



**WILEY-  
BLACKWELL**



---

Intentionalism and the Imaginability of the Inverted Spectrum

Author(s): Eric Marcus

Reviewed work(s):

Source: *The Philosophical Quarterly*, Vol. 56, No. 224 (Jul., 2006), pp. 321-339

Published by: [Blackwell Publishing](#) for [The Philosophical Quarterly](#)

Stable URL: <http://www.jstor.org/stable/3840945>

Accessed: 30/05/2012 13:44

---

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at <http://www.jstor.org/page/info/about/policies/terms.jsp>

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.



*Blackwell Publishing* and *The Philosophical Quarterly* are collaborating with JSTOR to digitize, preserve and extend access to *The Philosophical Quarterly*.

<http://www.jstor.org>

## INTENTIONALISM AND THE IMAGINABILITY OF THE INVERTED SPECTRUM

BY ERIC MARCUS

*Various thought-experiments have been offered as independent support for the possibility of intentionalism-defeating spectrum inversion, but they do not succeed. I refute what I take to be the four best arguments for holding that the thought-experiments do in fact provide such support: the implausible error argument, the symmetry argument, the no-inference argument and the best theory of representation argument. I thus offer a defence of intentionalism against a long-standing objection.*

Much has been written in recent years about whether a pair of subjects could have visual experiences which represented the colours of objects in the environment in precisely the same way, despite differing significantly in what it was like to undergo them, differing that is, in their qualitative character.<sup>1</sup> The possibility of spectrum inversion has been so much debated in large part because of the threat which it would pose to the more general doctrine of intentionalism, according to which the representational content of an experience fixes what it is like to undergo it.<sup>2</sup> In what follows, I shall argue that the thought-experiments offered as independent support for the possibility of intentionalism-defeating spectrum inversion fail in their aim. After an initial discussion of these thought-experiments (§I), I shall refute what I take to be the four best arguments for holding that the thought-experiments do in fact provide such support: the implausible error argument (§II), the symmetry argument (§III), the no-inference argument (§IV), and the best theory of representation argument (§V). I shall then resolve a potential difficulty for the intentionalist raised by considering the no-inference

<sup>1</sup> Recent discussions of the inverted spectrum include S. Shoemaker, 'Content, Character, and Color', *Philosophical Issues*, 13 (2003), pp. 253–78, and 'Colour and the Inverted Spectrum', in S. Davis (ed.), *Colour Perception: Philosophical, Artistic and Computational Perspectives* (Oxford UP, 2000), pp. 187–214; M. Tye, *Consciousness, Color, and Content* (MIT Press, 2000); M. Thau, *Consciousness and Cognition* (Oxford UP, 2002).

<sup>2</sup> Recent discussions of intentionalism (or 'representationalism') include N. Block, 'Mental Paint', in M. Hahn and B. Ramberg (eds), *Reflections and Replies: Essays on the Philosophy of Tyler Burge* (MIT Press, 2003), pp. 165–200; A. Byrne, 'Intentionalism Defended', *Philosophical Review*, 110 (2001), pp. 199–240; Tye, *Consciousness, Color, and Content*.

argument and the best theory of representation argument in combination (§VI). Though my argument does not, of course, constitute a proof of intentionalism, it does undermine what has been one significant source of doubt about the truth of that theory.

## I. IMAGINING THE INVERTED SPECTRUM

Intentionalism is typically understood as a claim of asymmetric dependence:

*Intentionalism.* Necessarily, if experiences  $e$  and  $f$  differ in their phenomenal character, they also differ in their representational content.

In discussions of colour, intentionalism is sometimes understood as also involving the converse:

*Converse intentionalism.* Necessarily, experiences  $e$  and  $f$  differ in their phenomenal character if they also differ in their representational content.

The relevant thought-experiments which threaten these principles are scenarios in which a pair of subjects (or single subjects at different times) undergo perceptual experiences  $e$  and  $f$  which satisfy one or other of the following conditions:

- (a)  $e$  and  $f$  contain the same representational colour-content, yet are inverted with respect to their phenomenal character
- (b)  $e$  and  $f$  are identical with respect to their phenomenal character, yet have inverted colour contents.

