A Relationalist's Guide to Error About Color Perception

Jonathan Cohen

Blue and yellow, bitter or sweet, can never be false ideas: these perceptions in the mind are just such as they are there, answering the powers appointed by God to produce them; and so are truly what they are, and are intended to be.
— Locke, *Essay Concerning Human Understanding*, II, 32, 16.

Abstract

Color relationalism is the view that colors are constituted in terms of relations to perceiving subjects. Among its explanatory virtues, relationalism provides a satisfying treatment of cases of perceptual variation. But it can seem that relationalists lack resources for saying that a representation of x's color is erroneous. Surely, though, a theory of color that makes errors of color perception impossible cannot be correct. In this paper I'll argue that, initial appearances notwithstanding, relationalism contains the resources to account for errors of color perception. I'll conclude that worries about making room for error are worries the relationalist can meet.

Color relationalism is the view that colors are constituted in terms of relations to perceiving subjects (possibly *inter alia*). The leading motivation for relationalism is that the view provides a ready solution to disputes involving perceptual variation: where S_1 and S_2 appear to disagree about x's color, the relationalist responds that both are right, insofar as x both bears the color *red for* S_1 *(in circumstance* C_1*)* and fails to bear the color *red for* S_2 *(in circumstance* C_2*)*. But the relationalist's ease in reconciling apparently conflicting representations of a single stimulus comes with a concomitant cost: while relationalism makes it extremely easy for a representation to be veridical (hence the possibility of reconciling apparently conflicting variants), it can seem that the relationalist lacks resources for saying that a representation of x's color is erroneous. Surely, though, a theory of color that makes errors of color perception impossible cannot be correct. Does this mean that we should give up color relationalism?

I want to argue that we should not. This paper is not principally an attempt to argue for relationalism about color (for that, see Cohen (2004)). Rather, it is an attempt to show how the relationalist can respond to an important anti-relationalist argument — namely, the argument that relationalism cannot

account for errors of color perception. I'll start by reviewing briefly the case for relationalism (§1), and then showing why relationalism threatens to make error impossible, and why that would be an unsatisfactory outcome (§2). Next, I'll show how, despite the worries, the relationalist has at her disposal the materials to account for several types of errors of color perception (§3). Finally, I'll consider objections (§4). I'll conclude that worries about making room for error are worries the relationalist can meet.

1 From Perceptual Variation to Color Relationalism

One of the most important motivations for color relationalism comes from the observation that there are wide interpersonal and intrapersonal (and interspecies) variations, along a number of different dimensions, in color perception.[1] To take just one widely discussed and widely observed example, a single stimulus in a fixed viewing condition can look unique green (i.e., it can look greenish without looking at all yellowish or at all bluish) to subject S_1 while failing to look unique green to subject S_2, even though S_1 and S_2 both have perfectly normally functioning visual systems (they are both non-anomalous trichromats, neither fails on standard discrimination tasks, etc.). Likewise, a single stimulus can look unique green to a single observer S under one viewing condition C_1 but fail to look unique green to S under another viewing condition C_2, even though both conditions fall well within the realm of normality. (For example, this can occur because C_1 and C_2 differ in illumination, background, viewing distance, angular subtense of visual field, or any of many other factors; cf. Cohen (2004) and Hardin (1988) for discussion.)

When such cases arise (as, in fact, they do all the time), there is a range of variants — viz., the way the stimulus looks to S_1 in C_1, the way the stimulus looks to S_2 in C_2, and so on. On the natural (and widely held) view that perception in general, and color perception in particular, represents the world, all of these variants represent the color of the (single) stimulus.[2] If you follow Hilbert (1987), Tye (1995), and Byrne and Hilbert (2003) in the anti-relationalist supposition that colors are subject- and circumstance-independent properties of surfaces, you are committed to saying that at most one of these representational variants veridically represents the color of the stimulus: for, on this view, at most one of the variants represents the stimulus as having the subject- and circumstance-independent surface property that it in fact does have. Unfortunately, it is hard to see that anything could make it the case that one of the variants is veridical at the expense of the other. For one thing, several hundred years of systematic efforts directed at this problem have failed to provide an answer (cf. Hardin (1988), 67–82); for another, it seems that any considerations that could be brought forward in support of the veridicality of one of the variants could be matched by considerations of equal force in favor of some other variant: what reason is there (other than *ad hoc* stipulation) for

thinking that the variant produced in S_2 in C_2, say, trumps the variant produced in S_1 in C_1? If there is no principled answer to this question, then this casts doubt on the anti-relationalist supposition.

What to do?

One reaction, favored by Byrne and Hilbert (2003), is to cling to the view that colors are subject- and circumstance-independent properties, and to insist that there is an unknown (or possibly unknowable) fact of the matter about which variant is veridical. While this position is coherent, it strikes me as pretty desperate.

An alternative, and less desperate, reaction is to give up the troublesome anti-relationalist insistence that colors are subject- and circumstance-independent properties. Instead, we should hold that colors are constituted in terms of relations to subjects and viewing conditions. In particular, on this view (henceforth, 'color relationalism'), a single stimulus can genuinely be unique green to S_1 in C_1, and genuinely fail to be unique green to S_2 in C_2. The virtue of color relationalism in the present setting is that, in the face of S_1's report that the stimulus is unique green in C_1 and S_2's report that the stimulus is not unique green in C_2, we can avoid having to choose between them — we can coherently say that both reports are correct. And that is a good thing because, as noted above, it's extremely hard to see that there is a metaphysically principled way of making this choice. Color relationalism, then, avoids the (apparently unsatisfiable) need to choose between the variants, and therefore doesn't require the unjustified optimism to which her opponent is driven. On the contrary, relationalism permits us to take seriously the wide range of empirically observed variation in color perception without giving up realism about color properties. This, it seems to me, is an attractive combination of virtues.

