fixing condition (by contrast, we are assuming color concepts to be natural kind concepts, such that the identity of the referent is not known *a priori* by competent users of the concept). If this is correct, no further analysis is required here.

Assembling this claim with our results about the senses of perceptual demonstratives, we conclude that the sense of 'seeing this object' is 'seeing the object that is right-causing the current visual experience,' while the sense of 'seeing this color' is of the form 'seeing the color instances of which when see typically cause, and an instance of which is now thereby causing, me to instantiate *R*' (varying *R* as appropriate).

We can see that the displaced judgement 'my current experience is a case of seeing something this color' is entailed *a priori* by Bill's ordinary color demonstrative judgement 'this object is this color [namely, red].' After all, the sense of the latter requires the subject at the center to be such that the object causing it by being seen has the color which is now causing and typically causes the subject to be *R* when seeing it, and the sense of the former requires the subject at the center to be experiencing so as to see something with the color which now and typically causes the subject to be *R* when seeing it. It is clear that the former requirement cannot be met unless the latter is. The displaced judgement is in this sense implicit in the ordinary demonstrative judgement. Using perceptual experience in even the most elemental judgements carries with it an element of self-knowledge.

Now that we have seen the account of the senses of displaced concepts, we should assess whether it genuinely respects the transparency of experience. One might be skeptical about its transparent credentials, due to the involvement in the senses in question of conditions concerning *R*, *G*, visual experiences, and other assorted internal conditions and entities. However, as framed, transparency concerns the *referential* properties of our acts of attention. We attend *to* instances of colors and objects, *with* internal properties, *along channels* opened by conscious visual openness to the world. On my displaced perception theory, there is no *reference* in our reflective judgments to any purely internal properties or conditions. The concepts in question solely concern external properties such as redness and harmless relations such as seeing. The internal conditions are solely mentioned in theories of the sense of our reflective concepts. Certainly, in some cases reference-fixing conditions are explicitly grasped. Consider the explicit introduction of a name as referring rigidly to the inventor of the zipper: One who grasps the sense of this name would plausibly need



to be consciously aware of it as referring to the inventor of the zipper. But it does not follow that this is so in all cases. In more typical cases, senses are best regarded as abstractions from tacitly understood rules that guide application of concepts. My claim about the sense of ‘that color [red]’ is intended solely to mean that one applies that concept as if guided by a rule requiring that seeing it now and typically causes, one to instantiate *R*. Acting in accord with the rule does not require any particularly explicit access to that rule. But if not, then the account does not require that when a subject employs a color demonstrative, the subject needs to have any particularly explicit awareness of *R*: *R* needs merely to causally influence the subject’s application of the demonstrative.

Before advancing to the next stage in the case for our epistemic descendant, let me pause for a brief aside concerning the broad parameters of the total theory I have developed. In certain respects, the view is Cartesian; in certain other respects, the view is Kantian. On a Cartesian view, perception cannot mislead, and empirical error is always the fault of overreaching thought; the theory here says much the same. Taking his current experience to be a typical case, Ilya judges his tomato to be red and his experience to present red; and yet his experience presents only the whiteness of the tomato. On a Kantian view, perceptual experience consists of a brute relation to an unconceptualized ‘given,’ and both empirical thought and thought about experience require bringing this condition of givenness under concepts. On the theory here, the given in question is not a classical internal sense-datum, but an external entity (exception: I take no stand on whether a hallucination requires givenness of something internal, some other condition, or nothing at all).

### 3.4. Displaced Discrimination of Veridical Experiences

I am now in a position to derive the observed patterns of discriminability and indiscriminability among veridical paradigms and antiparadigms of phenomenal sameness. The next subsection discusses illusion and hallucination.

What epistemic status does our displaced perception account assign to a reflective judgment about a veridical experience? Suppose that we judge on the basis of projection into Bill’s experience ‘this experience is a case of seeing this object.’ The former demonstrative refers to Bill’s experience, while the latter refers to tomato *t*; since Bill’s experience is a case of seeing *t*, the judgment is true. Is the judgment *knowledge*? Conclusively answering this question would require a nontrivial sufficient condition for knowledge. Still, supposing that Bill is not at risk of hallucinating or of having an indiscriminable tomato swapped for *t*, the judgment will be ‘safe’ (Williamson 2000); plausibly, safe true belief is a very good approximation to knowledge (the appeal to safety in this context follows Kennedy 2009).



(Is it plausible to suppose that Bill cannot *know* the experiential properties of his experience if he is at risk of hallucinating? This is a fairly straightforward prediction of externalism combined with the safety requirement on knowledge: The latter predicts that the subject at risk of hallucinating cannot know that this tomato is this color; if the subject's experiential properties include the tomato's thus-coloredness, an implausible denial of a 'closure' principle on knowledge would be required to accept that the subject can know that he has the experiential property. As far as the reflective discriminability of the risky experience from others is concerned, that matter should be assimilated to the discriminability of hallucinations from other experiences.)

And suppose that we judge on the basis of projection into Bill's experience 'this experience is a case of seeing this color.' Here the demonstratives refer to Bill's experience (once again) and to redness. Since Bill's experience is a case of seeing redness, the judgment is once again true. Supposing once again that Bill is not at risk of hallucination or illusion, the judgment is once again safe.

Parallel claims are true of our remaining veridical paradigms and anti-paradigms (Tina, Inez, and Greg).

Now let me develop the predictions made by this displaced perception theory for the discriminability status of our veridical paradigms and antiparadigms. Consider first Bill's and Tina's normal experiences of seeing visually identical but numerically distinct tomatoes. Intuitively, these experiences are reflectively indiscriminable (and they are paradigmatically phenomenally the same). The account predicts that this will be so just in case we can make knowledgeable reflective judgments on the basis of those experiences that are *a priori* inconsistent; and we cannot. Plausibly, *any* pair of displaced object judgments are *a priori* consistent. The sense of Bill's displaced object judgment is something like 'the experience at the center is a case of seeing the object that is right-causing it.' Indeed, this is the sense of any displaced object judgment, including Tina's. Since Bill's and Tina's senses do not differ, it is of course possible that these senses be mutually satisfied by a centered world. Accordingly, there is no *a priori* inconsistency between Bill's and Tina's displaced object judgments.