(a) would conflict with intentionalism, (b) with converse intentionalism. There are a variety of ways in which one might come to think that (a) and (b) are possible. One might have reasons for rejecting intentionalism or converse intentionalism that have nothing to do with the inverted spectrum. One might invoke other science-fictional cases – the possibility of zombies, for example<sup>3</sup> – or one might appeal to mundane cases, such as those raised by Peacocke in *Sense and Content*.<sup>4</sup> One might also argue for the possibility of (a) or (b) not on the basis of imagining scenarios in which these conditions hold, but rather by arguing that their possibility follows from other plausible assumptions about, for example, the physiology of colour perception.<sup>5</sup> I shall

<sup>3</sup> See, for example, D. Chalmers, *The Conscious Mind* (Oxford UP, 1996). I argue against the conceivability of zombies in my 'Why Zombies are Inconceivable', *Australasian Journal of Philosophy*, 82 (2004), pp. 477–90.

<sup>4</sup> C. Peacocke, *Sense and Content* (Oxford UP, 1984). In a different vein, see N. Block, 'Sexism, Racism, Ageism, and the Nature of Consciousness', in R. Moran, J. Whiting and A. Sidelle (eds), *The Philosophy of Sydney Shoemaker, Philosophical Topics*, 26 (1999), pp. 71–88.

<sup>5</sup> See, for example, Shoemaker, 'Content, Character, and Color'.

not be concerned with these other kinds of arguments here. I am concerned only with thought-experiments in which one is supposed to imagine the phenomenal character and representational content of an experience as coming apart in the manner of (a) or (b). My claim will be that these thought-experiments show very little.

The imaginative core of any such thought-experiment involves imagining one of the following scenarios:

1. A perceiver *S*, in an environment identical with mine, has an experience phenomenologically indistinguishable from an experience I would have in an environment in which all of the colours had been swapped with their complements
2. A perceiver *S*, in an environment identical with mine save that all of the colours have been swapped with their complements, has an experience phenomenologically indistinguishable from the experience I have in my normal environment.

It is the imaginability of (1) and (2) which can lure beginners or non-philosophers into thinking philosophically about the nature of colour and colour-experience. And it is the great plausibility of the notion that (1) and (2) can be imagined which is chiefly responsible for the sense amongst philosophers that the imaginability of (a) and (b) is just a datum from which further philosophizing on the topic must begin.<sup>6</sup> But imagining (1) and (2) is, on its face, a far cry from imagining (a) and (b). Without further details, the natural way to think about imagining (1) would be that one had imagined a subject whose perceptual experience is either simply erroneous, or in any case in plain disagreement with mine about the colours of objects in our respective environments. Similarly, without further details, the natural way to think about imagining (2) would be that one had imagined a scenario in which one, the other, or perhaps both of us are visually misled. Thus the imaginability of (1) and (2) do not by themselves threaten intentionalism.

What must be added to (1) and (2) to produce the threat? Judging from the classic tellings of inverted spectrum and inverted earth stories, we should add that the situations of (1) and (2) are the *normal ones* for the perceivers, that is, crimson objects normally prompt subjects to have experiences phenomenologically indistinguishable from ones that I would have when confronting chartreuse objects, and so forth. I shall not attempt to define the notion of 'normal' here. It might amount simply to a regular and/or natural correlation between the relevant colours and experiences, or there might

<sup>6</sup> See N. Block, 'Inverted Earth', in N. Block, O. Flanagan and G. Güzeldere (eds), *The Nature of Consciousness* (MIT Press, 1997), pp. 677–93, at p. 679, for a particularly clear statement of this sentiment.

need to be a suitable evolutionary story connecting them, or the perceivers might need to be ensconced in a community of perceivers similarly disposed, or some combination of the above. Be that as it may, according to proponents of the imaginability of intentionalism-defeating inverted spectrum scenarios, a subject whose normal response to seeing, for example, a chartreuse object is to have a perceptual experience phenomenologically indistinguishable from those we have when seeing crimson objects is a subject whose experience represents chartreuse objects *as* chartreuse, the qualitative inversion (relative to us) of this experience notwithstanding. Thus if we add to (1) and (2) what I shall call the 'normality condition', giving (1<sub>n</sub>) and (2<sub>n</sub>), then according to proponents of the imaginability of the intentionalism-defeating inverted spectrum, we have good reasons for thinking we have in mind the experiences described by (a) and (b).