2 Whither Error?

Despite these virtues, relationalism faces an extremely serious worry: it is unclear that relationalism is compatible with the possibility of errors of color perception.[3] Before I take on this worry, it is worth saying (i) why it would be so bad for a theory of color to preclude errors of color perception, and (ii) why relationalism might be thought to commit that sin.

2.1 Why Be Wrong?

Why, then, is it so important that we preserve the possibility of errors of color perception?

A first motivation comes from the straightforward evidence of our practices of color attribution: a perceiver will typically offer a second perceptually informed color attribution in an effort to correct an earlier one (by herself or someone else) that she thinks is mistaken. Surely the most direct way of explaining these practices would sustain the second reporter's thought that the first perceptually informed color attribution, and also the perceptual state that informs that first

attribution, is erroneous. If this is right, then the simplest explanation of our practice commits us to the view that color perception is liable to error.

A second, more theoretical, motivation comes from the plausible thought that color perception represents (some aspects of) the world to the organisms that possess it. This thought will, I think, strike many as platitudinous for reasons that have nothing to do with the relationalist/anti-relationalist debate about color. It is a thought that comes from a much more general conception of perception — one that goes at least as far back as Aristotle — as a mechanism that represents various aspects of the extramental, extra-perceptual world for perceivers, in order that they can think about and interact with that world. (Moreover, the thought that color perception is representational is explicitly invoked in the argument from perceptual variation rehearsed in §1 — namely, it is implicit in the argument's construal of perceptual variants as the sorts of things that have veridicality conditions. As such, the relationalist in particular has even stronger than usual reasons for endorsing the thought at issue.) But many have held that the possibility of error is a necessary element of the notion of representation. Indeed, the possibility of error is taken to be so important to representation that it is *the* central bone of contention over the adequacy of what is perhaps *the* leading approach to representation in contemporary philosophy of mind (see Fodor (1990)).[4] So, then, we have good reasons for thinking of color perception as representing the world, and we have good reasons for thinking that it couldn't possibly represent the world unless errors of color perception were possible. This leads pretty quickly to the conclusion that errors of color perception had better be possible.

2.2 Error Lost?

We have, then, strong reasons for thinking that an account of color that precludes error is *ipso facto* unacceptable. Why think that relationalism fails that test?

As we have seen in §1, relationalism is motivated by the need to reconcile distinct perceptual variants when (i) each purports to represent the color of a single stimulus, and (ii) there seems to be no non-arbitrary way of singling out one of those variants as veridical at the expense of the others. Relationalism allows us to accept all of the variants — to allow that each represents the stimulus veridically. But this treatment of perceptual variation threatens to over-generalize: the proposal seems so liberal in accepting perceptual variants that it is unclear how it could allow that any variant is *not* veridical. On reflection, this is unsurprising; relationalism is designed to be inclusive about perceptual variants, so it is hardly shocking to find that it is ill-placed to exclude variants at the end of the day.

Prima facie, relationalism seems not to have the resources to treat any perceptual variants as erroneous. On the other hand, we have seen that there are powerful reasons for insisting that an acceptable account of color must make room for that possibility. Something has got to give.[5]

3 How Can Something Feel So Right and Be So Wrong?

In this section I shall argue that, despite the considerations canvassed above, the relationalist has the resources to accommodate errors of color perception. In fact, I will argue for something even stronger (in §3.3) — not only that relationalism allows for the metaphysical possibility of error, but that it leads us to classify as erroneous the sorts of cases that are pre-theoretically described as perceptual errors.

As a way of organizing the exposition it will be useful to taxonomize perceptual errors around the traditional distinction between hallucination and illusion (a distinction that goes back to Esquirol's 1838 *Maladies Mentales*).[6] On this usage, a subject hallucinates an object x when she perceptually represents x as bearing some feature or features while, in fact, there is no object x that the subject perceives. For example, I hallucinate when, in my drunken state, I perceptually represent pink elephants in the room even though there are in fact no pink elephants in the room. Hallucinations are to be contrasted with illusions; an illusion occurs when the subject perceptually represents an object x that she is indeed perceiving, but errs in the features she perceptually represents x as bearing (either by perceptually representing x as bearing features that x in fact lacks, or by perceptually representing x as lacking features that x in fact bears).[7] Thus, I suffer from an illusion when I perceptually represent the (actually existing) immersed oar as being bent when it is in fact straight.

With this rough and ready distinction in hand, I want to explain how the color relationalist can account for several forms of error in color perception — including two relatively abstruse types of error in §§3.1–3.2 and a much more central type in §3.3 — and can correctly classify pre-theoretically erroneous cases as erroneous.[8]

3.1 Hallucination

First consider the example of hallucination already mentioned. Suppose I perceptually represent a pink elephant in the room even though there are in fact no elephants in the room. (Or, if I am a drunk and hallucinating relationalist, suppose I perceptually represent that an elephant is in the room and is pink to me in my perceptual circumstance.)[9]

I take it there is a clear sense in which my perceptual state is erroneous, and that this sense is available to color relationalists and non-relationalists alike. In particular, here the perceptual state involves the attribution of a property to an object (viz., an elephant) that, as it were, lacks the perfection of existence. Here, there is something that is (i) independent of the representational state itself, (ii) a standard of correctness for the representational state, and (iii) what makes it non-arbitrarily the case that the state counts as erroneous — namely, the paucity of local elephants. So it seems that the relationalist can indeed count this case as a perceptual error.

The relationalist can say the same thing about afterimages. Suppose that, because of what happens to my visual system at t_1, I experience a red afterimage when gazing at a white wall at t_2 (for relationalists: gazing at a wall that is white for me in the circumstances I'm in at t_2). The case would be an illusion rather than a hallucination if my visual system represented at t_2 some region of the wall as being red (/as being red for me in the circumstances I'm in). But it seems more true to the phenomenology to say that my visual system represents at t_2 local redness (/to me in the circumstances I'm in) as being exemplified by something other than the wall or any of its regions — it seems to float entirely free of the wall.[10] Once again, my perceptual state attributes the property in question to a merely intentional (i.e., non-existent) object. It is, then, a perceptual hallucination.