Accordingly, the displaced perception account predicts immediately that mere haecceitistic differences between experiences will never make for introspective discriminability.

Note also that Bill's and Tina's displaced *color* judgments will be *a priori* consistent. The sense of Bill's judgment is something like 'the experience at the center is a case of seeing the color instances of which typically cause, and an instance of which is now causing, the subject at the center to instantiate *R*.' This is also the sense of Tina's displaced color judgment. Since the senses of the two judgments are the same, they are of course *a priori* consistent.

Accordingly, the displaced perception account predicts that there will be no medically possible reflective judgments on the basis of Tina's experience that are *a priori* inconsistent with any medically possible reflective judgment on the basis of Bill's experience. Accordingly, the experiences are reflectively indiscriminable.

The explanation of the indiscriminability of Bill's experience of a red tomato and inverted Inez's experience of a green tomato is much the same. Once again, no displaced object judgments can discriminate them. And since Inez is inverted, when she is seeing a green thing, the sense of her displaced color judgment will incorporate *R*. Accordingly, the sense of her displaced color judgment will be the same as that of Bill's displaced color judgment. Accordingly, they too will be *a priori* consistent.

The displaced perception account additionally predicts immediately that mere *quidditistic* differences between experiences will never make for introspective discriminability.

Now consider Bill's experience in relation to Greg's experience of seeing a green thing: Intuitively, the two are reflectively discriminable, and we have advanced this pair as an antiparadigm of phenomenal sameness. Since Greg is a normal subject, when he sees a green thing, the sense of his displaced color judgment will be along the lines of 'the experience at the center is a case of seeing the color instances of which when seen typically cause, and an instance of which is now causing, the subject at the center to instantiate *G*'; and recall that the sense of Bill's displaced color judgment is 'the experience at the center is a case of seeing the color instances of which when seen typically cause, and an instance of which is now causing, the subject at the center to instantiate *R*.' For reasons I have already discussed, these senses cannot be mutually satisfied: This would require the subject at the center to instantiate both *G* and *R*, which is impossible.<sup>15</sup>

15. It is worth briefly contrasting this discussion with the predictions we would see on the basis of an account of reflective judgment like that of Chalmers (2003), according to which one conceptualizes one's phenomenal property *F* with *a priori* insight into its nature when one 'attends' to *F* or one of its determinates while entertaining a predicate that *F* alone among these determinates satisfies.

First, the Chalmers framework makes room, salutarily, for both ignorance and error. Ignorance is accommodated by an incapacity to attend in the right way (245). Error is accommodated by mismatch between the predicate and the property to which one attends: the conceptualized property is roughly the property satisfying the predicate which is most similar to the object of attention (237–8). This framework is not wholly amenable to our purposes. Accommodating the indiscriminability of experiences that differ merely haecceitistically or quidditistically would require imposing in an ad hoc manner restrictions on the reach of attention to exclude haecceitistic and (some, but not all) quidditistic properties of experiences. Accommodating the indiscriminability of hallucinatory and veridical experiences proceeds more naturally: the externalist could appeal to Hume's views on the 'animal creation' to justify imposing a medically necessary 'presumption of veridicality' on deployment of the relevant concepts.

### 3.5. On Becoming Alienated

I now turn to seeing what it takes to derive the observed patterns of discriminability and indiscriminability among pairs that include experiences that are not safely veridical.

The general difficulty is that two experiences are reflectively discriminable just when there are medically possible *knowledgeable* reflective judgments about them that are *a priori* inconsistent. As I discuss below, for experiences that are not safely veridical, displaced judgments do not generate knowledge; accordingly, such experiences cannot be reflectively discriminated from other experiences just by displaced judgments. But some experiences that are not safely veridical fall into pairs that are antiparadigms of phenomenal sameness. Unless we can come up with some way of discriminating those experiences other than by displaced judgment, the epistemic descendant falls. But in positing alternative accounts of reflective knowledge, we run the risk of running afoul of the transparency of experience.

I now fill in the details behind the claim that displaced judgment does not always provide knowledge. First, as I discussed above, someone with a veridical experience can fail to acquire knowledge via displaced judgments when in danger of hallucinating.

Next, consider illusion. Suppose that Ilya advances a displaced judgment about his experience. The sense of this judgment is the same as that of Bill's displaced judgment. But Ilya's displaced judgment is not determinately true. Recall that its truth condition is indeterminate as between *Ilya sees something red* and *Ilya sees something white*. So at the very least, Ilya's displaced judgment is not *determinately* a case of knowledge of what experiential property he has.

Hallucinating Dean is worse off still. While the sense of his displaced judgment is the same as that of Bill's displaced judgment, what is its truth condition? The judgment is 'I see something this color.' This judgment is problematic in two ways: First, Dean does not see; second, 'this color' does not refer, so the judgment has no truth value. Since truth is a requirement of knowledge, Dean's displaced judgment is not a case of knowledge of what experiential property he has.

Of these three difficulties, Ilya's is perhaps the easiest to resolve: Perhaps we could just jettison the requirement of determinate knowledge in the analysis, so that even indeterminate discrimination counts as sufficient for phenomenal difference. I do not speculate on the ultimate success of

---

Second, if we had *a priori* insight into the nature of our experiential properties, the difference of sense required to discriminate Greg's and Inez's experiences would ramify as a difference of referent; this difference could not reside in the color seen (in both cases, green), so it would have to reside in properties like *G* and *R*. This would make for a violation of the transparency of experience as we have understood it. Chalmers (2006), of course, recognizes this dialectic: his approach (i) opts for a weaker understanding of transparency; (ii) advances as consistent with that understanding an intricate and arcane theory of the natures of *G* and *R* involving multiple levels of content and self-misrepresentation of acquaintance with uninstantiated primitive color properties.

this approach, however, because it is piecemeal and does nothing to resolve the difficulties for hallucinating and endangered subjects. My solution for Dean's case is clearly effective in the other cases, so I turn directly to it. (Here I treat error, as found in Dean's case, differently from ignorance, as found in the remainder of our cases.)