The question I begin with, then, is this: what reason is there for thinking that in imagining (1<sub>n</sub>) and (2<sub>n</sub>), one has *thereby* imagined a scenario in which conditions (a) and (b) hold? Why not instead think of (1<sub>n</sub>) and (2<sub>n</sub>) as situations in which a subject's perceptual experiences misrepresent the colours of objects in the environment? In the sections that follow, I shall consider the following four answers to this question, here listed, and later discussed, in order of increasing plausibility:

*Implausible error argument.* To say of a subject *S* in (1<sub>n</sub>) or (2<sub>n</sub>) that *S*'s perceptual experience misrepresents the colours of objects would be to attribute to *S* widespread perceptual error. But widespread perceptual error would result in a diminished capacity to navigate in the environment. Since there is no such diminished capacity in (1<sub>n</sub>) and (2<sub>n</sub>), there is no perceptual error. Hence (a) and (b) hold in (1<sub>n</sub>) and (2<sub>n</sub>).

*Symmetry argument.* The same considerations as lead us to accept that our twin earth *Doppelgänger* has XYZ- rather than H<sub>2</sub>O-thoughts should lead us to accept that *S* in (2<sub>n</sub>) has crimson-representing experiences of crimson objects, the chartreuse feel of these experiences notwithstanding. Hence (b) holds in (2<sub>n</sub>).

*No-inference argument.* Since the subjects of (1<sub>n</sub>) and (2<sub>n</sub>) can non-inferentially form beliefs about their environment on the basis of their perceptions, and these beliefs are generally true, the perceptions on which they are based must also be true, regardless of their feel. Hence (a) and (b) hold in (1<sub>n</sub>) and (2<sub>n</sub>).

*Best theory of representation argument.* According to our best theories of representation, a perceptual content that is normally caused by chartreuse is thereby chartreuse-representing, regardless of its feel. Hence (a) and (b) hold in (1<sub>n</sub>) and (2<sub>n</sub>).

I shall argue that each of these arguments fails. Thus I shall argue that inverted-spectrum/inverted-earth thought-experiments pose no threat to intentionalism. The strategy I shall employ can be put into relief by contrasting it with others. One way for an intentionalist to respond to those thought-experiments is by accepting that the imagining of  $(1_n)$  and  $(2_n)$  amounts to the imagining of (a) and (b), accepting that the former pair is imaginable, but then denying that their imaginability is metaphysically significant – denying, that is, that their imaginability entails their possibility. A second way for an intentionalist to respond is by accepting that the imagining of  $(1_n)$  and  $(2_n)$  amounts to the imagining of (a) and (b), but then denying that the former pair are genuinely imaginable. My strategy is rather to accept that  $(1_n)$  and  $(2_n)$  are imaginable, but to deny that in imagining them one is thereby imagining (a) and (b). This strategy has clear advantages over the previous two. All other things being equal, a defence of intentionalism which does not force us to reject either the core idea behind inverted-spectrum/inverted-earth cases or the idea that these possess metaphysical significance is preferable to one which does.

I shall not rely here on various *empirical* objections which have been raised against the possibility of spectrum inversion, from the asymmetry of our colour-quality space, for example.<sup>7</sup> According to these objections, there would be behavioural and functional differences between you and your colour-inverted twin. If representational content were determined functionally, then this would render the inverted spectrum inconceivable. I do not take these empirical objections into account here, in part because I do not believe one needs to assume functionalism about representational content in order to defeat arguments which would take us from  $(1_n)$  and  $(2_n)$  to (a) and (b). More importantly, however, I believe it will be instructive to consider non-empirical philosophical reasons for rejecting the conceivability of (a) and (b).

I also take this strategy to be preferable, *ceteris paribus*, to another that has been deployed recently. According to this intentionalist strategy, it is not the representation of *colour* that is so deeply wedded to the phenomenal character of visual experience, but rather the representation of some other property, albeit one that is closely associated with colour; and intentionalism is true because the representation of this other kind of property supervenes on phenomenal character.<sup>8</sup> I cannot here address fully the considerations which make this last strategy appear attractive. I do hope to make it clear, however, that inverted spectrum thought-experiments lend no support to it,

<sup>7</sup> See, for example, Shoemaker, 'Color and the Inverted Spectrum'.