On reflection, it is unsurprising that all this is available to the relationalist about color, insofar as what goes wrong in hallucinatory states — the perceptual representation of the inexistent — is more far reaching than any issues about color in particular. Since this is so, one can entirely bracket the question of whether there are errors that have specifically to do with representations of *color*, and still end up with the conclusion that hallucinatory states of the sort under discussion are erroneous (that's why our explanation of its being erroneous didn't have to say anything specifically about color, as manifested for example by its neutrality between relationalist and non-relationalist formulations of the perceptual property ascriptions). In one sense, then, a relationalist who points to the possibility hallucination has not thereby allowed for errors of color perception *per se*. On the other hand, the color relationalist might reasonably respond, hallucinatory representations involving the attribution of color properties are, after all, a perfectly good class of erroneous perceptual representations of color; therefore, if relationalism makes room for these states, it allows for what we worried it doesn't allow for in §2, and thereby answers our worry. But, on the first hand again, it would be unsatisfying if the only possible errors of color perception were hallucinatory states, so the relationalist needs to say more.

3.2 Illusion Through Deviant Causation

I want to argue that the color relationalist can make room for illusory, as well as hallucinatory, errors of color perception. Recall that my representation of x's color is illusory only if (i) there is an x such that I perceive x, and (ii) I perceptually represent x as bearing a color it lacks, or as lacking a color it bears.[11]

A first sort of illusion that the relationalist can recognize involves deviant causal chains. For example, consider the (admittedly remote) case of the telekenetically chromatic tomato. This tomato, T, like other ripe tomatoes, has the capacity to look red to me under the conditions I am now in. However, T has, in addition, a very unusual telekinetic capacity: it has the capacity to affect nearby visual cortices directly — i.e., without retinal stimulation of any kind; indeed, the way in which T affects those cortices is to induce in them the state

they normally undergo when they experience a very unripe tomato of the same size and shape as T. When I visually attend to T, its telekinetic capacity instantly swamps the effect of its ordinary, non-telekinetic capacity, so I end up perceptually representing that T is green to me in my perceptual circumstances.

What should the relationalist might say about my representation of T's color?

One option is to endorse it. For, just as the relationalist endorses the perceptual representations of object colors produced in low illumination (for example) as veridical representations of the colors objects bear in those circumstances, she might hold that T genuinely is green to me in the (admittedly remote) circumstance in which T can telekenetically affect my visual cortex.

A more plausible option, to my mind, is to say that it is erroneous. The thought here would be that, when I visually represent objects as being green to me in my perceptual circumstance (or as bearing any color), this representation carries a commitment to the non-deviancy of the causal process that produces it. And, while it beyond the scope of this paper to say what distinguishes deviant from non-deviant causation, it strikes me that any adequate account of that distinction should classify the telekinetic process that connects T to my visual cortex (not via my retina) as deviant. Why think that visual representations of color carry a commitment to the non-deviancy of their mode of production? Because it's hard to see why these representations count as *visual* if they can be brought about by non-visual (e.g., telekinetic) causal pathways.[12] So, then, the idea is that my representation of T's color, *qua* visual, carries a commitment to the non-deviancy of its etiology (or, at least, a commitment to its etiology's involving the usual mechanisms of the visual system), but that that commitment is in fact not satisfied, so the representation in question is erroneous. In particular, since, in the case described, T exists and is perceived by me, the current treatment of the case would classify it as a perceptual illusion.[13]

If this is right, then cases involving deviant causal chains provide another form of erroneous perceptual representation involving the attribution of color that lies within the grasp of the color relationalist. However, once again, this seems a pretty unusual (i.e., cooked up by overactive philosophical imagination) class of errors; we should hope, therefore, that the relationalist's account of error extends to more ordinary cases. Luckily, it does.

3.3 Ordinary Illusion

To explain the relationalist's treatment of a much more important and ubiquitous class of errors of color perception, I need first to say something about the relationalist's account of ordinary color ascriptions.

According to relationalism, as expressed in §1, the ripe tomato is not red *simpliciter*, but red for a subject in a circumstance; that is, relationalism takes colors to be relational properties rather than monadic properties. On the other hand, linguistically expressed color ascriptions in English and other natural languages ordinarily treat colors as monadic predicates. *Prima facie*, it seems that the linguistic evidence is at odds with the theory. But, as I have argued

elsewhere Cohen (2004), there are reasons for thinking that ordinary color ascriptions (those that occur in our ordinary thought and talk about the colors of objects) treat colors not as *non*-relational, but as *tacitly* relational.

In particular, I suggest that color ascriptions are tacitly relativized to (vague) parameters fixed by our (vague) pragmatically presupposed interests in making those ascriptions — viz., we are interested in delimiting the range of perceivers and viewing conditions to those that matter to us when we make the ascriptions. Thus, we say/think that x is red *simpliciter* just in case x is red for perceivers pretty much like ourselves, in circumstances pretty much like those we encounter.[14]

What happens when our pragmatic interests shift — say, because we become interested in the perceptual systems of other creatures or of our fellows in other perceptual circumstances? Just what you would expect if there are the tacit default presuppositions in place that I claim there are: we use explicit qualifiers to cancel presuppositions that would otherwise remain in place. Thus, we say/think that the ripe tomato is red for us, but not for a bee, or red in bright sunlight, but grey when illuminated by candlelight.