There are two reasons why the displaced color judgment ('this experience has *seeing something this color* as an experiential property') is not true of Dean's experience. The first is that Dean does not see anything; the second is that 'this color' does not refer. The difficulties are independent: Even if the color predicate were not demonstrative, the first difficulty would make the judgment false; even if the relation to the color were less demanding than seeing, the second difficulty would still rob any judgment containing 'this color' of a truth value. What is needed, then, are ways of transforming the initial judgment, first, so that the relation to the color is less demanding than seeing, and second, so that the color term does not require an actual present referent in order to contribute to the truth condition. If we could transform the displaced judgment in these ways, the result would still be 'based on' the displaced judgement. Moreover, this would seem to be in accordance with the 'minimal added conceptual material' requirement of the transparency thesis: the aim of all this maneuvering is to manufacture a judgement with a content that is less specific than the displaced judgement.

On the first difficulty, notice that the sense of the reflective judgment about Dean's experience contains more information than is strictly necessary to establish *a priori* its distinctness from Greg's experience of seeing a green thing. All that is really necessary to establish this are a sense applied to Dean's experience along the lines of 'the subject at the center instantiates *R*' and a sense applied to Greg's experience along the lines of 'the subject at the center instantiates *G*' (or perhaps some sense of an abstractness level between these and the senses of displaced judgments). Each of these senses is true, and plausibly safe. And while we are not in a position to directly grasp either of those simple senses, each is entailed by the more complex sense of the actual displaced reflective property judgments. The straightforward remedy here is to abstract from the initial judgment by *quantifying*. This recalls Smart's (1959, 149) famous claim that 'when a person says, 'I see a yellowish-orange after-image,' he is saying something like this: '*There is something going on which is like what is going on when I have my eyes open, am awake, and there is an orange illuminated in good light in front of me, that is, when I really see an orange.*'"

Plausibly, the Smart-like judgment 'there is something going on in me which is intrinsically like what is going on when I see something red' abstracts away the requirement that one *see* the color that regularly causes *R* in one from the sense of the initial displaced-color judgment. But the involvement of the context-independent expression 'red' prevents the judgment in question of being tied to the nature of the experience in the way displaced judgments are, so the Smart-like judgment does not count as purely reflective in the desired way. A bit more specifically: first, this sort of judgment has nothing to do with the transparency of experience, because

no attention to anything is involved, and second, we have no guarantee that the Smart-like judgment will be true in any particular case: The thought can be entertained when one is seeing a blue thing, so the story we have been providing to explain the truth of our reflective judgments (viz., that they cannot be entertained in the absence of the internal property they concern) would not be applicable; some other explanation would be needed.

The second difficulty for the displaced color judgement as an account of reflective knowledge of Dean's experience is the problem of simultaneously preserving the tie contributed by the color demonstrative to the internal properties *R*, *G*, and the like, and also eliminating the color demonstrative's attendant dependence of a truth value of the judgment on the existence of an actual present referent. The sense of this color demonstrative is along the lines of 'the color that now and typically causes me to instantiate *R* when seen.' The clause that is responsible for both the cost and the benefit is the current-causation clause: In the veridical case, it assures the tie to the relevant internal properties, but in the hallucinatory case, it deprives any judgment containing the concept of a truth value. What is needed is a concept that is tied to the internal property, perhaps by way of being implicitly understood in terms of the demonstrative concept, but that can refer even without a present instance of its referent.

Something like the following explanation seems plausible. A number of authors, including Kripke (1972), have suggested that the sense of a context-independent color concept like 'red' is something like 'the color that typically causes me to instantiate *R*.' This sense is more abstract than the sense of our demonstrative concept, in that it lacks the current-causation clause. It is also more *objective* than our sense, in the sense that it envisages the property in a way that prescind from the subject's current circumstances (while still involving the subject as a more statistical measure of the property). It is plausible that we learn to grasp this sort of concept by, as it were, considering a color of which we demonstratively conceive 'as unseen,' by abstracting away from the fact that the color is a current object of vision and focusing exclusively on its instances that are not currently present to us. The result of this sort of consideration would be a concept that is pedagogically rooted in the demonstrative concept but that is also semantically self-standing, in that it does not rely compositionally for its semantic properties on the semantic properties of the original concept.

We can represent the psychological process by which we learn to grasp a context-independent color concept by abstraction from a color-demonstrative concept by a noncompositional operator 'O' (for 'objective') on color-demonstrative concepts: If the sense of 'this color' is 'the color that now and typically causes and the subject at the center to instantiate *X* when seen,' 'O(this color)' is a concept with the sense 'the color that typically causes the subject at the center to instantiate *X* when seen,' and the referent the color, whatever it may be, that typically causes the subject at the center to instantiate *X* when seen. So, relative to Bill and Dean, the sense of 'O(this color)' is 'the color that typically causes the subject at the center to instantiate *R* when seen,' and its referent is redness; for Greg, the sense is 'the color



that typically causes the subject at the center to instantiate *G* when seen,' and its referent is greenness, and for Inez the sense is 'the color that typically causes the subject at the center to instantiate *R* when seen,' while the referent is greenness. Importantly, the tie between a concept *C* and the concept 'O(*C*)' is a *pedagogical* tie rather than a *semantic* tie. It is a tie at the level of sense, rather than at the level of reference. Because 'O' does not function at the level of reference, standard assumptions about compositionality in the determination of truth-conditions need not apply. In particular, there is no need to suppose that in order for 'O(*C*)' to refer, '*C*' must also refer. Rather, what is true is that in order for 'O(*C*)' to have a sense, '*C*' must have a sense. And, as we have seen, Dean's concept 'that color' does have a sense. Accordingly, using 'O(*C*)' is a way to form a judgment 'by way of' a reference-free concept *C* without infecting the full judgment with semantic defectiveness.

