

Color and the Inverted Spectrum

DAVID R. HILBERT and MARK ELI KALDERON

If you trained someone to emit a particular sound at the sight of something red, another at the sight of something yellow, and so on for other colors, still he would not yet be describing objects by their colors. Though he might be a help to us in giving a description. A description is a representation of a distribution in a space (in that of time, for instance).

Ludwig Wittgenstein
(1945, p. 187)

1. Intentionalism and Transparency

When we open our eyes, one of the most striking features of the occurrent scene before us is the apparent multitude of colored surfaces, transparent volumes, and radiant light sources.¹ Grass, emeralds, and certain lava lamps, for example, all appear green to us. If something visually appears green to us, and it is in fact green, then our visual experience is *veridical*, at least in part.² If, on the other hand, something visually appears green to us but is not in fact green, then our visual experience is, at least in part, *illusory*. Specifically, if some object *o* visually appears green to a normal subject, then that color experience is veridical only if its total content includes the proposition that *o* is green and that proposition is true.³ Such experiences are in this sense *green-representing*.⁴ The world appears colored to us because our visual experience *represents it* as colored. One way of typing visual experience, then, is in terms of synonymy or sameness of representational content.

Moreover, our experiences of emeralds, grass, and the like have a subjective character that is immediately present to consciousness—there is a particular way it is like to undergo such experiences. They are *green-feeling* in the sense that such experiences enjoy a distinctive phenomenology. We shall refer to the distinctive way it is like to undergo an experience as its *phenomenal* character. In so doing, we do not mean to commit myself, at least not initially, as to the nature of phenomenal properties. We shall not assume, for instance, that the phenomenal features of color experience are, or are constituted by, *qualia*, *i.e.*, intrinsic, non-representational features of experience

that are immediately present to consciousness. Talk of phenomenal properties is meant only to capture the intuitive, pre-theoretic sense that our visual experience has a subjective character. Another way of typing visual experience, then, is in terms of phenomenal similarity.

In describing two ways to type visual experience—in terms of synonymy and phenomenal similarity, we are *not* assuming that they are *distinct*. Indeed, we disavow belief in primitive phenomenal properties—instead we believe that the phenomenal properties of experience admit of a reduction to its content. *Intentionalism*, as this identification has come to be known, has recently been proposed by Gilbert Harman (1990). According to intentionalism, the phenomenal properties of experience are a subclass of its intentional or representational features. From the intentionalist perspective, the two ways of typing visual experience are, at bottom, one.

Intentionalism, broadly conceived, identifies the phenomenal character of experience with (some aspect of) representational content. There are at least three dimensions along which intentionalist reductions may vary. Intentionalists may, and in fact do, disagree about with *which* intentional features of experience to identify its phenomenal properties. The phenomenal character of color experience might be identified with its *color* content, as Harman proposes, or in terms of some *other* aspect of its total content. Thus, for instance, Sydney Shoemaker (1994) has claimed that color phenomenology is not explained by the representation of color properties by visual experience, but by its representation of some *other* property. *Phenomenal properties*, in Shoemaker's special sense of the term, are the relations experiences bear to the objects that elicit them. And it is the representation of *these* properties that accounts for the subjective character of our visual experience. Thus the first dimension along which intentionalist reductions vary is with which aspect of the total content of visual experience to identify its phenomenal character. Shoemaker is driven to ascribe such contents to visual experience in part by his conception of color content. According to Shoemaker, color contents are *Millian* in the sense that the content of a color experience is exhausted by the color property it denotes. In contrast, one might claim that color contents are *Fregean*—that there is more to the content of a color experience than the designated color property. The Fregean does not deny that our visual experience *refers* to color properties; he denies, instead, that its content is exhausted by such a reference. The second dimension, then, along which intentionalist reductions vary is in terms of how they individuate the content of experience. The third dimension concerns the *determinants* of visual content. What makes it the case that a visual experience has the content that it has? Some intentionalists, like Fred Dretske (1995), believe in a

teleological version of *correlational psychosemantics*—that it is because a state of a normally functioning visual system counterfactually covaries with the instantiation of some property that it comes to represent that property. Other intentionalists, like Harman (1990), believe that an experience has its content in virtue of its *conceptual role*, *i.e.*, the function that experience plays in the subject's psychology. Perhaps other accounts are possible as well. The important point is that one's views about the determinants of visual content will constrain the resources available to one's intentionalist reduction.

Why believe that phenomenal character is a function of representational content? One motivation is the so called *transparency* of experience. G.E. Moore (1903, pp. 37, 41) writes:⁵

In general, that which makes a sensation of blue a mental fact seems to escape us: it seems, if I may use a metaphor, to be transparent—we look through it and see nothing but the blue.

When we try to introspect the sensation of blue, all we can see is the blue: the other element is as it were diaphanous.

Moore is emphasizing the difficulty one encounters in distinguishing introspective awareness of one's color experience from introspective awareness of what it represents. When attending to what it is like to undergo a color experience, one attends to how things appear visually; and it would seem that what is manifest is the content of that experience and not some non-intentional feature of it. The transparency property is not merely confined to introspective awareness. If asked to *imagine* what it is like to undergo a particular color experience one could hardly do better than to simply imagine that color. That is, in representing to oneself the phenomenal character of a color experience it apparently suffices to represent the color itself. A good explanation of this, the intentionalist contends, is that attending to, imagining, and so on, the phenomenal character of a visual experience *just is* to attend to, imagine what it represents.⁶ Intentionalism is not merely an attempt to salvage functionalism from the difficulties it faces in accounting for the qualitative character of our experience (see, *inter alia*, Block, 1978; Jackson, 1982; and Nagel, 1974)—it enjoys an independent phenomenological motivation.

A *caveat* is in order. Moore's phenomenological observation should not be conflated with the intentionalist explanation of it. Moore acknowledges the *difficulty* encountered in distinguishing introspective awareness of color experience from introspective awareness of what it represents. One ought not, on this basis, make him out to be an intentionalist. Such an interpretation sits ill, for instance, with the

argumentative drift of Moore's essay. Moore criticizes the idealist for not distinguishing an experience from what it is an experience of. He claims that the idealist is misled in this conflation by transparency. It is because it is difficult to attend to one's experience without attending to what it represents that the idealist does not take sufficient care to distinguish an experience from the state of affairs it represents. In making this diagnosis, Moore is clearing the way for an act-object analysis of conscious experience—a conception of experience at odds with contemporary intentionalism. The intentionalist claims, not only that it is *difficult* to attend to the phenomenal character of visual experience without attending to what it represents, but that it is also *impossible* and provides an explanation for why this is so—the phenomenal character of experience just is (some aspect of) its representational content.

The phenomenological argument for intentionalism is not decisive, however, in that no reason has been given for thinking that transparency might not be vindicated in some *other* way. For example, as should be evident from Moore's discussion, there is an *idealist* explanation of transparency—the reason it is difficult to distinguish an experience from what it is an experience of is that the represented state of affairs *existentially depends on* the experience that represents it. There is, moreover, a *projectivist* explanation of transparency. According to the projectivist, when a normal subject visually represents an object *o* as green, he is *misrepresenting* *o*—he is misattributing to *o* a property that properly belongs to the subject's visual field.⁷ (See Boghossian and Velleman, 1989, 1991.) The projectivist would claim that the reason it is difficult to attend to a color experience without also attending to the color it represents is that color, properly conceived, is an intrinsic feature of visual experience. While providing some weight to the intentionalist thesis, the phenomenological argument does not conclusively establish it. Moreover, intentionalism itself faces a number of challenges. The full case for intentionalism will consist in its success in dealing with the challenges it faces and in illuminating the nature of experience.