As far as I can see, the relationalist needs to tell something like this story about color ascription — otherwise she has no way of reconciling the apparently unrelativized form of color predicates with her commitment to the relationality of color properties. Moreover, I think this story has quite a lot going for it: for example, it recognizes the centrality of visual systems like our own in our thought about color without giving in to extreme anthropocentrism that would deny the possibility of deference to other sorts of visual systems (cf. Matthen (1999)), and suggests a natural understanding of the various scientific and industrial specifications of standard observers and viewing conditions as precisifications designed for particular purposes (which is just what they appear to be). In addition, this story provides what we need to account for ordinary illusions of color perception.

Here's how. Let it be that Sally the subject is invited to the psychophysics lab and is asked to view the stimulus — a ripe tomato as it happens — under viewing condition C. The stimulus, let us suppose, is red for Sally in condition C. Now, Sally will report that the tomato is red *simpliciter* just in case she takes it to be red for perceivers pretty much like herself, in circumstances pretty much like those she normally encounters. Of course, she thinks she herself is a perceiver quite a lot like herself, and she takes her present perceptual circumstance C to be pretty much like those she normally encounters, so she thinks the tacitly presupposed conditions for the ascription of *red simpliciter* are met. Hence, she represents the tomato as being red *simpliciter*, and reports as much to the experimenter. It turns out, however, that Sally has been fooled: C was constructed by the clever psychophysicist so that (i) C would lie outside the range of perceptual circumstances pretty much like those she encounters, (ii) the tomato's appearance in C to Sally would be entirely distinct from the very same tomato's appearance in perceptual circumstances pretty much like those she encounters, and (iii) there would be no visual clues to tip off Sally to these facts about C. (Psychophysicists have many techniques for implementing such

subtle manipulations; they include the use of contrast effects, unusual illumination, and so on; cf. Hardin (1988) for a description of some of the parameters of the perceptual circumstances that affect color appearance. This is old hat to psychophysics, and it is no surprise that Sally can be fooled.)

In this case, I claim, Sally represents the color of the tomato erroneously. As it happens, the tomato *is* red for Sally in C; so if she had represented only that it is red for her in C, she would have avoided error. But she did not so confine herself. Rather, because the experimental manipulation was subtle enough not to tip her off, she represented it less cautiously as being red *simpliciter* — which it was not. This is just to say that Sally's representation of the tomato's color is erroneous; and since (we are supposing) there really is a tomato that Sally perceives, the error is a textbook case of perceptual illusion.

A similar recipe for illusion would manipulate Sally herself, rather than the perceptual circumstance under which she perceives the tomato.[15] Suppose that the experimenter alters Sally's visual system (e.g., by causing chromatic adaptation in her retina, or, for that matter, by slipping something in her coffee) and then asks for her report about the tomato's color in the (unaltered) perceptual condition C'. Assume that the tomato is red for Sally in C'. If Sally is unaware of the manipulation, she is likely to assume that she is a perceiver pretty much like herself; but the experimental manipulation has made it the case that, in the relevant sense, Sally is not any longer pretty much like herself. Because of her false assumption, then, Sally will represent not merely that the tomato is red for her in C' (which is true) but that the tomato is red *simpliciter* — i.e., red for perceivers pretty much like herself, in circumstances pretty much like those she normally encounters (which is false). Here, too, we have a case of illusory representation of the tomato's color.

What I am suggesting, then, is that we represent the exemplification of colors on two simultaneous layers: we represent that x is red for S in C, and, often, we additionally represent that x is red *simpliciter*. The illusions now under discussion occur when the subject makes the (subdoxastic) transition from the first representation, which is veridical, to the second representation, which is erroneous. It turns out, then, that relationalists have a standard against which perceptual representations can be judged after all — namely, the standard supplied by the pragmatic presuppositions that are incorporated in the ascription of red *simpliciter*, green *simpliciter*, etc. at the second representational level. The existence of this standard is enough to secure for the relationalist an account of illusions of color perception; but since the standard is pragmatically governed and flexible, and since it doesn't choose between representations at the first level, it leaves the relationalist free to embrace the wide range of observed perceptual variation in a way that the non-relationalist cannot (see §1).

It is worth observing that the sort of illusion discussed above is not confined to the psychophysics lab. For example, it arises in cases that are naïvely described as color illusions, such as those in popular books of illusions (e.g., Seckel (2002)). One such case is the famous Benham disk, depicted in figure 1.[16] When the Benham disk is rotated about its center at a rate of about 6–8 Hz, rings of desaturated chromatic colors appear on its surface. I think there

Figure 1: The Benham disk (from Hardin (1988), 73).

is a reasonable pre-theoretical case for counting our perceptual representation of desaturated chromatic bands on the surface of the spinning Benham disk as illusory: this is, for example why the effect is surprising to learn about. Moreover, the case seems amenable to the (relationalist-friendly) account of illusion now under discussion. The story would be that, when you look at the spinning disk, some region of it *is* blue for you under the highly constrained perceptual circumstance you're in. As it happens, you represent the relevant region of the disk's surface as being blue *simpliciter* — viz., blue for perceivers pretty much like yourself in circumstances pretty much like those you normally encounter. But this is false: the disk is *not* blue for perceivers pretty much like yourself in circumstances pretty much like those you encounter, but only blue for such perceivers in the highly constrained circumstance in which the disk is rotating at rate in the 6–8 Hz range. Thus, your representation of the disk's color is correctly classified by the account I'm defending as an illusion of color perception.

An even more naturalistic example of this sort of illusion caused me great frustration one night when I walked up and down the parking lot, searching for (but repeatedly walking past) the Chevrolet I had rented. The Chevy had been red *simpliciter* when I parked it in daylight illumination, so I was looking for something red to me in the nighttime perceptual circumstance; unbeknownst to me, the nighttime circumstance involved illumination by sodium vapor lights, and consequently the rented Chevy looked (not red, but) grey to me in that perceptual circumstance. Because I was unaware of relevant facts about the perceptual circumstance, I erroneously represented the rented Chevy as being grey *simpliciter*, and wrongly judged that it could not be the car I had parked earlier in the day (after all, that car had been red *simpliciter*). It is natural to describe the error I made in the parking lot as an error of color perception; it

is to the credit of relationalism that it explains that natural and pre-theoretical description of this and other commonplace cases.