Both of these devices—'smartening,' converting a judgement about what is seen into a Smart-like judgement, and 'objectivizing'—enable one to 'alienate' one's internal situation from its current immediate perceptual tie to one's surroundings: Smartening enables one to consider situations internally like seeing that need not involve seeing, while objectivizing enables one to consider a property—putting it a bit roughly—as looking a certain way dispositionally (though perhaps not occurrently).<sup>16</sup> By exploiting both of these alienating devices, we can entertain the *alienated judgment* that results from slotting 'O(this color)' into the smartened displaced judgment: 'There is something going on in me that is intrinsically like what is going on when I see something O(this color).'

Alienated judgments can play the required role in explaining the reflective discriminability of Dean's and Greg's experiences. First, an alienated judgment is a case of knowledge. The alienated judgment is true of Dean's experience because 'O(this color)' refers to redness, and there is something going on in Dean that is like what is going on when he sees something red: He instantiates *R*. Moreover, the judgment is safely true (and, in general, alienated judgments will be safely true even when their corresponding displaced judgments are dangerously true). Accordingly, the alienated judgment is a case of knowledge.<sup>17</sup>

16. These devices bear a certain distant resemblance to Loar's (2003a) notion of 'oblique reflection,' perhaps.

17. More of this sort of semi-technical work remains to be done, I fear, if the externalist is to accommodate the full range of internalist intuitions: to build a style of judgment which *any* subject we might judge to be *R* could exploit to acquire a safe belief that they are *R*. After all, Victor, who is *R* and envatted, and therefore never caused to be *R* by seeing any color, would not refer to any property with 'O(this color),' and nor is there anything that typically happens when he sees colors. Thanks to Chris Mole for raising this objection.

Another concern addresses *normativity*. Safety is a mere statistical notion; what does it have to do with *justification*? On Chalmers's (2003) account of introspection, one who is *R* is 'acquainted' with *R*, which bestows justification on judgements that one is *R*. It is not clear that my account cannot appeal to a similar maneuver: accept that *R* is an experiential property and assert one is acquainted with all of one's experiential properties. I would need to deny, of course, that one is in a position to conceptualize in a relatively straightforward way every object of one's acquaintance. But this is a position that Chalmers also advances, though in a rather more restricted form (§4.1). Thanks to David Chalmers for raising this objection.



Second, the requisite patterns of *a priori* inconsistency hold. The sense of Dean's alienated judgment is something like 'there is something going on in me that is intrinsically like what is going on when I see something the color of which typically causes *R* in me when seen'—which straightforwardly entails the sense 'the subject at the center instantiates *R*.' This sense is *a priori* inconsistent with the sense of the displaced color judgment applied to Greg's experience, so that the experiences are discriminable on the basis of the alienated and displaced judgments.

So the alienated judgment does seem to be a (more-or-less reflective) way of discriminating among experiences that are not safely veridical.

### 3.6. Qualia and Consciousness

On the basis of such an alienated judgment, we can construct concepts that refer to the internal qualities *R*, *G*, and so forth. Beginning with an alienated judgment of the form 'there is something going on in me that is intrinsically like what is going on when I see something *O*(this color),' we can form the description 'the property instantiated in me that is intrinsically like what is going on when I see something *O*(this color).' We can then use this description to fix the referent of various property terms; for instance, on the basis of Bill's experience, we could define in this way the expression 'q-red,' on the basis of Greg's experiences, we could define 'q-green,' and so on. In this way, we enable ourselves to refer to internal qualities, from something like a first-person perspective.<sup>18</sup> This constructed referential ability gives rise to a series of questions that will be relevant in developing the ontological part of our interpretation, and in assessing the broader philosophical impact of the views developed here.

1. Is our ability to reflectively refer to these internal 'q-properties' consistent with the key thesis of our displaced perception theory of reflection?

Yes: The initial statement of the theory required only that any reflective knowledge of which experiential properties a subject has is *based in* displaced judgments. Set aside the question whether the q-properties are experiential, and whether q-concepts genuinely provide a sufficiently robust 'knowledge which' properties are experiential; these referential q-concepts are in fact 'based in' displaced judgments (via smartening, unseeing, abstraction of property descriptions, and reference fixation), in the sense of either semantically or pedagogically defined in terms of displaced judgments.

2. Do q-concepts enable 'knowledge of which' properties the q-properties are?

On a weak sense of 'knowing which,' yes: They enable reference to q-properties. On a very strong sense (Lewis 1995), maybe. On behalf of

18. This is perhaps something like what Shoemaker (1991, 522) refers to as the 'reconceptualization of what one was aware of all along' that 'produce[s] explicit awareness of qualia as such.'

an affirmative answer, q-concepts appear to be ‘epistemically rigid’: no property other than R might be the property instantiated in one that is intrinsically like what is (typically) going on when one sees something of the color that when seen typically causes one to instantiate R. After all, get rid of the characterization, and it becomes clear that (modulo certain presuppositions) this is just a fancy way of specifying the condition of being equal to R. On behalf of a negative answer, though, the cognitive complexity built into the concept renders it a poor substitute for the ideal of acquaintance, in which the subject simply meets the property face-to-face, as it were.

### 3. What is the nature of a q-property?

The only restriction I can see on theory here is that the q-properties be able to serve in modes of presentation in demonstrative concepts of color. (Moreover, it is far from obvious whether a single filler of this role must have a unified essence.) Though limitations of space prevent full consideration of this point, philosophical considerations that have been advanced on behalf of various answers to this question tend to strike me as not germane.<sup>19</sup>

### 4. Is this reference to q-properties consistent with the transparency of experience?

First, recall that I phrased the transparency doctrine ‘most reflective knowledge of which experiential property one’s experience has requires attention to entities—objects or property tokens—in one’s environment.’ Q-concepts provide no knowledge of the true nature of hallucinatory experience as hallucinatory, of the experiential properties that it has *in exclusion* of the external properties an indiscriminable veridical experience would have. Moreover, as I discussed above, it is not clear either that q-properties are experiential or that q-concepts afford ‘knowledge which.’