2. Intentionalism and NECESSITY

Intentionalism has important modal consequences. The intentionalist believes that content and phenomenology could not vary independently. Since phenomenal character is identified with (some aspect of) visual content, the intentionalist is committed to the following:

COLOR REP → COLOR FEEL

There could be no difference in color phenomenology without a relevant difference in visual content.

COLOR FEEL → COLOR REP

There could be no relevant difference in visual content without a difference in color phenomenology.

The restriction to *relevant* differences is required to state these modal commitments in full generality. We want to allow that intentionalists may identify phenomenal properties with different intentional features of experience. So, for instance, whereas Harman would claim that there could be no difference in *color* content without a difference in phenomenology, Shoemaker would claim that what makes for a difference in subjective character is a difference in *phenomenal* content (in his idiosyncratic sense of the term). The disagreement between Harman and Shoemaker concerns which aspect of visual content properly explains color phenomenology. We do not want to prejudge this issue in an initial statement of these modal commitments. Taken together, COLOR REP → COLOR FEEL and COLOR FEEL → COLOR REP are equivalent to:

NECESSITY

Necessarily, the visual experiences of two normal subjects share the same relevant content iff they share the same phenomenology.⁸

While a consequence of intentionalism, NECESSITY does not itself imply intentionalism. To see this consider *the imagistic theory of meaning* popularly attributed to the British Empiricists.⁹ According to the imagistic theory of meaning, the content of experience is identified with, or constituted by, its phenomenal properties where phenomenal properties are conceived to be qualia. Clearly the imagistic theory of meaning is inconsistent with intentionalism—the intentionalist denies that there are any phenomenal properties so conceived.¹⁰ Though inconsistent with intentionalism, the imagistic theory of meaning, if true, would imply NECESSITY.

There is an apparent difficulty with NECESSITY. Suppose content and phenomenal properties were distinct. If the content and phenomenal character of visual experience were distinct features of it, then, by NECESSITY, there would be a necessary connection between distinct existences. But how could this be? Isn't this just the kind of mysterious necessary connection of which Hume complained? Though this difficulty is not decisive, it does emphasize an explanatory debt that the advocate of NECESSITY must discharge. The believer in NECESSITY must either explain why there could be no

difference in content without a difference in phenomenology or accept commitment to non-Humean necessary connections. Notice, the problem only gets going on the assumption that content and phenomenal properties are distinct features of visual experience. One familiar way to provide the requisite explanation and so avoid commitment to non-Humean necessary connections is to *identify* color content and color phenomenology. And this is precisely what intentionalism and the imagistic theory of meaning do, albeit in different ways.

3. Inverting the Spectrum

The believer in NECESSITY faces a more serious difficulty. Let CONTINGENCY be the thesis that both COLOR REP \rightarrow COLOR FEEL and COLOR FEEL \rightarrow COLOR REP (and hence NECESSITY) are false. According to CONTINGENCY there is at best a contingent connection between green-feeling experiences and green-representing experiences. John Locke has described a case which is the basis of an argument for CONTINGENCY. We are to suppose, according to Locke (1689, II, xxxii, 15), that:

...by the different Structure of our Organs, it were so ordered, that *the same Object should produce in several Men's Minds different Ideas at the same time; v.g., if the Idea, that a Violet produced in another Man's Mind by his Eyes, were the same that a Marigold produced in another Man's, and vice versa.*

Locke is suggesting that we can conceive of a subject who has yellow-feeling experiences when looking at violets, and blue-feeling experiences when looking at marigolds. Since this state of affairs is conceivable, we have a positive, if defeasible, reason to believe it to be possible.

Let Nonvert be a normal subject—Nonvert enjoys blue-feeling experiences when looking at violets and yellow-feeling experiences when looking at marigolds. Invert, however, is phenomenologically inverted in the way that Locke describes. Invert has yellow-feeling experiences when looking at violets and blue-feeling experiences when looking at marigolds. In normal subjects blue-feeling and blue-representing experiences coincide. Consider now Invert's visual experience of a violet. It is a yellow-feeling experience, but is it also yellow-representing? Or are Invert's yellow-feeling experiences, in fact, blue-representing? There are at least four lines of thought that suggest that Invert's yellow-feeling experiences are blue-representing.

(a) The color perceptions of normal subjects are mostly veridical. Thus when Nonvert, by hypothesis a normal subject, visually represents violets as blue he is not suffering from any illusion—violets do in fact have the color that Nonvert's visual experience attributes to them. The reason that the blue-representing experiences of normal subjects are mostly veridical is that they reliably track blue things, and it is

because they are standardly caused by the instantiation of blueness that such experiences are blue-representing. Invert's yellow-feeling experiences reliably track blue things. Because Invert's yellow-feeling experiences are standardly caused by the instantiation of blueness, Invert's yellow-feeling experiences are, in fact, blue-representing.

(b) Invert and Nonvert are both competent English speakers. They at least apply their color words to the same range of things. It is plausible to suppose that Invert means by "blue" what the rest of us mean. Invert describes violets as looking blue to him and as not looking yellow. If, however, Invert's yellow-feeling experiences were also yellow-representing, then Invert would be radically mistaken about how things look to him. But surely this is implausible. We might occasionally be wrong about how things appear to us, but surely we could not be *systematically* mistaken about the appearance of things. And so Invert's yellow-feeling experiences are, in fact, blue-representing.

(c) Invert's experience of violets is yellow-feeling. Violets are in fact blue, and so if Invert's experience were also yellow-representing, his perceptual experience would be illusory. To claim that Invert's yellow-feeling experiences are yellow-representing would be to attribute widespread and systematic error to Invert's perceptual experience. The principle of charity, however, is partly constitutive of our practice of interpretation. When interpreting the thought and talk of others we are rationally constrained to ascribe mostly true contents. And so too for perceptual contents. We are therefore rationally constrained to interpret Invert's yellow-feeling experience as blue-representing.

(d) Suppose that Invert's yellow-feeling experience when looking at a violet is yellow-representing. If violets are in fact blue, then Invert's experience is, at least in part, illusory. Invert, however, can use his color vision to successfully navigate through the world. Given his success when dealing with his environment as visually represented, there is, at the very least, a *presumption* that Invert is not systematically misrepresenting the colors of things. And so Invert's yellow-feeling experiences are, in fact, blue-representing.

If any of these arguments are persuasive, then one could argue for CONTINGENCY as follows. Consider Invert looking at a violet. Invert's visual experience of a violet is yellow-feeling. But his experience represents the violet as blue. So Invert's experience is yellow-feeling and blue-representing, and Nonvert's experience is blue-feeling and blue-representing. There thus could be two subjects whose visual experience shared the same total content but whose color phenomenology

differed. And hence COLOR REP \rightarrow COLOR FEEL fails, and NECESSITY is false. Consider Invert looking at a marigold. Invert's visual experience of a marigold is blue-feeling but is also yellow-representing. Nonvert, however, when looking at a violet has a blue-feeling experience that is also blue-representing. There thus could be two subjects with the same color phenomenology, but whose experience differed in content. And hence COLOR FEEL \rightarrow COLOR REP fails, and NECESSITY is false. Since both COLOR REP \rightarrow COLOR FEEL and COLOR FEEL \rightarrow COLOR REP are false, CONTINGENCY is true.

If NECESSITY is false and CONTINGENCY true, then intentionalism is false as well. If there could be a difference in color phenomenology without a difference in content, then there could be no question of identifying phenomenal properties with representational features of experience. We do not believe, however, that content phenomenology can vary independently.