3.4 Vagueness and Ordinary Illusion

In §3.3 I claimed that x is red *simpliciter* just in case x is red for perceivers more or less like ourselves, in circumstances more or less like those we encounter.

Now, it is a vague matter whether something is a perceiver more or less like ourselves. For there are perceivers determinately like ourselves (human beings with normal trichromatic visual systems), perceivers determinately unlike ourselves (dichromatic squirrels (Thompson (1995), 145), decachromatic mantis shrimp Cronin and Marshall (1989)), and perceivers that are neither determinately like ourselves nor determinately unlike ourselves (anomalous trichromat human beings). Similarly, it is a vague matter whether a circumstance is more or less like those we encounter. For there are circumstances determinately like those we typically encounter (from 90 degrees at a distance of one meter under flat illumination of 7000 degrees Kelvin with uniform surround), circumstances determinately unlike those we typically encounter (at a depth of 3 miles underwater), and circumstances that are neither determinately like those we typically encounter nor determinately unlike those we typically encounter (lit by a candle, or with a surround field consisting of thin diagonal lines).

It follows that whether something is red *simpliciter* (or green *simpliciter*, etc.) is itself vague. Indeed, these ascriptions should be vague along multiple dimensions, insofar as there are multiple vague dimensions of comparison between perceivers (e.g., number of cone types, cone spectral sensitivity curves, cone adaptation rates, total visual angle, acuity, etc.) and between perceptual circumstances (illumination, surround contrast, visual angle, eye-to-stimulus distance, hallucinogenic drugs ingested, etc.). And indeed, this prediction of the account seems correct: if the perceiver continues to classify a stimulus as red *simpliciter* while we manipulate the perceiver and perceptual circumstances in less and then more severe ways, the classification strikes us as first determinately correct, and eventually determinately incorrect, with a penumbral region in between where the classification seems neither determinately correct nor determinately incorrect.

Now, on the account presented, ordinary color illusions occur when subjects represent stimuli as bearing the color red *simpliciter* (as it might be) that those stimuli in fact lack those properties. Consequently, if it is vague whether x lacks red *simpliciter*, then it is vague whether S's representation of x as bearing that property is illusory. That, too, strikes me as a plausible consequence of the theory. There are some cases that seem to fall determinately on the illusory side of the illusory/veridical distinction (I offered the Benham disk and my error in the parking lot in §3.3), some that seem determinately to fall on the veridical side (when you represent a ripe tomato as red *simpliciter*), and some that fall determinately on neither side.

I suggest that the predictions of the account of ordinary illusion presented here are correct, and that this is a reason for believing the account.

4 Objections

4.1 Overintellectualization

On the view I have presented, the most important class of errors of color perception (viz., the illusions discussed in §3.3) arises only in the context of the pragmatically governed presuppositions undergirding our thought and talk about colors — our anthropocentric presuppositions about the sorts of perceivers and perceptual conditions we care most about. But, by making error beholden to the way we think and talk about color, this treatment can seem to overintellectualize the phenomenon. One way to see this is to ask whether prelinguistic infants and non-human animals can suffer from illusions of color perception. To say they cannot would commit to the curious consequence that creatures become more susceptible to perceptual illusion as their cognitive sophistication increases. But to say they can (assuming we wish to maintain a uniform account of these illusions) would require them to make presuppositions about certain kind of perceivers and perceptual conditions. That does seem to impose surprising intellectual demands on infraverbals; no?

I'm prepared to bite the bullet here. Infraverbals don't do a lot of talking about color, so the presuppositions in question would have to be presuppositions of their thought about color. But why deny that infraverbals have thoughts about color, or that these thoughts carry such presuppositions? Needless to say, the claim here is not that infraverbals (or adult human subjects, for that matter) are aware of these sorts of presuppositions — something that would be implausible. Rather, the position requires only that infraverbal representation of colors of objects is tacitly committed to such presuppositions. But that sort of commitment is hardly extravagant or unusual; indeed, developmental and comparative psychologists routinely uncover such tacit commitments to extremely sophisticated presuppositions in infraverbals (e.g., infraverbals seem to be committed to the principle that objects move in continuous spacetime trajectories, to the principle that causal influence requires contact, and so on Spelke (1990)).

To the extent that the account of illusion proposed here turns on locating occurrent perceptual representations against a backdrop of presuppositions, it seems fair to characterize the view as an intellectualized account of color illusion. Whether it amounts to an *over*intellectualized account depends a lot on where we draw the bounds of the intellectual. The suggestion I am making is that, because the sorts of intellectual demands we are considering are modest and well precedented, it is appropriate to construe 'intellectual' broadly; if so, then the present proposal is reasonably regarded as an intellectualized account of color illusion, rather than one that is overintellectualized.

4.2 It's the Visual System, Stupid

I am proposing that (the most important class of) color illusions occur only in the context of our thought and talk about color, and only against a background

of presuppositions organizing this thought and talk. Plausibly, such presuppositions are contributions of the cognitive system, or at least the cognitive-*cum*-perceptual system as a whole rather than the visual system *per se*.[17] And if that is right, then such illusions are not aptly described as errors of the visual system *per se*, but only as errors of the cognizing/perceiving system as a whole.[18] Consequently, given the assumption that representation requires the possibility of error (see §2), it seems that the present account entails that the visual system *per se* doesn't represent the colors of objects.

There are a number of things that should be said in response.