Second, the very same capacity to attend is exploited in q-concepts as in displaced concepts. Accordingly, the situation ratifies the following: *Attempting* to acquire reflective knowledge of experiential properties requires *attempting* to attend to entities in one’s environment. Accordingly, to the extent that we are tempted to apply the transparency thesis even to hallucinatory experiences, this can be explained away by its *seeming* true, from the first-person perspective, in such cases.

Third, the q-concept does not bring in any heavy-duty cognitive resources that offend against the spirit of the transparency thesis: Smartening merely quantifies at the level of the judgment, unseeing merely quantifies at the level of the sense, and descriptive reference fixation is an elementary and highly general semantic capacity. Nothing remotely like direct attention to internal experiential qualities is involved.

19. In particular, (Shoemaker 1994, 2006) and Chalmers (2006) seem to represent our cognitive access to q-properties as being much too direct. In an earlier incarnation, Shoemaker (1991, 521) was right on, regarding this access as ‘theoretical,’ although ‘bringing out something that is implicit in folk psychology.’

5. What then can be concluded about the nature of perceptual experience on the basis of the transparency of experience?

Not very much. First of all, transparency is compatible with veridical spectral inversion. My story does not discriminate between Inez and the rest of our subjects.

Second, transparency is compatible with introspective knowledge of intrinsic qualities, such as *R*. It is very plausible that alienated judgments count as 'introspective,' and *R* is, we are assuming, an intrinsic quality, so alienated judgments bring introspective knowledge of intrinsic qualities of experience, and our capacity to make alienated judgements is compatible with transparency. And, as I argue in the concluding section, all of this is independent of whether *R* is understood as an experiential property.

Third, transparency is compatible with *R* being any of a wide range of properties: a quale, a relation to sense-data, an intentional property, or some mixture of these. The only constraints I have imposed on *R* are that it be intrinsic, conceptually irreducible, and in a position to serve in the modes of presentation of colors in the way specified. Nothing in either feature seems to be relevant to these metaphysical debates. (Attempts to constrain the nature of *R* are discussed in the prefatory text of section 2, and footnotes 5 and 15.)

6. Why have these points been elusive?

I speculate that the following may be at work. Philosophers have conflated two styles of reflection: reflection by displaced judgment, and reflection by alienated judgment. Reflecting by displaced judgement, we think we find external aspects to experience: the colors we see. Reflecting by alienated judgment, we think we find internal aspects: the q-properties. But the logical structure of our concepts is opaque. Accordingly, we erroneously take ourselves to have done some unified thing and thereby accessed a unified domain. As a result, we search for theories that eliminate one or the other side (qualia theory, direct realism), or attempt to bundle them together (representationism). But once we recognize that we have done two things, we need neither eliminate nor bundle. Debate about metaphysics fuelled by apparently conflicting phenomenological results is dissolved if the results do not conflict.

7. Should q-properties be regarded as experiential properties?

This question is a loose end from section 1.4. It is not easy to answer.

What pictures are attached to the answers to the question? If we answer negatively, we would think of the nature of Bill's (Greg's) consciousness as being something like *seeing a red (green) thing*; if affirmatively, something like *R-ly (G-ly) seeing something of some color or other*. On the former picture, veridical visual consciousness is a simple relation to the environment, without any intrinsic features, and is characterized solely by one's openness to various objects and features in one's environment. (The same could also be said of illusory consciousness. There is no such thing as hallucinatory visual consciousness: a hallucinating subject is simply not visually conscious.) On the

latter picture, visual consciousness, whether veridical, illusory, or hallucinatory, is more like a 'field' with intrinsic features of its own, which serves as the basis of one's openness to various aspects of the environment. On this picture, visual consciousness involves something like 'paint that points,' to use Loar's (2003b, 251) memorable phrase. The former picture is compatible with externalism, the latter is not. (Note an ignorance/error alternation. According to the externalist, error is radically different from correctness; ignorance is inevitable and forces symmetric treatment of distinct cases. The 'field' conception treats error and truth as on a par, assimilating both to cases of ignorance.)

Obviously the approaches could be combined to form a view compatible with externalism, by thinking of the nature of Bill's (Greg's) consciousness as something like *R-ly (G-ly) seeing a red (green) thing*. The picture of veridical perceptual consciousness then would be something like a field with external entities dotted around in it. This approach raises obvious if perhaps somewhat vague worries about overdetermination, redundancy, excessive complication, and wishy-washiness, so I will set it aside.

On either picture, one's openness in perceptual consciousness to some object or feature serves as a basis for reference to that entity via perceptual demonstrative concepts. But the relation between perceptual consciousness and the sense of such a concept differs significantly. On the field view, forming such concepts is largely a matter of thought's borrowing or exploiting resources already available within perceptual consciousness. On the externalist view, forming such concepts is a rather more creative enterprise: thought brings q-properties from without perceptual consciousness to assist in grappling with the entities conveyed to the subject within perceptual consciousness.

What considerations motivate answering one way or another? In section 1.3, recall, I argue that q-properties generate an explanatory gap. It does not follow from this that q-properties are experiential properties; after all, in Chalmers's (2006) Eden, 'perfect colors' generate an explanatory gap but are qualities of mindless entities such as trees.

There is certainly a sense in which q-properties count as 'sensational' or 'perceptual'; after all, which q-properties one instantiates typically tracks one's perceptible environment. But it still does not obviously follow that q-properties are experiential; after all, there does not seem to be any obvious contradiction in the notion of an irreducible property of an organism that tracks the perceptible environment without characterizing the organism's conscious state.

It is certainly desirable to provide distinct characterizations of Bill's and Greg's conscious states. But this does not settle whether the distinction should be drawn by appeal to the differing q-properties they instantiate or by appeal to the differing colors they see.