4. The Inadequacy of Existing Formulations

We should emphasize that the inverted spectrum argument comes in two parts. The advocate of CONTINGENCY must establish:

- (1) that there could be two subjects, Nonvert and Invert, whose color experience is phenomenologically inverted, and
- (2) that there would be no relevant difference in content between their perceptual experiences.

(1) is underwritten by an initial conceivability claim. To complete the argument, however, the advocate of CONTINGENCY must then *argue* for (2)—that the perceptual experience of phenomenologically inverted subjects would share the relevant content.¹¹ Unfortunately none of the arguments for (2) are cogent.¹²

(a) One reason for believing that Invert's yellow-feeling experience would also be blue-representing was that a causal theory must be the correct account of the determinants of color content. It is because Invert's yellow-feeling experiences are standardly caused by the instantiation of blueness that they are also blue representing. Notice, however, this argument *assumes* that Invert's experience could be yellow-feeling independently of its being yellow-representing. But this is supposed to be a *consequence* of the inverted spectrum argument, not a premise for it. Within this broader dialectical context, the argument, in effect, presupposes what it seeks to establish. So formulated the inverted spectrum argument is question begging and so should be rejected.

(b) Another reason for believing that Invert's experience of a violet would be blue-representing, if yellow-feeling, was that to suppose otherwise would allow for the possibility that Invert is systematically mistaken about the appearances of things. We are supposing that Invert is a competent English speaker and would apply his color words to the same range of things as Nonvert. Invert would, like Nonvert, describe violets as looking blue to him and as not looking yellow. If when looking at a violet his experience were nonetheless yellow-representing, Invert would be radically mistaken about how the violet appears to him. The central mistake of this argument is its conflation of public language meaning with mental content. If Invert were to conform his usage to his linguistic community's, then he would mean by "blue" what Nonvert means. If he were to then describe violets as looking blue to him, he would, unbeknownst to him, be *misdescribing* his experience *in English*. But that is *not* to say that the character of his experience is introspectively unavailable to him, only that he does not fully understand the meanings of his public language color words. Suppose Invert were to use his color words to designate whatever color property is manifest in the corresponding color experience. Then by "blue" he would mean yellow, and hence something different from Nonvert's use of "blue". If this were the case, then he would *correctly* describe violets as looking blue to him *in his idiolect*. In either case, to say that Invert's yellow-feeling experience would also be yellow-representing is not to say that he is mistaken about the appearances of things.

(c) Another reason for thinking that Invert's yellow-feeling experiences of must also be blue-representing was that the contrary supposition would violate the principle of charity. Violets are in fact blue, and so if Invert's experience of blue things were yellow-representing, he would be systematically mistaken about the colors of things. But it is a consequence of charity that the colors that vision attributes to objects are properties that they generally have. So the claim that Invert's yellow-feeling experiences are also yellow-representing would violate a constitutive principle of our interpretive practices. We shall not here question whether charity, suitably understood, really is a constitutive principle of interpretation. It suffices to point out that even if it were, the envisaged application of such a principle would be a *misapplication* of it. Charity applies, if it applies at all, to an *entire* system of representation and not to some proper part of it. An interpretation may be charitable and still ascribe widespread error to some large region of the system. (See Boghossian and Velleman, 1989.)

(d) A further reason for believing that Invert would correctly represent the color of his environment was Invert's adaptive success. We are assuming, for the sake of argument, that Invert successfully navigates through the world by means of his color

vision and are asked to conclude on this basis that his color experience must be mostly veridical. There is, however, no direct inference from Invert's success when dealing with his environment as visually represented, to his color experience being mostly veridical. The argument is supposed to be an instance of inference to the best explanation. But it is by no means clear that the veridicality of Invert's color vision is the *best* explanation of his successful action. It is unclear that his adaptive success might not be explained at least as well by his tracking features of the environment by means of a systematic *illusion*. (Indeed, projectivists are committed to the cogency of such an explanation.)

None of the arguments for (2) that we have considered are both sound and convincing. So these versions of the inverted spectrum argument fail. Of course, this is not yet to establish that there could be no cogent formulation of the inverted spectrum argument—only that the most familiar formulations are inadequate. But merely to claim that there might be an adequate version of the inverted spectrum argument is not yet to give an argument for CONTINGENCY. Nevertheless we shall argue that there could be no sound and convincing version of the inverted spectrum argument. Or more accurately, that if a particular account of color content is correct, then no cogent version of the argument could be forthcoming.

5. The Grammar of Our Color Concepts

Among the abilities involved in color vision is the ability to treat visible objects as more or less similar in respect of color. One manifestation of this is the consequent disposition to sort visibilia into classes given their color similarities. This ability is conceptually primitive in the sense that it requires neither the possession of any sophisticated concepts nor a theoretical understanding of color. Thus bees, for example, display this kind of behavior as well as human beings. It is universal in the sense that normal subjects display nearly the same set of behavior-guiding reactive dispositions. For example, a standard test of color vision, the Farnsworth-Munsell 100 hue test, asks subjects to arrange 100 colored samples in terms of color similarity. One sample is fixed and the rest must be arranged in a circle beginning and ending with the fixed sample. Normal subjects make very nearly the same arrangement, and abnormal subjects differ from normal ones in predictable ways. The ability is also universal in another sense. *All* visible objects find a place in the structure of similarities and differences determined by our visual reactions and exhibited in our classificatory behavior.

Our behavioral disposition to sort things by color is grounded in our disposition to visually classify them into similarity classes; and it is plausible to suppose that this

disposition is *itself* grounded in our visual system's propensity to treat the colors themselves as more or less similar. The perceived color similarities among objects are determined by the colors they are visually classified as instantiating and the perceived similarities among the colors themselves. Red things, for example, look similar to orange things in part because red itself looks similar to orange. Furthermore, red things look similar to purple things because red itself looks similar to purple. Red occupies a unique region of the space determined by the relations of similarity in hue, saturation, and brightness.¹³ What unites the elements of a similarity class for the visual system is the relations the similarity class is perceived to bear to other relevant similarity classes.

It is well known that the *internal* relations of similarity and difference among perceived colors can be represented by *external* relations of distance in a space, and that an adequate representation requires a space of three dimensions—one dimension each for hue, saturation and brightness. A complete, empirically adequate color space is an *exhaustive* representation of the similarities in hue, saturation, and brightness—all and only such similarities are represented by the structure of the color space. An adequate representation of the internal relations of similarity and difference suffice to represent the identity and distinctness of perceived color and their order of determination. Roughly speaking the points of the space represent minimal differences that normal subjects can perceptually distinguish. So the identity and distinctness of particular shades can be represented by the identity and distinctness of points in the color space. The order of determination among perceived colors can be represented by mereological relations among regions of the color space. A particular shade of red is represented as a determinate of the determinable red by the particular shade's location being a proper part of the region associated with the more general property. Red and orange, though not specific shades, are represented by distinct regions of the color space. That they are codeterminates is represented by the fact that they are non-overlapping parts of the broader region of the color space that represents the property of being reddish. A representation of the internal relations of similarity and difference suffice in this way to represent the identity and distinctness of color properties and the structure of determinables and determinates in which they stand.

When psychophysicists describe the structure of the color space they take themselves to be describing a biologically-determined classificatory function of the visual system, one that, in effect, sorts similarity classes in terms of their comparative similarity in hue, saturation, and brightness. And the present suggestion is that the similarity classes are determined, in part, *by* this antecedent classificatory function. The human visual system determines a network of higher-order comparative similarities.