First, it is worth recalling that the relationalist makes room for errors of color perception that don't depend on cognitive presuppositions — namely, the hallucination cases of §3.1 and the illusions of §3.2. These forms of error are naturally described as attaching to the visual system *per se* insofar as they don't depend on cognitive presuppositions of any kind. Therefore, this is enough to ensure the possibility of error, and therefore the possibility of representation, for the visual system

Now, one might object that, this response is unsatisfactory because it appeals to a quite restricted set of cases, and therefore leaves a wide range of cases untouched. In particular, if we restrict attention to those cases where the visual system does not suffer from either hallucination or illusion by deviant causation, we can explain errors within this restricted set of cases only by appeal to cognitive factors; but if so, it might seem as if it is impossible that the visual system *per se* errs about cases in this restricted range, so representation is impossible in this restricted range by the visual system *per se*. But why accept that representation requires the possibility of error in this restricted range? Surely it would be too strict to require the possibility of error in arbitrary subsets of cases. For, on anybody's story, the visual system is incapable of error in the restricted class of cases where it operates veridically; but surely this shouldn't be taken to show that the visual system is incapable of representation in the restricted class, nor to show that the visual system is incapable of representation at all. I take this to show that, while it is reasonably required that a representational system make room for error in general, it is not a reasonable requirement to insist that it make room for error in some particular range of cases. In particular, it is enough for the representational *bona fides* of the visual system (*per se*) that it make room for error in some cases rather than in all cases, and I have shown how the relationalist can secure this desideratum.

However, even granting that the relationalist treatment of error allows for representation by the visual system, it can still seem counterintuitive to claim, as I have, that in ordinary cases of color illusion the locus of error is the whole cognitive/perceptual system, rather than the visual system in particular. Of course, the relationalist can recognize a derivative sense in which the visual system is erroneous — viz., that the state of the visual system is liable to lead to a false conclusion given the existing background of presuppositions. (Compare: when a witness produces a true but misleading piece of testimony, we are inclined not only to say that conclusions reached on the basis of the testimony are false, but also that there is something derivatively wrong about the testimony itself,

despite its truth when considered on its own.) But, an opponent will suggest, this sort of derivatively erroneous status is not enough: what is needed is that there are illusions involving the representations of color in the visual system *per se* — i.e., in the visual system considered on its own, rather than considered as part of a larger cognitive/perceptual system.

But I think this objection depends on treating our intuitions about error with much more evidential authority than they deserve. It is hard to deny the authority of the intuition that we make errors in color perception; for reasons explained above, I think we would be justified in rejecting any account of color that could not be reconciled with this intuition. On the other hand, it is hard to see why we should trust intuitions about how the labor of producing these errors is divided between the visual system and other components of the cognitive/perceptual system. Surely that's something to be sorted out by (broadly) empirical inquiry, not by the armchair consultation of intuitions.

By way of analogy, consider what the linguist says about acceptability judgments. It is reasonable to insist, on the basis of considering your own reactions to the cases, that *the bulldogs fight* is acceptable and that *the bulldogs the bulldogs the bulldogs fight fight fight* is unacceptable; moreover, it is incumbent on linguistic theory (in the broad sense) to explain those verdicts. But it is not reasonable to insist, on the basis of considering your own reaction to the case, that the unacceptability of the latter string is due to its failure to conform to the grammar of the language in particular. On the contrary, the standard story goes, acceptability judgments are the result of the interaction of the grammaticality faculty with other components in the cognitive system (in this case, attention and memory thresholds are likely to be important to the story), and it is up to systematic empirical inquiry, as opposed to armchair consultation of intuitions, to dole out the labor of explaining the phenomena.

Likewise, I don't see that our intuitions about illusions of color perception come marked as intuitions that have to be explained by the operation of the visual system alone. Rather, there is an unmarked intuition to the effect that such illusions arise, and all we can insist based on that intuition is that an acceptable account of color perception allow for such illusions. If relationalism meets this demand, as I have suggested that it does, we have no further complaint against it on this score.

5 Conclusion

The idea that there are errors of color perception is so fundamental to our (naïve and scientific) thinking about the visual system that it would be very difficult to accept a theory of color that failed to sustain it. In particular, even though color relationalism has many virtues, I would be willing to give it up if I thought it could not accommodate errors of color perception. As it happens, however, it seems to me that relationalism contains the resources to account for errors of color perception, and correctly classifies as erroneous those cases that are naturally and pre-theoretically described as erroneous. It seems, after all,

that relationalism can indeed get things wrong; in so doing, it might get things right.[19]

Notes

[1]I present and defend the argument for color relationalism sketched here in much greater detail in Cohen (2004).

[2] The view that color perception is representational in this sense is extremely widespread (but see Smith (2002) and Travis (2004) for dissent), and I'll be assuming it in what follows. As far as I know, it is not disputed by any of the authors whose views I'm criticizing in the present section.

[3]

Versions of the accusation that relationalism (typically some particular form of relationalism) precludes error occur in e.g., (Hilbert (1987), 88), (Watkins (2002), 93), (Matthen (2001), note 10) and (Byrne and Hilbert (2003), 57–58), among other places.

[4]The possibility of error should only be regarded as a necessary element of the notion of representation for semantically atomic (uncomposed) states. Presumably, after all, we wouldn't want to say that the belief that 2+2=4 is non-representational; likewise for so-called cogito-like judgments (which can be contingent; see Burge (1988), Burge (1996)). I take the assumption that the visual system states representing colors are atomic to be reasonably plausible (it is, after all, hard to see what their constituents would be). Consequently, it is plausible that the possibility of error should indeed constrain theories of how colors are represented by the visual system. (Thanks to an anonymous referee for helping me to see the need for this qualification.)

[5]Objection: Perhaps what should give is the premise that color perception is representational (cf. note 2). This premise is not only at the root of the difficulties raised for anti-relationalists (§1), but, we are now seeing, seem to threaten relationalism as well.