Nor is it uncontroversial that, despite the fact that Greg is G and Inez is R, we should provide distinct characterizations of Greg's and Inez's conscious states (ignoring the numerical difference between their tomatoes): while reflection on Greg's and Inez's experiences reveals them to instantiate incompatible properties, this does not settle whether these properties are aspects of *consciousness*.

There are semantical considerations both for and against each candidate. Extracting reference to colors from perceptual experience is completely straightforward; by contrast, extracting reference to q-properties from perceptual experience requires extensive theoretical sophistication. It is possible in principle (via alienated judgement) to extract a direct conceptualization of the nature of q-properties from perceptual experience; by contrast, a direct conceptualization of the nature of colors cannot be extracted from perceptual experience. Both of these features are ones it is attractive to attribute to consciousness and its various manifestations. On standard internalist views (compare Chalmers 2003), q-properties are both candidate objects of straightforward reference and candidate objects of nature-revealing conception. But if externalism is correct, this is too much to hope for. (Note here an ignorance/error alternation. Perceptual demonstrative thought leaves us ignorant of the natures of the features and objects we see, and in error when we are hallucinating. The ignorance is indelible; the error is washed out only when we willfully render ourselves ignorant.)

Methodological considerations pertaining to the transparency of experience may provide a way of breaking the tie. If q-properties are not aspects of consciousness, visual consciousness of external objects simply involves no intrinsic aspects. This latter doctrine is plausibly regarded as a metaphysical formulation of the doctrine of the transparency of experience. Ordinarily the doctrine is cashed in dispositional epistemic or semantic terms resembling those in section 3. A certain pleasing theoretical unification would be attained by providing these doctrines with a categorical ground along the lines of the metaphysical transparency doctrine; and the metaphysical doctrine is perhaps a bit more in line with the metaphor of transparency (the transparency of glass is a feature of *glass* which explains but is not identical to our epistemic situation in regard to glass).

A great deal remains to be said if this question is to be settled. If we have notionally distinguished consciousness and phenomenality, we lose a way of operationalizing questions about whether a certain property is an aspect of consciousness. Although the externalist must deny that q-properties are aspects of consciousness, the impartial observer may find it still too early to pronounce definitively on the matter. (We may ultimately decide that there is no real distinction between these views. If so, we would need to formulate a strategy for thinking about consciousness rather different from the one employed in this chapter; perhaps a strategy hinted at in the quotation from Hinton with which this chapter begins.)

#### 4. THE METAPHYSICS OF PHENOMENALITY

This completes the conceptual part of our project: The notion of reflective indiscriminability I have constructed seems to square up with the paradigms of phenomenal sameness. So now on to the metaphysical part.



Our questions are: Which properties are the phenomenal properties, and how are they related to the externalist's experiential properties? The answers here depend on a conceptual issue and an ontological issue.

On the conceptual issue, the question is what we mean by 'phenomenal properties': A first parameter is whether the metaphysical or epistemic descendant of 'phenomenal' is intended; a second is whether we insist that 'phenomenal' can apply only to an experiential property. On the ontological issue, the question is what the experiential properties are: in particular, whether the q-properties are counted among the experiential properties.

Suppose first that the metaphysical descendant is intended. Then, in effect, 'phenomenal property' means 'experiential property.' In that case, assessing what the phenomenal properties of an experience are requires assessing what the experiential properties of the experience are. The externalist's theory here is clear, at least in outline, up to the question of whether the q-properties are counted among the experiential properties.

Now suppose that the epistemic descendant is intended. On the epistemic descendant, two (projectable) experiences are phenomenally the same just if they are reflectively indiscriminable. Now, one thing we have noted is that reflective indiscriminability crosscuts all aspects of the externalist's characteristic experiential properties. Bill's and Tina's experiential properties differ haecceitistically in their object but are indiscriminable; Bill's and Inez's experiential properties differ quidditistically in their object but are indiscriminable; Bill's and Dean's experiential properties differ with respect to factivity but are indiscriminable. Conversely, Greg's experience is discriminable from Inez's despite being just alike with respect to factivity, and haecceitistically and quidditistically in their object. Rather, phenomenal sameness tracks q-properties.

Suppose either that q-properties are experiential or that it is not analytic that phenomenal properties are experiential. In that case, we can say that the phenomenal properties are q-properties (or, perhaps, q-properties qualifying properties sufficiently abstract to be shared among our paradigms of phenomenal sameness, such as quantified or disjunctive properties).

If q-properties are not experiential, and it is analytic that phenomenal properties are experiential, then the metaphysics of phenomenality will end up looking more functionalist: We might end up saying things like 'Inez's and Greg's tokens of *seeing t's greenness* realize different phenomenal properties while Inez's and Bill's distinct tokens experiential properties realize the same phenomenal property' (how exactly to refine the formulation of this theory does not influence our overall aim).<sup>20</sup>

20. Or maybe the concept of phenomenality, or the only satisfiable concept of phenomenality, is a concept of a quantified higher order response-dependent property of the form 'having some property that makes for reflective indiscriminability from *F*.'



My story in the foregoing sections of this chapter is independent of these nuances of ontological bookkeeping, however. So let me leave these nuances behind and reflect more generally on the explanatory role phenomenality plays in that story. In a nutshell, the view is that *most of the work of phenomenality is done by appropriate conceptual first-person thought about perception of external qualities*. Fix the senses of such thoughts, and you fix the facts about phenomenality (up to issues of ontological bookkeeping). Phenomenal difference and sameness tracks difference and sameness of such thoughts, crosscutting difference or sameness of experiential properties; phenomenality can perhaps even exist in the *absence* of experiential properties. (What experiential properties does Dean have? We don't know; maybe he has none.)

Phenomenality may be therefore regarded as a projection of this conceptual structure onto experience. In this respect, my notion of 'phenomenality' corresponds closely to Kant's, with the 'noumenal' structure amounting to the nonconceptual relations to the outside world that constitutes the uncontroversial ontological structure of experience. Perhaps some nonartificial, internal aspect of experience corresponds to that structure (i.e., if the q-properties get counted as experiential), but it is far from clear what difference this would make.