Red is experienced as more similar to orange than to yellow. Red is also experienced as being more similar to purple than to blue. Indeed, there are indefinitely many such relations of comparative similarity that obtain among red and other pairs of colors. Our proposal is that the content of red-representing experiences is partly determined by *all* the relations of comparative similarity that red is represented as standing in to the other colors. The content of a color experience is partly determined by its position in the psychological color space. What makes a color experience an experience of a particular color is the similarities and differences, determined by the character of our visual system, the represented color stands in to the other colors. There are a multitude of relations (taken in extension) that obtain among classes of visible objects. Normal subjects are equipped with an innate, biologically-determined set of similarity relations that selects from among the multitude those relations that *count as* similarities in hue, saturation, and brightness. The relevant similarity classes are determined, in part, by this classificatory function. *To be* similar in respect of color is *to look* similar. And in so determining what counts as color similarity, visual experience comes to represent a given color property.

It is consistent with the present account that the relations of similarity in hue, saturation, and brightness supervene on color properties that exist independently of viewing subjects. In speaking of the visual system of *selecting* these relations we are not claiming that visual experience *brings them into being*. We are merely claiming that a pre-existing relation *only counts as* similarity in hue, saturation, or brightness in virtue of an antecedent classificatory function of the visual system. Facts about visual experience fix which of the similarities among certain objective properties *count as* similarities in respect of color. But that is consistent with higher-order color similarities, *so understood*, supervening on objective color properties. The properties represented by color experiences are ontologically independent of those experiences as are the relations of similarity and difference between them. What is determined by the characteristics of our visual system is which properties are represented in color experience and what relations of similarity and difference are perceived to obtain among those properties. Color similarities are thus *anthropocentric* in something like Hilbert's sense of the term (Hilbert, 1987). But being anthropocentric makes color similarities neither less real nor less objective.¹⁴

So far we have claimed that color experience denotes color properties by determining a similarity class. We have further claimed that this is accomplished, in part, by an antecedent classificatory function of the visual system that selects which relations among these classes are to count as similarities in hue, saturation, and

brightness. Mark well we say *in part*. All that the visual system determines is that the similarity classes jointly exhibit the structure of the color space. And it's possible that *different* systems of similarity classes jointly exhibit the *same* structure. The similarity classes, the extensions of our perceptual color categories, are also determined by relations normal subjects bear to their native environment. An experience's position in the color space partly determines its content by determining constraints governing its extension. If color properties are associated with unique regions of the color space, then this pattern of similarity and difference must be preserved when that property's extension is settled by relations born to the subject's environment. Whatever determinate extension is assigned to red, for example, it cannot overlap with the extension assigned, say, to green. Similarly these structural constraints also determine that whatever extensions are assigned to our perceptual color categories, it can never be false that red is more similar to orange than to blue.¹⁵

The content of color experience is determined by two sorts of factors:

- (A) function in the color space, and
- (B) relations born to the subject's environment.

The structural constraints derived from similarity relations do not *wholly* determine color content. Position in the color space will not by itself determine the members of the similarity classes. But where the relations of similarity and difference among the colors leaves the denotations of our color experiences indeterminate, the relations that a normal subject bears to his environment resolves this indeterminacy. It is important to recognize that we are not ascribing two contents to color experience—one determined by external relations and the other determined by internal similarity relations. We are not proposing a *dual aspect semantics* according to which color experience has a wide and narrow content; rather, our color experience has a unitary content determined by two sorts of factors. Monadic color contents, in effect, supervene on a complicated set of dispositional relations some of which are born to the environment, others of which are born to actual and potential color experiences.

Color contents so determined are *Fregean*. In describing color contents as Fregean, we only mean that they are not exhausted by the designated color property. The content of a green-representing experience, for example, is not completely characterized by its designating green. We do not mean to claim, in addition, that color contents are identical to descriptive contents or clusters of descriptive contents or play any of the other roles traditionally associated with Fregean senses. Our visual

experiences refer to color properties—a green-representing experience refers to the color green—it is just that their content is not exhausted by such reference.

How is Fregeanism a consequence of the present view? If color contents were *Millian*, if they were exhausted by the designated color property, then they could not determine the constraints that govern our color experience. While position in the color space does not completely determine the denotations of our color experiences, it does impose constraints on how these denotations are settled by relations born to the subject's environment. In determining the extensions of our color contents, the relevant external relations are thus constrained to select collections of things that jointly exhibit the structure of the color space. If the denoted color were all there were to the content of our color experience, then such a content would determine no such constraint.

The analysis of color contents that we propose is importantly analogous to Fregean conceptual role semantics and shares an important feature with such approaches to semantic theory. A Fregean *conceptual role* semantics has one clear advantage over traditional Fregean theories. Fregean contents, say of names in public language, are often thought to explain the conceptual role of such expressions. In particular the content of a name, its "sense", is thought to explain and predict the pattern of substitution failure occurring within propositional attitude constructions. But the content of a name could only explain and predict the pattern of substitution failure if one could specify that content independently of such substitution failures. Unfortunately, Fregeans have either failed to provide such an independent specification or have provided an inadequate one (in terms of descriptive contents, for example). If, however, content supervenes on conceptual role, this problem is *dissolved*. According to conceptual role semantics, the apparent difficulty gets the order of explanation wrong. It is *not* that content *explains* conceptual role; rather, conceptual role *explains* content. And it might be held that the content of a name cannot be specified independently of the pattern of substitution failure because its content is determined in part *by* such substitution failures. Similarly on our view of color content it is not that color contents explain location in color space but rather that location in color space (partly) explains color content. Thus our Fregean color contents cannot be specified independently of the role that the relevant experiences play in the psychological color space. There is no saying in other terms what such Fregean contents are. We should emphasize that we are only committed to *color* contents being Fregean. We remain silent on whether other mental contents or the contents of public language expressions are also Fregean.

It would be a fatal consequence for the present account if it turned out that a subject is automatically credited with knowledge of all the similarity and difference

relations constitutive of his color space on the basis of a single color experience. Boghossian and Velleman (1991, pp. 128–30) have charged that an account superficially similar to the present one is guilty of such a consequence. On the account they consider, the content of a color experience is specified by a description of the color's location in the quality space. Since the content of the experience is given by a description that explicitly refers to the structure of the color space, the subject of such an experience would know without further reflection the full extent of that space. Notice, however, that this is *not* a consequence of the present view. Though an experience's position in the color space partly determines its content, it is not because its content is given by a description of its location. What partly determines color content is an experience's *location* in color space and not its *describing* the color space or its location in it.

Furthermore, while the correct location in color space of a visual experience may partly determine its content, that is by itself no guarantee that the subject automatically knows or has any reasonable beliefs about that function or even that he has the concepts required to describe it. When a subject undergoes, say, an orange-representing experience, he will usually know certain *immediate* consequences of its having a particular position—that orange is yellowish, for example. But that is not to say that the subject *automatically* knows the *complete* pattern of phenomenal similarity and difference that characterizes orange-representing experiences. He would have to reflect on the contents of *further* experiences in order to explicitly recognize that certain browns, say, are low-grade oranges. That an experience's position in the color space is content-determining only guarantees that its pattern of phenomenal similarity and difference is *a priori discoverable* (that is, discoverable purely by reflecting on the contents of visual experience). But the *explicit* recognition of certain content-determining similarities may require that the subject reflect on the contents of further, contrastive experiences.

So far we have been discussing the nature of color content, but what of its relation to phenomenology? What of NECESSITY? Not only do we believe that fixing the *content* of visual experience settles its subjective character; we further believe that some *particular aspect* of color content is responsible for color phenomenology. That the visual experiences of normal subjects share the same color phenomenology is, of course, a matter of their being relevantly similar. We would suggest, however, that the relevant dimension of similarity consists in the fact that the contents of their respective experiences are partly determined by their occupying the same region of the psychological color space. NECESSITY, then, is underwritten by a substantive identification—the phenomenal properties of color experience are identified with the structural constraints governing them (as determined by their position in the color

space). The phenomenal similarities among color experiences *just are* perceived similarities among the colors.