<sup>8</sup> See Shoemaker, 'Introspection and Phenomenal Character', *Philosophical Issues*, 28 (2000), pp. 247–73, and *Consciousness and Cognition*.

as against the natural view that it is the representation of colour in particular which depends on the phenomenal character of visual experience. For those thought-experiments, I hope to show, do not tell against this latter view at all.

## II. THE IMPLAUSIBLE ERROR ARGUMENT

In the previous section I said that adding the normality condition to (1) and (2) has been thought to lend support to the idea that we have in mind the experiences described by (a) and (b). But how exactly does the normality condition help to confer this support? The first answer goes like this:

*Implausible error argument.* To say of subject  $S$  in  $(1_n)$  or  $(2_n)$  that  $S$ 's perceptual experience misrepresents the colours of objects would be to attribute to  $S$  widespread perceptual error. But widespread perceptual error would result in a diminished capacity to navigate in the environment. Since there is no such diminished capacity in  $(1_n)$  and  $(2_n)$ , there is no perceptual error. Hence (a) and (b) hold in  $(1_n)$  and  $(2_n)$ .

Those of a pragmatist, if not behaviourist, bent will see a subject whose perception enables coping perfectly well with the environment as one whose perception is by and large veridical. If  $S$  were really always mistaken about the colours of objects in the environment, according to this way of thinking, then this would be bound to show up in  $S$ 's behaviour. Coping perfectly well with the environment is at odds with always misperceiving the colours of objects.

But this argument is weak. Whereas mundane cases of systematic misperception of colour, like colour-blindness, can render perceivers unable to cope with their environments, total colour inversion cannot. A colour-inverted subject  $S$  has a colour quality-space of richness equal to ours. So any pair of objects that we can distinguish,  $S$  can distinguish too. And  $S$ 's colour experience is also as stable as ours. Objects look the same to  $S$  unless they have been painted, or the lighting has been changed, or  $S$  has put on sunglasses, etc. Objects appear to  $S$  to be differently coloured in just those circumstances in which they appear to us like that as well. With colour experience of the same richness and stability,  $S$  will track and reidentify objects just as well as those who get the colours of objects right.  $S$  is mistaken, but it is a function of the ingenuity of the thought-experiment that the world never makes  $S$  pay for this defect. There is no problem here.<sup>9</sup>

<sup>9</sup> This argument is criticized more extensively in Shoemaker, 'Color and the Inverted Spectrum'.

It may be that *S*'s spectrum inversion is behaviourally undetectable (though it may not be, as Shoemaker argues). If it is, our normal standards of evidence for the attribution of colour content will justify our attributing to *S* the same sorts of (veridical) perceptions in the same sorts of situations as we do to ordinary perceivers.<sup>10</sup> But however compelling the behavioural evidence is for the claim that *S* is intentionally identical with normal perceivers, it is by no means obviously sufficient to outweigh the stipulation that *S*'s experience is qualitatively inverted relative to normal perceivers – not unless we assume some form of behaviourism or verificationism about intentional content (but not qualitative character).

The implausible error argument poses the following would-be rhetorical question: what more do you want from perceptual experience beyond its enabling you to deal with the world successfully, and to behave as veridical perceivers do? The intentionalist answer is simple: we want it to reveal the true colours of objects.

### III. THE SYMMETRY ARGUMENT

*Symmetry argument.* The same considerations as lead us to accept that our twin earth *Doppelgänger* has XYZ- rather than H<sub>2</sub>O-thoughts should lead us to accept that *S* in ( $2_n$ ) has crimson-representing experiences of crimson objects, the chartreuse feel of these experiences notwithstanding. Hence (b) holds in ( $2_n$ ).<sup>11</sup>

The symmetry argument is an attempt to export a plausible line of thinking concerning our perception of natural kinds to our perception of colour. Suppose Oscar is suddenly and without his knowledge moved to twin earth, where the watery stuff is XYZ. For some period of time immediately following the move, Oscar's claims to the effect that local pools of liquid are water will be false, as will the thoughts these claims express. Eventually, after a certain amount of time has passed, the theory goes, it is no longer plausible to say that he is speaking and thinking of H<sub>2</sub>O. His claims regarding what he calls 'water', together with the thoughts that these claims express, now target XYZ. And it does not seem too much of a stretch from this to say that the content of Oscar's perceptual experience has also changed: it used to represent pools of water in his immediate environment; now it represents pools of twin water in his immediate environment. If one accepts

<sup>10</sup> See Shoemaker, 'The Inverted Spectrum', *Journal of Philosophy*, 79 (1982), pp. 357–81.