Response: The relationalist who gives up the premise in question does not thereby escape the worries about providing for errors of color perception. On the contrary, the difficulty generalizes, insofar as giving up the premise makes it much more difficult to say what it means for color perception (or perception more generally speaking) to be veridical or not on any occasion. Consequently, I join the consensus in thinking that denying the premise is not the answer to our troubles.

[6]

Despite its historical precedent, the following formulation of the illusion/hallucination distinction is arguably somewhat revisionary. E.g., it will follow from this formulation that afterimages and amodal completion are hallucinations rather than (as they are sometimes called) illusions. My sense is that 'illusion' often causally used as a general term for both of what I'll calling illusions and hallucinations (and even cases that on my view, are neither illusions nor hallucinations: e.g., the Rubik's cube "illusion" (287), and Jastrow's ambiguous duck-rabbit figure included in (Seckel (2002) 287, 145)). I'm willing to stipulate about the terminology in this way since (i) I don't see that the revisionism carries serious costs, and (ii) it will help me distinguish between cases that I want to treat differently.

[7]Here we should distinguish between S's perceptually representing x as lacking feature F, on the one hand, and S's failing to represent x as bearing F, on the other. Presumably subjects fail to represent all sorts of features that the objects they perceive bear — e.g., they might represent that x is triangular without representing that x is isosceles, or might represent that x is bigger than a breadbox without representing that x is 3.27 cubic meters in volume. But representational omissions of this sort shouldn't count as illusions. What I have in mind, in contrast, are errors of representational commission — e.g., what happens when S represents that the oar lacks the property *straightness* that the oar in fact bears (and not simply what happens when S fails to make representational commitments about the oar's shape).

[8]McLaughlin (2003) proposes a further type of relationalist error (over and above the types I shall consider), that is tied closely to his specific version of relationalism (and so not available to other relationalists). On McLaughlin's form of relationalism, "Redness for a visual perceiver of type P in circumstances of visual observation C is that property which disposes its bearers to look red to P in C, and which [is] had by everything so disposed" (122). According to McLaughlin, the final clause of the quoted proposal, which distinguishes this view from other forms of relationalism, provides an extra road to errors of color perception: a state representing that x is red for P in C can be erroneous even if x looks red for P in C if it turns out that x fails to share the property required of red things.

I don't see how this proposal can work. In order that McLaughlin's second clause can have any teeth, he must be construing 'property' in a way that would count different bases for the disposition to look red to P in C as different properties (otherwise it will be trivially true that bearers of the disposition satisfy the second clause — in which case nothing could satisfy the first clause without satisfying the second clause, so there could be no additional path to errors of color perception here after all). But it seems clear as a matter of empirical fact that the class of things that have the disposition to look red to P in C are heterogeneous at every level of description lower than 'all of them have the disposition in question' (cf. Nassau (1980)). Or, in other words, it seems

clear that the disposition is realized by a range of distinct basis properties. Consequently, given McLaughlin's proposal, it looks like the empirical facts force the conclusion that nothing is red. But if McLaughlin's view amounts to a form of color eliminativism, rather than a form of relationalism, it's not true that he has provided for the relationalist a form of error other than those I shall discuss.

[9]There are, of course, hallucinations caused by other means as well. In particular, many of the states Hardin (following reasonably common usage) calls "subjective colors" — e.g., those caused by eyeball pressure, migraines, cosmic ray bombardment, electrodes (Hardin (1988), 91–96) — are naturally described as hallucinatory color experiences, and can be treated by the account in this section. On the other hand, some of the examples listed above strike me as non-erroneous (e.g., those involving opponent interactions, 91) or as illusions (e.g., the Benham disk, about which see §3.3).

[10]Larry Hardin has pointed out to me that, while this phenomenological report is correct in the vast majority of cases, it seems incorrect as a report about the afterimage produced by Bidwell's disk (see Hurvich (1981), 191 for a description), which seems to be confined to a quite specific location — viz., the location of the light behind the disk.

[11]As stated, this criterion is too strict, since it would preclude cases of so-called veridical illusion (cf. Lewis (1980)). I'll explain how such cases might be counted as illusory nonetheless in note 13.

[12]I'm not assuming the deviant/non-deviant distinction is principled: I can imagine many possible causal pathways that diverge from the standard one less dramatically than does the telekinetic pathway by which T acts, and I doubt that there is a crisp line past which a pathway counts as deviant. In any case, all I need for present purposes is that there might be some visual representations produced by deviant causal pathways; the reader who disagrees that the case under discussion qualifies is invited to substitute one that does.

[13] Indeed, it is natural to extend this treatment to cases of veridical illusion. Consider a second telekinetic tomato T^*; T^* is unripe, and, like other unripe tomatoes, has the capacity to look green to me under the conditions I am now in. However, like T, T^* has an additional, capacity to induce directly in nearby visual cortices the state they normally undergo when they experience a very unripe tomato of the same size and shape as T^* in the normal way (i. e., by a causal pathway involving the retina). Moreover, as with T, T^*'s telekinetic influence on a cortex immediately swamps T^*'s visual influence on that cortex.

The same reasons that make it plausible to think that my representation of T's color is erroneous extend to my representation of T^*'s color, even though in the latter case T^*'s color happens to be exactly what I represent it to be (viz., green to me in the circumstance I am in).

[14]There is reason for thinking that these presuppositions work differently for those who know that their visual systems are statistically anomalous; indeed, there is some evidence that they make tacit presuppositions about the visual systems of people without the anomaly; cf. Cohen (2004), note 40. I'll ignore this complication in what follows.

[15]There's a natural sense in which manipulations to the subject can be redescribed as manipulations to the perceptual condition, and vice versa. Consequently, its unclear whether what follows is a second recipe for illusion, or just a redescription of the first.