We should remark that this is less a worked out theory of color content than a sketch of one. Much more would have to be said to fully defend such a view. It does have, however, two notable consequences that are important to what follows. First, the structure of the color space partly determines the content of our perceptual color categories. Second, it is conceptually necessary that the identity of perceived color is given by position in the color space. Given these operative assumptions, we shall argue that no version of the inverted spectrum argument could be sound.

6. Reinverting the Spectrum

Recall that the inverted spectrum argument came in two parts. The advocate of CONTINGENCY must make out two things:

- (1) that there could be two subjects, Nonvert and Invert, whose color experience is phenomenologically inverted, and
- (2) that there would be no relevant difference in content between their perceptual experiences.

If the present account of color content is correct, then anyone who would argue in this way faces a dilemma. If the structure of a quality space were asymmetrical, then (2) would be false—there would be a relevant difference in content. If, however, it were symmetrical, then (1) would be false—phenomenological inversion would not be genuinely possible. Or so shall we argue.

6.1 *Asymmetry and Content Inversion*

We are assuming that position in the color space partly determines the content of color experience. Given this assumption, we shall argue that if the color space is asymmetrical, then there would be a relevant difference in content between phenomenologically inverted subjects.

The psychological color space of normal subjects is *irregular*. There are, notoriously, asymmetries in our psychological color space. Consider the following two asymmetries that occur along the yellow/blue axis of the color space. Blue continues to appear blue at all levels of brightness. Light blue and navy blue, for instance, each appears blue to us. A decrease in the level of brightness does not affect the hue of blue-feeling experiences. Not so for yellow-feeling experiences. Yellow-feeling experiences have a comparatively high level of brightness. If the level of brightness is sufficiently

decreased, the hue will shift. While darkened blues (such as navy blue) continue to look blue to us, darkened yellows (that is, browns) do not continue to look yellow. The phenomenal similarity between blue-feeling experiences at high and low levels of brightness is much greater than the phenomenal similarity between yellow- and brown-feeling experiences. Indeed the phenomenal *difference* between yellow- and brown-feeling experiences is so great as to suggest a shift in chromatic content. Not only do yellow-feeling experiences only occur at comparatively high levels of brightness when compared to blue-feeling experiences, but the brightest yellow-feeling experience is much brighter than the brightest blue-feeling experience. There is consequently no blue-feeling experience at the corresponding level of brightness as the brightest yellow-feeling experience.

What do such asymmetries tell us about pairs of spectrally inverted subjects? One consequence is that the phenomenological difference between Nonvert and Invert would be in principle detectable. A psychophysicist could determine, for instance, that marigolds appear brighter to Nonvert than to Invert. But that the phenomenological inversion be undetectable was no part of the argument for CONTINGENCY. Our present concern is less with *epistemology*—with whether we could know, or reasonably believe, that another’s color experience is phenomenally similar to our own—than with *metaphysics*—with whether the phenomenal and representational features of experience are necessarily connected and so may be identified.

More germane to the present purpose is the *functional* difference between Invert and Nonvert if our color space is asymmetrical. If our color space is asymmetrical, then Invert and Nonvert could not be functional duplicates. The same stimulus would produce a visual representation of color with distinct functional roles in Invert and Nonvert. Invert’s and Nonvert’s visual experience of a marigold differ, for instance, in their pattern of phenomenal similarity and difference. Nonvert’s visual experience of a marigold is less phenomenally similar to his experience of mahogany, than Invert’s corresponding experiences of marigolds and mahogany. Whereas Nonvert’s visual experiences of marigolds and mahogany differ in chromatic content (they are yellow-feeling and brown-feeling, respectively), Invert’s visual experiences of marigolds and mahogany are uniformly blue-feeling. If our psychological color space is asymmetrical, then Invert and Nonvert could not be functional duplicates.

While the argument for CONTINGENCY does not assume that Invert and Nonvert *could be* functional duplicates, that they *could not be* is nonetheless relevant given what we believe to be the determinants of color content. We suggested that *part* of what makes it the case that a normal subject’s visual experience of a violet is blue-

representing is the role that experience plays in the subject's psychological color space. Part of what makes an experience blue-representing, for instance, is its opponent opposition to yellow-representing experiences. If Invert and Nonvert differ in the functional organization of their color spaces, and if among the determinants of color content is an experience's role in the subject's color space, then Invert's and Nonvert's visual experience of a violet must differ in color content. Nonvert's blue-feeling experience of a violet is blue-representing, but Invert's yellow-feeling experience of a violet could not also be. So the claim that NECESSITY is false because there could be a relevant difference in phenomenology without a relevant difference in content fails. Moreover, Nonvert's blue-feeling experience of violets are blue-representing, but Invert's experience of marigolds while blue-feeling could not be yellow-representing. So the claim that NECESSITY is false because there could be a difference in content without a corresponding difference in phenomenology fails.¹⁶

6.2 Symmetry and Phenomenological Inversion

A standard response to the appeal to color asymmetries is that the conclusion of any such argument must fall short of its mark. The most that empirically discovered asymmetries could establish is that there are no actual cases of quality inversion without a relevant difference in content and not that such cases are *impossible*. For we can certainly conceive of subjects capable of differential color discriminations whose psychological color space is *symmetrical*. (See Shoemaker, 1981 and Block, 1990.) If the only obstacle to there being a relevant difference in content between Nonvert and Invert is the asymmetry of our color space, then there could be a relevant difference in content between phenomenologically inverted subjects with symmetrical color spaces.

The advocate of CONTINGENCY who would respond in this way now faces the second horn of the dilemma. We are assuming that the identity and distinctness of perceived color (the phenomenal character of color experience) is exhaustively represented by the identity and distinctness of the points in the color space. Given this assumption, we shall argue that if the color space is symmetrical, then phenomenological inversion is impossible--there would be no phenomenal difference between Nonvert's and Invert's experience on which to contrast the supposed inversion. Not only is phenomenological inversion not possible for creatures with symmetrical quality spaces, but every possible quality space *must* be asymmetrical. The case comes in two parts. We shall argue that the standard conceivability arguments that underwrite belief in the possibility systematic quality inversion fail if the quality space is symmetrical. If these are our reasons for believing in the possibility of

phenomenological inversion, then we are not entitled to this belief. We shall further argue that *there could be no* adequate description of quality inversion in a symmetrical quality space. The difficulty is that there could be no phenomenal difference between inverted subjects with symmetrical quality spaces. In order for an experience to have a distinctive phenomenal character, and indeed to have a phenomenal character at all, the associated quality space *must* be asymmetrical.¹⁷

To avoid the first horn of the dilemma it is enough that there could be a pair of subjects, $\text{Nonvert}_{\text{SYM}}$ and $\text{Invert}_{\text{SYM}}$, whose color space has a *single* symmetry axis along which they are phenomenologically inverted. Since the quality inversion occurs along a symmetry axis, such an inversion is a *symmetry transformation*—a mapping of the color space onto itself by means of a continuous rigid motion. To dramatize the difficulty in conceiving of this we shall argue that it is impossible to conceive of pairs of phenomenologically inverted subjects whose color space is *perfectly* symmetrical (or *supersymmetrical*). The root difficulty for each is the same, but it is easier to see in the more radical case.