<sup>11</sup> This argument is advanced explicitly in Shoemaker, 'Qualia and Consciousness', repr. in his *The First-Person Perspective and Other Essays* (Cambridge UP, 1996), pp. 121–40, at p. 134, and implicitly in 'Inverted Earth', p. 683.



this view of Oscar and natural kinds, then according to the anti-intentionalist, one should accept the analogous view of Jean, suddenly and without her knowledge moved to inverted earth, where the colours of all objects are inverted relative to ordinary earth colours.<sup>12</sup> She is fitted with colour-inverting lenses, and so notices no difference. According to the anti-intentionalist, by parity of reasoning, we should also say that the content of Jean's colour experience eventually changes. But we have no reason to think that the qualitative character of her experiences changes – at least no reason which does not just assume intentionalism. Thus the symmetry argument takes us from (2<sub>n</sub>) to (b).

This argument is problematic for a number of reasons. As stated, it simply presupposes that perceptual representational content depends on features of the subject's *recent* environment. But some have argued that content depends on features of the environment in which our perceptual system was designed by nature to function.<sup>13</sup> Since Jean's trip to inverted earth will not affect her evolutionary history, it cannot, according to this sort of view, put her in a context in which her crimson-feeling experience comes to represent chartreuse. Since I do not want to rest my defence of intentionalism on any particular theory of perceptual representation, I shall employ a different strategy in this section. My focus here will be what I take to be the problematic appeal to symmetry itself.

The fundamental point at issue concerns the extent to which we should treat natural kinds and colours similarly when thinking about the relationship between the qualitative and representational aspects of experience. A crucial *asymmetry* between natural kinds and colours is suggested by a salient difference between the twin earth and inverted earth stories. They are similar, of course, in that both involve victims placed in a new world whose differences from our world are to them undetectable. The differences must be undetectable, because it is only if they are undetectable that the case will reveal whether the phenomenal aspects of experience can remain constant while its representational aspects shift. But whereas no alterations need to be made to Oscar's perceptual apparatus to foster his illusion, Jean must be fitted with colour-inverting contact lenses to foster hers. Why must her perceptual apparatus be interfered with in this way? Because she will not mistake chartreuse for crimson unless either viewing conditions have been rendered non-standard, or she has been rendered a non-standard observer. If neither of these has been done, the same colour will look the same, and hence Jean will be in a position to catch the shift.

<sup>12</sup> See Block, 'Inverted Earth'.

<sup>13</sup> See F. Dretske, *Naturalizing the Mind* (MIT Press, 1995), and Tye, *Consciousness, Color, and Content*, for different versions of this view.

What, if anything, does this difference between twin and inverted earth cases tell us about the difference between the relationship of phenomenology and representation in the cases of natural kinds and colours? For a start, it tells us that whereas veridical perceptions of water and XYZ are phenomenologically indistinguishable in standard conditions by standard observers, veridical perceptions of crimson and chartreuse are not. That veridical perceptions of water and XYZ are phenomenologically indistinguishable in standard conditions by standard observers suggests that the following principle is false:

*Converse natural-kind intentionalism.* Necessarily, if experiences  $m$  and  $n$  differ in their (visual) representation of natural kinds, they differ in their phenomenal character.

That veridical perceptions of crimson and chartreuse are *not* phenomenologically indistinguishable in standard conditions by standard observers at least leaves it open that the following principle is true:

*Converse colour intentionalism.* Necessarily, if experiences  $m$  and  $n$  differ in their (visual) representation of colour, then they differ in their phenomenal character.