This is not to say that phenomenality and experience are utterly disconnected, of course. Phenomenality is projected onto experience by a cognitive structure culminating in the q-concepts. The base-level concepts that get this structure up and running are the displaced concepts; these displaced concepts have the externalist's experiential properties as their ordinary referents, and this referential relationship requires a certain degree of regular association between the experiential properties and the internal features that characterize the senses of those concepts. So the relation between phenomenality and experience is not completely free.

While many philosophers have made phenomenality into a paradigm of consciousness, this seems to me to invert the true conceptual order, on which our primary concepts of consciousness concern its external aspects. The explanatory gap arises in the first instance through these concepts, which concern *seeing*, an irreducible sort of conscious visual openness to the world. External consciousness is the subject matter of the judgments from which our concepts of phenomenal properties are reached by abstraction; phenomenality shapes the contours of our demonstrative access to the external world through—and therefore of our reflective access to—external consciousness.<sup>21</sup>

21. As I ordered a Brazilian Clover at Toronto's Manic Coffee to commence the final day of work on this chapter, I spotted a man wearing a Husserl T-shirt with the slogan 'Phenomenology: leave the world outside.' But the world outside is still present, if abstractly, in phenomenology.

## REFERENCES

- Bayne, Tim. 2007. Conscious States and Conscious Creatures: Explanation in the Scientific Study of Consciousness. *Philosophical Perspectives* 21: 1–22.
- Block, Ned. 1995. On a Confusion about the Function of Consciousness. *Behavioral and Brain Sciences* 18: 227–247.
- Byrne, Alex. 2001. Intentionalism Defended. *Philosophical Review* 110: 199–240.
- Byrne, Alex. 2004. What Phenomenal Consciousness Is Like. In *Higher-Order Theories of Consciousness*, ed. Rocco Gennaro. Amsterdam: John Benjamins.
- Byrne, Alex. 2005. Introspection. *Philosophical Topics* 33: 79–104.
- Byrne, Alex. 2009. Experience and Content. *Philosophical Quarterly* 50: 429–451.
- Byrne, Alex, and David R. Hilbert. 1997a. Colors and Reflectances. In *The Philosophy of Color*, ed. Alex Byrne and David R. Hilbert. Vol. 1 of *Readings on Color*. Cambridge, Mass.: MIT Press.
- Byrne, Alex, and David R. Hilbert. 1997b. Introduction. In *The Philosophy of Color*, ed. Alex Byrne and David R. Hilbert. Vol. 1 of *Readings on Color*. Cambridge, Mass.: MIT Press.
- Byrne, Alex, and Heather Logue. 2008. Either/Or. In *Disjunctivism: Perception, Action, Knowledge*, ed. Adrian Haddock and Fiona Macpherson. Oxford: Oxford University Press.
- Campbell, John. 2002. *Reference and Consciousness*. Oxford: Oxford University Press.
- Chalmers, David J. 1996. *The Conscious Mind*. Oxford: Oxford University Press.
- Chalmers, David J. 2003. The Content and Epistemology of Phenomenal Belief. In *Consciousness: New Philosophical Essays*, ed. Quentin Smith and Alexander Jokic. Oxford: Oxford University Press.
- Chalmers, David J. 2004. The Representational Character of Experience. In *The Future for Philosophy*, ed. Brian Leiter. Oxford: Oxford University Press.
- Chalmers, David J. 2005. The Foundations of Two-Dimensional Semantics. In *Two-Dimensional Semantics: Foundations and Applications*, ed. Manuel Garcia-Carpintero and Josep Macià. Oxford: Oxford University Press.
- Chalmers, David J. 2006. Perception and the Fall from Eden. In *Perceptual Experience*, ed. Tamar Szabó Gendler and John Hawthorne. Oxford: Oxford University Press.
- Chalmers, David J. 2009. The Two-Dimensional Argument against Materialism. In Chalmers, *The Character of Consciousness*. Oxford: Oxford University Press.
- Chalmers, David J. Unpublished manuscript. Verbal Disputes and Philosophical Progress. Revision of 1 September 2009.
- Crane, Tim. 2006. Is There a Perceptual Relation? In *Perceptual Experience*, ed. Tamar Szabó Gendler and John Hawthorne. Oxford: Oxford University Press.
- Dickie, Imogen. 2010. We Are Acquainted with Ordinary Things. In *New Essays on Singular Thought*, ed. Robin Jeshion. Oxford: Oxford University Press.
- Dretske, Fred I. 1995. *Naturalizing the Mind*. Cambridge, Mass.: MIT Press.
- Evans, Gareth. 1982. *The Varieties of Reference*. Oxford: Oxford University Press.