There is an important difference between the Lockean conceivability claim and the present invitation to conceive of phenomenologically inverted subjects with supersymmetrical color spaces. We know from our own case what it is like to undergo experiences of violets and marigolds. We know that our visual experience of violets are blue-feeling and that our visual experience of marigolds are yellow-feeling. We certainly can imagine that our experience of violets might be yellow-feeling and that our experience of marigolds might be blue-feeling, and it is on the basis of this that we feel confident in our ability to conceive of pairs of subjects that are phenomenologically inverted in the way that Locke describes. And it is because we can conceive of pairs of phenomenologically inverted subjects that we believe such pairs to be possible (even if, like us, you doubt that the visual experience of each could have identical contents).

We can conceive of the Lockean case since we can imagine violets looking like marigolds and *vice versa*. We can then imagine a subject for whom this phenomenological inversion is *systematic*. But we can engage in no similar imaginative exercise with respect to the color experience of subjects with perfectly symmetrical color spaces. If color content is partly determined by position in a color space, then $\text{Nonvert}_{\text{SYM}}$'s and $\text{Invert}_{\text{SYM}}$'s color contents differ from our own. We cannot then take even the first step in the imaginative exercise that leads to the alleged conceivability of the phenomenological inversion. For we have no idea whatsoever of how colors appear to $\text{Nonvert}_{\text{SYM}}$. And if we have no idea of how colors appear to $\text{Nonvert}_{\text{SYM}}$ then we have no idea of the phenomenal character of such appearances. We cannot imagine

what the subjective character of Nonvert_{SYM}'s experience of a marigold would be, and so cannot imagine what it would be for violets to look like marigolds to Nonvert_{SYM}. And if we cannot conceive of this, neither can we conceive of a subject, Invert_{SYM} for whom this inversion is systematic. So we can neither imagine the phenomenal character of Nonvert_{SYM}'s color experience nor Invert_{SYM}'s. So in what sense are pairs of phenomenologically inverted subjects with perfectly symmetrical color spaces even conceivable?

It might be objected that we can indeed imagine phenomenologically inverted subjects with symmetrical quality spaces. Begin with our own, admittedly, asymmetrical color space. Now imagine a symmetrical subregion of this space and suppose that this subregion could be the entire color space for some possible subject. Since we know how the colors in the subregion appear, we know how such colors would appear to Nonvert_{SYM}.

We know how such colors appear, however, because the identities of the colors in the subregion are given, in part, by the similarity relations they bear to the broader embedding region. Once the pattern of similarities and differences are restricted in the way imagined, then we no longer know how such colors would appear, for they would be different colors. To see this consider the "missing shades"—the colors in the broader embedding region. We could have been raised in an environment such that the only colors with which were perceptually acquainted were the colors of the target subregion. Even given this chromatic impoverishment, we certainly could have imagined a missing shade of blue, call it "Hume blue". We could have imagined Hume blue because it would have participated in a pattern of similarity and difference with the colors of the subregion. Given the structure of his color space, Nonvert_{SYM}, however, could not imagine Hume blue. If Nonvert_{SYM} could imagine Hume blue, it would be located within his color space. The perceived shades aligned along the points of his color space do not participate in a pattern of similarity and difference that includes Hume blue. This could only be the case if they were different from the colors that occur within the subregion of *our* color space. Imagining a world that instantiates only the colors within the subregion is not yet to imagine what our visual experience would be like if the subregion were our entire color space.

Perhaps to imagine a symmetrical color space – all we need do is imagine a two-quality color space. After all there is no great difficulty in conceiving of what it would be like, for example, to see the world in black and white. We know what black looks like, and we know what white looks like. So, *a fortiori*, we know what it would be like to see the world *only* in black and white. Despite its initial plausibility, nothing new has

been accomplished by the present imaginative exercise. It is just a degenerate version of the symmetrical subregion case discussed above. The problem once again is that the target state of affairs has not in fact been conceived. Imagining a black and white world is not to imagine what visual experience would be like for a creature whose color space consisted entirely of these two qualities.

When advocates of CONTINGENCY claim that symmetrical color spaces are conceivable, they might have in mind something like the following. Simply imagine the perceived colors arrayed on the points of a sphere. As there are infinitely many symmetry axes in a sphere, a spherical color space would itself be symmetrical. However, a moment's reflection will reveal that this is an inadequate means to conceive the target state of affairs. The contrary supposition turns on a conflation of geometric and chromatic structure. The points on the surface of the sphere are, for example, geometrically but not chromatically continuous. When we imagine the colors arrayed on a sphere, we picture it as striped much as a rainbow is (or, better yet, a beach ball); but this is no consequence of geometrical structure. And what this shows is that imagining the colors arrayed on the points of the sphere is to imagine, *inter alia*, the identity and distinctness of perceived color as being given independently of the geometrical structure. But this is just to imagine a colored spherical space and not yet to imagine a spherical color space.

We have yet to discover a satisfactory means to imagine symmetrical color spaces. And if they are not conceivable, with what right do we believe them to be possible? We suspect that some advocates of CONTINGENCY would remain unconvinced.¹⁸ To these we would suggest that such considerations show, *at the very least*, that the case of symmetrical quality inversion is crucially underdescribed—especially given its role in the *argument* for CONTINGENCY. Not only do we believe that to merely claim that there could be subjects with symmetrical color spaces is to underdescribe the case, but that no adequate description is possible. In a perfectly symmetrical quality space the pattern of phenomenal similarity and difference of any given point in that space is exactly the same as the pattern of similarity and difference of any other point. In particular, the set of relations that a point bears to every other point in the color space will be structurally isomorphic to the set of relations associated with the corresponding point along a symmetry axis. What gives lie to its putative conceivability is that, unlike the Lockean case, there is no distinction in the *pattern* of phenomenal similarity and difference with which to contrast the supposed inversion. If an experience's phenomenal character is determined by the pattern of phenomenal similarity and difference in which it participates, then a visual experience occurring

along any point in the symmetrical quality space will have the same phenomenal character as any other visual experience. Daniel Dennett (1993. p. 927) has vividly expressed this thought:

What *anchors* our naive sense that there are such properties as qualia are the multiple, asymmetrical, interdependent sets of reactive dispositions by which we acquaint ourselves with the sensible world. Our sense that the color red has, as it were, an identity, a “personality” all its own is *due* to the host of *different* associations that go with each color. Shoemaker’s envisaged creatures, lacking all such reactive landmarks in their dispositional make-up, would not think that they had qualia at all—what it was like to have one sort of experience would not differ at all from what it was like to have a different one!¹⁹

If Dennett is right, then Nonvert_{SYM}’s and Invert_{SYM}’s visual experience of an object under fixed viewing conditions would not differ in phenomenal character. Of course we cannot merely assert this—that would be question begging. But elementary reflections on the nature of a supersymmetrical quality space suffice to establish this conclusion. Recall, a complete, empirically adequate quality space is an *exhaustive* representation of the phenomenology associated with a given sensory modality. The phenomenal character of an experience aligned on a point of that space is represented by its position in the quality space. But if the points in a quality space are relationally individuated, then there is no way to individuate one point from another. Consider a geometrical sphere. There is no way to individuate the points on the surface of a sphere *purely* in terms of their relations to one another. And the situation with respect to a supersymmetrical quality space is strictly analogous.²⁰

It would do no good to claim that phenomenal properties are *intrinsic* features of experience and so are in no need of a *relational* characterization in terms of phenomenal similarity. Belief in intrinsic, non-representational features of experience that determine its qualitative character is supposed to be a consequence of the inverted spectrum argument. If the inverted spectrum argument is one’s reason for believing in qualia, one could not then rely on the existence of such features as a premise.