[16]I've chosen the Benham disk because the chromatic effects it produces are often counted as illusory, and because I can easily reproduce it here without incurring the cost/difficulty of chromatic printing. But there are many other cases that are amenable to the account in the text. Some examples in Seckel (2002) are the Hermann grid illusion (title page), neon color spreading (245), the Bezold effect (299), etc. On the other hand, there are several cases in this book that, on the account propounded here, are not illusions of color perception (e.g., the Rubik's cube "illusion" (287; cf. the demonstrations at `http://www.purveslab.net/`)), and even several that seem (independently of the views about color I'm defending) not to be illusions at all (e.g., Jastrow's ambiguous duck-rabbit figure, 145).

[17]It might be suggested that this conclusion is too quick. After all, the ascription of presuppositions to the visual system has been fruitful in other areas of vision science. For example, as Ullman (1979) proved, the visual system has enough information to recover structure from motion only if the object moves rigidly; but the visual system uniformly applies its structure from motion algorithm whether the condition is met or not, obtaining a correct representation of object structure only when it is met. It is natural here to say that the visual system is presupposing that objects move rigidly; why not, on the strength of this precedent, avoid the current objection by attributing the presuppositions at issue to the visual system *per se*?

Unfortunately, I doubt that this treatment can be extended to the presuppositions about perceivers and visual circumstances at issue in the present account of errors of color perception. A significant disanalogy is that Ullman's sort of presupposition, precisely because it occurs at the level of the visual system, is incapable of being overridden at the discretion of the perceiver; it is, in the terminology of Pylyshyn (1984), cognitively impenetrable. Whereas, in contrast, it is essential to the story I'm telling that the presuppositions at issue can be canceled when our pragmatic interests shift (see §3.3); this is just to say that these presuppositions are cognitively penetrable. For that very reason, it seems reasonable to think of the presuppositions I have in mind as cognitive contributions.

[18]Cf. Descartes' account of (intellectual) error in Fourth Meditation, accord-

ing to which the deliverances of the (properly confined) intellect are free from error, and it is only the (insufficiently confined) application of the will to the deliverances of the intellect that lead us into error.

[19]I am indebted to Craig Callender, Larry Hardin, Mohan Matthen, Brian McLaughlin, and two anonymous referees for discussion and comments on this paper, and to my co-participants in the Workshop on Colour Ontology and Colour Science at the University of British Columbia in 2003, who helped me to think through some of the initial ideas.

References

Burge, T. (1988) "Individualism and Self-Knowledge," *The Journal of Philosophy* 85:11, 649–663.

Burge, T. (1996) "Our Entitlement to Self-Knowledge," *Proceedings of the Aristotelian Society* XCVI, 91–116.

Byrne, A., and D. R. Hilbert (1997) *Readings on Color, Volume 2: The Science of Color*, MIT Press, Cambridge, Massachusetts.

Byrne, A., and D. R. Hilbert (2003) "Color Realism and Color Science," *Behavioral and Brain Sciences* 26:1, 3–64.

Cohen, J. (2004) "Color Properties and Color Ascriptions: A Relationalist Manifesto," *The Philosophical Review* 113:4, 451–506.

Cronin, T. W., and N. J. Marshall (1989) "A Retina With At Least Ten Spectral Types of Photoreceptors in a Mantis Shrimp," *Nature* 339, 137–140.

Fodor, J. A. (1990) "A Theory of Content, I: The Problem," in J. Fodor, ed., *A Theory of Content and Other Essays*, MIT Press, Cambridge, Massachusetts, 51–87.

Hardin, C. L. (1988) *Color for Philosophers: Unweaving the Rainbow*, Hackett, Indianapolis.

Hilbert, D. R. (1987) *Color and Color Perception: A Study in Anthropocentric Realism*, CSLI, Stanford.

Hurvich, L. M. (1981) *Color Vision*, Sinauer Associates, Sunderland, Massachusetts.

Lewis, D. (1980) "Veridical Hallucination and Prosthetic Vision," *Australasian Journal of Philosophy* 58, 239–249. Reprinted in Lewis (1986), 273–286.

Lewis, D. (1986) *Philosophical Papers, Volume II*, Oxford University Press, New York.

Matthen, M. (1999) "The Disunity of Color," *The Philosophical Review* 108:1, 47–84.

Matthen, M. (2001) "Our Knowledge of Colour," in J. Macintosh, ed., *Naturalism, Evolution, and Intentionality*, Canadian Journal of Philosophy, 215–246. Supplementary volume 27.

McLaughlin, B. (2003) "Color, Consciousness, and Color Consiousness," in Q. Smith and A. Jokic, eds., *Consciousness: New Philosophical Perspectives*, Oxford University Press, New York, 97–154.

Nassau, K. (1980) "The Causes of Color," *Scientific American* 243, 124–154. Reprinted in Byrne and Hilbert (1997), 3–29.

Pylyshyn, Z. W. (1984) *Computation and Cognition: Toward a Foundation for Cognitive Science*, MIT Press, Cambridge, Massachusetts.

Seckel, A. (2002) *The Great Book of Optical Illusions*, Firefly, Toronto.

Smith, A. D. (2002) *The Problem of Perception*, Harvard University Press, Cambridge, Massachusetts.

Spelke, E. (1990) "Principles of Object Perception," *Cognitive Science* 14, 29–56.

Thompson, E. (1995) *Colour Vision: A Study in Cognitive Science and the Philosophy of Perception*, Routledge, New York.

Travis, C. (2004) "The Silence of the Senses," *Mind* 113:449, 57–94.

Tye, M. (1995) *Ten Problems of Consciousness: A Representational Theory of the Phenomenal Mind*, MIT Press, Cambridge, Massachusetts.

Ullman, S. (1979) *The Interpretation of Visual Motion*, MIT Press, Cambridge, Massachusetts.

Watkins, M. (2002) *Rediscovering Colors: A Study in Pollyanna Realism*, Dordrecht, Boston.