- Fish, William. 2009. *Perception, Hallucination, and Illusion*. New York: Oxford University Press.
- Gendler, Tamar Szabó, and John Hawthorne. 2006. Introduction. In *Perceptual Experience*, ed. Tamar Szabó Gendler and John Hawthorne. Oxford: Oxford University Press.
- Hawthorne, John, and Karson Kovakovich. 2006. Disjunctivism. *Proceedings of the Aristotelian Society* 80(Supplement): 145–183.
- Hellie, Benj. 2004. Inexpressible Truths and the Allure of the Knowledge Argument. In *There's Something about Mary*, ed. Peter Ludlow, Yujin Nagasawa, and Daniel Stoljar. Cambridge, Mass.: MIT Press.
- Hellie, Benj. 2005. Noise and Perceptual Indiscriminability. *Mind* 114: 481–508.
- Hellie, Benj. 2007a. Factive Phenomenal Characters. *Philosophical Perspectives* 21: 259–306.
- Hellie, Benj. 2007b. 'There's Something It's Like' and the Structure of Consciousness. *Philosophical Review* 116: 441–463.
- Hinckfuss, Ian. 1970. J. M. Hinton on Visual Experiences. *Mind* 79: 278–280.
- Hinton, John Michael. 1973. *Experiences: An Inquiry into Some Ambiguities*. Oxford: Clarendon Press.
- Hume, David. 1777. *An Enquiry Concerning Human Understanding*. Reprint, Indianapolis, Ind.: Hackett, 1993.
- Jackson, Frank. 1977. *Perception: A Representative Theory*. Cambridge: Cambridge University Press.
- Johnston, Mark. 2004. The Obscure Object of Hallucination. *Philosophical Studies* 103: 113–183.
- Kennedy, Matthew. 2010. Explanation across Good and Bad Experiential Cases. In *Hallucination*, ed. Fiona MacPherson. Cambridge, Mass: MIT Press.
- Kennedy, Matthew. 2009. Heirs of Nothing: The Implications of Transparency. *Philosophy and Phenomenological Research* 79: 574–604.
- Kripke, Saul Aron. 1972. *Naming and Necessity*. Reprint, Cambridge, Mass.: Harvard University Press, 1980.
- Levine, Joseph. 1983. Materialism and Qualia: The Explanatory Gap. *Pacific Philosophical Quarterly* 64: 354–361.
- Lewis, David. 1995. Should a Materialist Believe in Qualia? *Australasian Journal of Philosophy* 73: 140–144.
- Loar, Brian. 2003a. Transparent Experience and the Availability of Qualia. In *Consciousness: New Philosophical Essays*, ed. Quentin Smith and Alexander Jokic. Oxford: Oxford University Press.
- Loar, Brian. 2003b. Phenomenal Intentionality as the Basis of Mental Content. In *Reflections and Replies: Essays on the Philosophy of Tyler Burge*, ed. Martin Hahn and Bjorn Ramberg. Cambridge, Mass.: MIT Press.
- Lycan, William G. 2001. The Case for Phenomenal Externalism. In *Metaphysics*, ed. James Tomberlin. *Philosophical Perspectives*, vol. 15, 17–36. Oxford: Blackwell.
- Martin, Michael G. F. 1998. Setting Things before the Mind. In *Contemporary Issues in the Philosophy of Mind*, ed. Anthony O'Hear, 157–179. Cambridge: Cambridge University Press.
- Martin, Michael G. F. 2004. The Limits of Self-Awareness: Disjunctivism and Indiscriminability. *Philosophical Studies* 120: 37–89.

- Martin, Michael G. F. 2006. On Being Alienated. In *Perceptual Experience*, ed. Tamar Szabó Gendler and John Hawthorne. Oxford: Oxford University Press.
- Nagel, Thomas. 1974. What Is It Like to Be a Bat? *Philosophical Review* 83: 435–450.
- Peacocke, Christopher. 1983. *Sense and Content: Experience, Thought, and Their Relations*. Oxford: Clarendon Press.
- Raffman, Diana. 2000. Is Perceptual Indiscriminability Nontransitive? *Philosophical Topics* 28: 153–175.
- Shoemaker, Sydney. 1990. Qualities and Qualia: What's in the Mind? *Philosophy and Phenomenological Research* 50: 109–131.
- Shoemaker, Sydney. 1991. Qualia and Consciousness. *Mind* 100: 507–524.
- Shoemaker, Sydney. 1994. Phenomenal Character. *Nous* 28: 21–38.
- Shoemaker, Sydney. 2006. On the Ways Things Appear. In *Perceptual Experience*, ed. Tamar Szabó Gendler and John Hawthorne. Oxford: Oxford University Press.
- Siegel, Susanna. 2004. Indiscriminability and the Phenomenal. *Philosophical Studies* 120: 90–112.
- Siegel, Susanna. 2005. The Contents of Perception. In *The Stanford Encyclopedia of Philosophy*, ed. Edward N. Zalta. Online. Available: <http://plato.stanford.edu/archives/spr2005/entries/perception-contents/>.
- Siegel, Susanna. 2006. Which Properties Are Represented in Perception? In *Perceptual Experience*, ed. Tamar Szabó Gendler and John Hawthorne. Oxford: Oxford University Press.
- Siegel, Susanna. 2008. The Epistemic Conception of Hallucination. In *Disjunctivism: Perception, Action, Knowledge*, ed. Adrian Haddock and Fiona Macpherson. Oxford: Oxford University Press.
- Siewert, Charles. 2003. Consciousness and Intentionality. In *The Stanford Encyclopedia of Philosophy*, ed. Edward N. Zalta. Online. Available: <http://plato.stanford.edu/archives/sum2003/entries/consciousness-intentionality/>.
- Siewert, Charles. 2004. Is Experience Transparent? *Philosophical Studies* 117: 15–41.
- Smart, John Jamieson Carswell. 1959. Sensations and Brain Processes. *Philosophical Review* 68: 141–156.
- Soteriou, Matthew. 2005. The Subjective View of Experience and Its Objective Commitments. *Proceedings of the Aristotelian Society* 105: 193–206.
- Speaks, Jeffrey J. 2009. Transparency, Intentionalism, and the Nature of Perceptual Content. *Philosophy and Phenomenological Research* 79: 539–573.
- Stalnaker, Robert C. 2008. *Our Knowledge of the Internal World*. Oxford: Oxford University Press.
- Sturgeon, Scott. 2008. Disjunctivism about Visual Experience. In *Disjunctivism: Perception, Action, Knowledge*, ed. Adrian Haddock and Fiona Macpherson. Oxford: Oxford University Press.
- Thompson, Brad. 2006. Colour Constancy and Russellian Representationalism. *Australasian Journal of Philosophy* 84: 75–94.
- Thompson, Brad. 2007. Shoemaker on Phenomenal Content. *Philosophical Studies* 135: 307–334.
- Thompson, Brad. 2009. Senses for Senses. *Australasian Journal of Philosophy* 87: 99–117.

- Tye, Michael. 2000. *Consciousness, Color, and Content*. Cambridge, Mass.: MIT Press.
- Williamson, Timothy. 1990. *Identity and Discrimination*. Oxford: Blackwell.
- Williamson, Timothy. 2000. *Knowledge and Its Limits*. Oxford: Oxford University Press.