Worse still, there *could be no* intrinsic difference in the points of a supersymmetrical quality space. That an experience instantiates a quale, if such there be—that it has an intrinsic phenomenal character—is supposed to determine the pattern of phenomenal similarity in which that experience participates. But if the pattern of phenomenal similarity constitutive of color phenomenology is exhaustively represented by a perfectly symmetrical color space, then each point in that space must be

intrinsically like any other. It is only on *this* basis that the symmetry of the color space could be determined by intrinsic features of experience. To distinguish the phenomenal character of an experience, such an intrinsic feature would have to determine a distinctive pattern of phenomenal similarity. But by hypothesis, no point in a perfectly symmetrical quality space participates in a pattern of phenomenal similarity different from any other point. So not only is there an *epistemological* problem in appealing to an intrinsic difference, but there is also a *metaphysical* problem in that each point of the color space must be an intrinsic duplicate.

It is worth emphasizing that this is a consequence of the fact that the color space suffices to represent the identity and distinctness of perceived color and their order of determination. To see this consider the following objection. If two points of a supersymmetrical color space were picked out in some way other than their position, then we could individuate all the points from one another in terms of their distance relations plus the antecedent identification of the two points. Thus, the objection continues, there could be intrinsic differences between the points of a supersymmetrical quality space. The observation upon which the objection rests is surely correct, but the conclusion that there could be intrinsic differences only follows (if it does at all) if position in the quality space does *not* suffice to represent the identity and distinctness of phenomenal properties. If position in a quality space suffices to represent the identity and distinctness of phenomenal properties, then the points in a supersymmetrical quality space must be intrinsic duplicates.

To summarize, avoiding the first horn of our dilemma depends on our putative ability to conceive of pairs of subjects with symmetrical color spaces that are phenomenologically inverted with respect to one another. And we have argued that we cannot, in fact, conceive of this state of affairs and so have no reason to believe it to be possible. As Dennett apparently believes, any possible quality space *must be* asymmetrical.

7. Conclusion

The inverted spectrum argument has been one of the greatest obstacles to intentionalism. If the experience of two subjects could differ in color phenomenology without differing in content, then there could be no question of identifying phenomenal properties with intentional features of experience. Intentionalist reductions of the phenomenal are only possible if the inverted spectrum argument is unsound. If the present account is correct, then the resources are on hand to establish this. The inverted spectrum argument comes in two parts. There is an initial conceivability claim that

underwrites belief in the possibility of systematic quality inversion. To complete the argument, however, the advocate of CONTINGENCY must then argue that there could be phenomenological inversion without a relevant difference in content. We have argued that the extant versions of the inverted spectrum argument have failed to make a sound and convincing case for there being a relevant difference in content between Nonvert and Invert. We have further argued that if the present account of color content is correct, then the advocate of the inverted spectrum argument faces a dilemma: if color space were asymmetrical, then there would be no relevant difference in content; if color space were symmetrical, then quality inversion would be inconceivable. The present account, if true, thus goes a long way to vindicating the truth of intentionalism.

Acknowledgments

We wish to thank Fiona Cowie and Mark Johnston for helpful conversations on these issues. Jim Pryor provided exceptionally helpful comments on an earlier draft for which we are very grateful.

References

- Block, N. (1978). Troubles with Functionalism. In C.W. Savage (ed.), *Perception and Cognition. Issues in the Foundations of Psychology, Minnesota Studies in the Philosophy of Science*, vol. 9, 261-325. Minneapolis: University of Minnesota Press. Reprinted in Block (1980), pp. 268-305.
- Block, Ned, ed. (1980). *Readings in the Philosophy of Psychology*, vol. 1. Cambridge, MA: Harvard University Press.
- Block, N. (1990). Inverted Earth. *Philosophical Perspectives* 4: 53–79.
- Boghossian, P. and D. Velleman (1989). Color as a Secondary Quality. *Mind* 98, 81–103. Reprinted in Byrne and Hilbert (1997a, pp. 81–103).
- Boghossian, P. and D. Velleman (1991). Physicalist Theories of Color. *Philosophical Review* 100: 67–106. Reprinted in Byrne and Hilbert (1997a, pp. 105–136).
- Byrne, Alex and D. Hilbert, eds. (1997a). *Readings on Color, Vol. 1: The Philosophy of Color*. Cambridge, MA: MIT Press.
- Byrne, A. and D. Hilbert (1997b). Colors and Reflectances. In Byrne and Hilbert (1997a, pp. 263–288).
- Clark, Austen (1993). *Sensory Qualities*. New York: Oxford University Press.
- Dennett, D. (1993). The Message is: There is No Medium. *Philosophy and Phenomenological Research* 53: 919–931.
- Dretske, Fred (1995). *Naturalizing the Mind*. Cambridge, MA: MIT Press.
- Evans, Gareth (1982). *The Varieties of Reference*. New York: Oxford University Press.
- Firth, R. (1949–1950). Sense-Data and the Percept Theory. *Mind* 58, 59: 434–465, 35–56. Reprinted in Swartz (1965, pp. 204–270).
- Grice, H. P. (1962). Some Remarks about the Senses. In R.J. Butler (ed.) *Analytic Philosophy, First Series*. Oxford: Basil Blackwell. Reprinted in Grice (1989, pp. 133–53).
- Grice, H.P. (1989). *Studies in the Ways of Words*. Cambridge, MA: Harvard University Press.
- Hardin, C.L. (1988). *Color for Philosophers*. Indianapolis: Hackett Publishing Company.
- Hardin, C.L. (1997). Reinverting the Spectrum. In Byrne and Hilbert (1997a, pp. 289–301).
- Harman, G. (1990). The Intrinsic Quality of Experience. *Philosophical Perspectives* 4: 31-52.
- Harrison, B. (1973). *Form and Content*. Oxford: Basil Blackwell.
- Hilbert, David (1987). *Color and Color Perception*. Stanford: CSLI.
- Hume, David, (1740/1978). *Treatise of Human Nature*. Oxford: Clarendon Press.

- Jackson, F. (1982). Epiphenomenal Qualia. *Philosophical Studies* 32: 127–36.
- Johnston, M. (1991). Explanation, Response-Dependence and Judgement-Dependence. In P. Menzies (ed.) *Response-Dependent Concepts*. Working Papers in Philosophy, 1, 122-183. Canberra: Philosophy Program, Research School of Social Sciences, ANU.
- Johnston, M. (1992). How to Speak of the Colors. *Philosophical Studies* 68: 221–261. Reprinted in Byrne and Hilbert (1997a, pp. 137-172).
- Johnston, M. (1993). Objectivity Refigured. In J. Haldane and C. Wright, *Reality and Representation*. New York: Oxford University Press.
- Johnston, M. (1997). Impressionism, Primitivism and Relativity. Unpublished paper, Department of Philosophy, Princeton University, Princeton, NJ.
- Kripke, Saul (1972). *Naming and Necessity*. Cambridge, MA: Harvard University Press.
- Locke, John (1689/1975). *An Essay Concerning Human Understanding*. Oxford: Clarendon Press.
- McGinn, C. (1996). Another Look at Colors. *The Journal of Philosophy* 11: 537–553.
- Moore, G.E. (1903). The Refutation of Idealism. *Mind* 12: 433–53. Reprinted in Moore (1922).
- Moore, G.E. (1922). *Philosophical Studies*. London: Routledge & Kegan Paul.
- Nagel, T. (1974). What is it Like to be a Bat?. *Philosophical Review* 81: 127–36.
- Shoemaker, S. (1990). Qualities and Qualia: What's in the Mind?. *Philosophy and Phenomenological Research* 50: 109–31. Reprinted in Shoemaker (1996a).
- Shoemaker, S. (1991). Qualia and Consciousness. *Mind* 100 507–24. Reprinted in Shoemaker (1996a).
- Shoemaker, S. (1994). Phenomenal Character. *Noûs* 28 21–38. Reprinted in Byrne and Hilbert (1997a, pp. 227-245). A version of this also appears in Shoemaker (1996a).
- Shoemaker, S. (1996a). *The First Person Perspective and Other Essays*. Cambridge: Cambridge University Press.
- Shoemaker, S. (1996b). Intrasubjective/Intersubjective. In Shoemaker (1996a), pp. 141–154.
- Swartz, Robert J., ed. (1965). *Perceiving, Sensing and Knowing*. Berkeley: University of California Press.
- Tappendon, J. (1993). Analytic Truth—It's Worse (or Perhaps Better) than You Thought. *Philosophical Topics* 21: 233–261.
- Tye, M. (1992). Naturalism and the Mental. *Mind* 101: 421–41.
- Tye, M. (1994). Qualia, Content, and the Inverted Spectrum. *Noûs* 28: 159–83.

Tye, M. (1995). A Representational Theory of Pains and their Phenomenal Character. *Philosophical Perspectives* 9.

Wittgenstein, Ludwig (1945). *Philosophical Investigations*, translated by G.E.M. Anscombe. New York: MacMillan Publishing Co., Inc.

Notes

¹ For the sake of convenience, however, we will primarily speak of surface colors.

² “At least in part” because visual experience might nevertheless *misrepresent*, say, the shape of the thing.

³ Since we will be exclusively concerned with *visual* experience, we will occasionally drop the qualification.

⁴ Throughout we follow the terminology and presentation of Byrne and Hilbert (1997a), (1997b).

⁵ Moore is not the only historical figure to have observed the transparency of experience. Wittgenstein (1945, §275–7) has also taken note that our experience is “as it were diaphanous”. For more recent sources see also Evans (1982, pp. 227–30), Firth (1949–1950), Grice (1962, p. 259), and Harman (1990, pp. 34–40).

⁶ A number of recent authors have cited Moore in this context: Dretske (1995, p. 62); Shoemaker (1990, pp. 112–3), (1991, p. 516), (1994, pp. 30–1); Tye (1992, pp. 160, 167), (1994, p. 162), (1995, pp. 30–1, 135–6).

⁷ Hence the Humean metaphor: “the mind has a great propensity to spread itself on external objects” (1740, p. 167).

⁸ We should point out that NECESSITY, COLOR REP → COLOR FEEL, and COLOR FEEL → COLOR REP, are different, if related, theses than the theses that go under the same names in Byrne and Hilbert (1997a, 1997b). Byrne and Hilbert (1997a, 1997b) assume a Millian conception of color content—the content of a color experience is taken to be exhausted by the color property it designates, and NECESSITY, COLOR REP → COLOR FEEL, and COLOR FEEL → COLOR REP are formulated in terms of this conception of color content. In the present essay we have dropped this assumption and adopted a more general formulation of these theses.

⁹ Such an attribution is controversial, however. Locke (1689, Book II, Ch. VIII), for example, explicitly denies that our ideas of secondary qualities resemble the powers that produce them.

¹⁰ In denying that there are any qualia, the intentionalist does not deny that there are phenomenal properties—that there is a distinctive way it is like to undergo a conscious experience. Indeed intentionalism is an account of what phenomenal properties *are*. Though an irrealist about qualia, the intentionalist is not, thereby, committed to the existence of *zombies*—creatures whose experience lacks an associated phenomenology. Quining qualia need not involve quining phenomenal consciousness.

¹¹ This underscores an important difference between the inverted spectrum argument and so-called “absent qualia” arguments. Like the inverted spectrum argument, the absent qualia argument is really a family of arguments against NECESSITY. Consider a normal subject whose experience of *o* under fixed viewing conditions is both green-representing and green-feeling. The advocate of an absent qualia argument claims that we can conceive of a subject whose “experience” of *o* under identical viewing conditions is green-representing but which lacks phenomenal character. There thus could be two subjects whose experiences share the same content but differ in phenomenal character (trivially, since one of the subjects is *a zombie*—his “experience” lacks a distinctive phenomenology). Thus COLOR REP → COLOR FEEL and, hence, NECESSITY fails. The absent qualia argument, unlike the inverted spectrum argument, consists *entirely* in a conceivability claim. As such, we believe that absent qualia arguments are as problematic as the Cartesian argument for dualism. Just as Descartes argues that since we can conceive of our mind existing without our body, they must be distinct, the advocate of the absent qualia argument argues that since we can conceive of a state being green-representing without also being green-feeling, these properties must also be distinct. In each instance, however, the inference from conceivability to possibility is fallacious. Oedipus can conceive of Jocasta existing without his mother also existing, but it doesn’t follow that they are non-identical. Indeed, unbeknownst to Oedipus, Jocasta *is* his mother.

¹² The advocate of NECESSITY is committed to there being no difference in color phenomenology without a *relevant* difference in content. All of the arguments that we have considered for (2) assume that the relevant aspect of visual content is an experience’s *color* content, *i.e.*, its representation of the color property *o* instantiates. It is open to the intentionalist, however, to deny that color phenomenology is to be explained by the visual representation of color—color phenomenology might be explained by the representation of some *other* property. Call such a version of intentionalism *the extra property view*. Advocates of the extra property view include Shoemaker (1994) and Michael Thau in his forthcoming Ph.D. thesis. But if we are right

in contending that none of the arguments for (2) are cogent, then the extra property view is *unmotivated*. Part of the reason for the view consists in accepting some version of the inverted spectrum argument as sound.

¹³ We use hue, saturation, and brightness merely to illustrate the three dimensions of similarity found in a human color space. Nothing we say depends on hue, saturation and brightness actually being the best way of characterizing our color space.

¹⁴ An account of color similarity that uses color content to fix the similarity metric is developed in Byrne and Hilbert (1997b). That account differs from the present one in not making use of the relations among the colors as contributing to the content of color experience.

¹⁵ Red is more similar to orange than to blue because while red and orange are reddish, blue is not. Specifically, then, the constraint is this: whatever extensions are assigned to our perceptual color categories, the extensions assigned to red and orange should be disjoint proper subsets of the extension assigned to the determinable reddish and the extension assigned to blue should be a proper subset of the determinable's anti-extension. These constraints are analogous to what Jamie Tappendon (1993) has described as *pre-analyticity*.

¹⁶ Harrison (1973) and Hardin (1988, 1997) also argue against the possibility of spectrum inversion on the basis of such asymmetries of the color space, though their operative assumptions are different from our own. Hardin, for example, appeals to asymmetries involving cross-modal comparisons. In contrast, we confine our attention to the autonomous domain of color phenomenology.

¹⁷ There are two ways to defeat the inference from conceivability to possibility. One could argue that the putative state of affairs while conceivable, is not, in fact, possible. Or one could argue that the target state of affairs has not in fact been conceived, and hence the conceivability claim provides no reason for believing it to be possible. (For more information on the role of conceivability in modal epistemology see Yablo, 1993.) We shall pursue this latter strategy with respect to the alleged possibility.

¹⁸ Like the nominalist, the qualia-freak's conviction outstrips his reasons. And this suggests that the source of their conviction lies elsewhere—perhaps in a philosophical *picture*.

¹⁹ We endorse the conclusion of this passage with the exception of one minor infelicity. The issue is not whether subjects with supersymmetrical quality spaces could *tell* whether there is a phenomenal difference between distinct experiences, but whether there would be a phenomenal difference *at all*.

²⁰ One could not object that Nonvert_{SYM}'s and Invert_{SYM}'s experience must differ in phenomenal character because of some behavioral or cognitive difference between them. Nonvert_{SYM} and Invert_{SYM} could be functional duplicates (given the symmetry of their color space), and if they were the phenomenal character of their respective experiences could be responsible for no behavioral or cognitive difference between them.