The fact that standard veridical perceptions of crimson and chartreuse are not phenomenologically indistinguishable does not, of course, prove converse colour intentionalism. But it does suggest a closer tie between representation and phenomenology in the case of colour than in that of natural kinds. And this closer tie renders very dubious the argument that what holds for Oscar and natural kinds must therefore also hold for Jean and colours. At the very least, it opens up a hole in the symmetry argument. Given the closer tie between representation and phenomenology in the case of colour perception, 'parity of reasoning' arguments cannot be accepted without additional support. The symmetry argument, as stated baldly above, fails.

#### IV. THE NO-INFERENCE ARGUMENT

*No-inference argument.* Since the subjects of  $(1_n)$  and  $(2_n)$  can non-inferentially form beliefs about their environment on the basis of their perceptions, and these beliefs are generally true, the perceptions on which they are based must also be true, regardless of their feel. Hence (a) and (b) hold in  $(1_n)$  and  $(2_n)$ .

In Shoemaker's original presentation in 'The Inverted Spectrum' of the inverted spectrum argument, Fred undergoes a series of colour inversions

that eventually put him in the position described in (1). Afterwards, Fred is a perceiver in a normal environment who has experiences which are phenomenologically indistinguishable from experiences we would have in an environment in which all of the colours had been swapped with their complements. For some period after the change, although Fred knows that the sky is blue, that pumpkins are orange, etc., he is still tempted to think that the sky is orange, that pumpkins are blue, etc. He might express this by saying that the sky 'looks orange'. Eventually, however, he gets used to the change and says sincerely that the sky 'looks blue' in just the same circumstances as ordinary perceivers say this. In later work, Shoemaker characterizes this 'semantic accommodation' as involving Fred's applying colour words 'spontaneously, and without inference ... in a way that accords with the usage of others and [his] own previous usage' ('Qualia and Consciousness', p. 123).

Immediately after the inversion, then, when Fred gets the colours of objects right, this is because he *infers* from the fact that a table looks yellow, for example, that it is violet. So long as his correct violet-judgements are still the product of such an inference, Shoemaker's thought seems to be, then the perceptual experiences that prompt those judgements continue to represent objects as yellow. After a while, however, Fred's experience of a violet object automatically prompts in him the belief that it is violet, and does so without any conscious inference – without any inference at all, we may suppose. According to Shoemaker, Fred's perceptual experience represents the relevant object as violet. (This is still not quite right. Fred must also, according to Shoemaker, forget that the qualitative character associated with perceptions of specific colours has changed. Only if this further condition is met does Fred's yellow-feeling state occupy the same functional role as his former violet-feeling state. But I shall ignore this further requirement for the purposes of this discussion.)

Shoemaker's understanding of the significance of this shift is supported in part by his functionalism about representational content. According to his functionalism, part of what makes a perceptual experience of violet a perceptual experience of violet is that it causes the belief that the object is violet, and does so without the mediation of any inference. But one need not be a functionalist to find a point in this vicinity compelling. Certainly, so long as Fred is using his own experience as evidence from which he infers the colours of objects, how he comes to know the colours of objects is out of the ordinary, to say the least. Immediately after the inversion, Fred, it seems right to say, does not really *see* that the table is violet; rather, he infers that the table is violet on the basis of what he does see. One way of putting the point which emerges from Shoemaker's discussion is this: if Fred's visual

experience really perceptually represents the table as violet, then he can come to believe that the table is violet non-inferentially. So even if, immediately after the inversion, Fred's perceptual experience *evidentially* represents the table as red, it does not *perceptually* represent the table as red. I shall return to this distinction later in this section.

But while it is plausibly a necessary condition of a perception's having a certain content that it is a possible basis for a non-inferentially formed belief with the same content, it is not so plausibly a sufficient condition. Certainly, Fred's non-inferential coming to believe that the table is violet on the basis of his visual experience appears to provide very good evidence that this experience represents the table as violet. After all, to resist this conclusion, we must hold both of the following: first, Fred's visual experience systematically misrepresents colour; secondly, the beliefs which Fred non-inferentially forms on the basis of those mistaken perceptions are true. And this is surely peculiar. But the appearance of peculiarity is, I shall argue, connected to a failure to think through the idiosyncratic nature of the non-inferential connection between Fred's experience and his belief.

Immediately following the inversion, the beliefs Fred inferentially forms on the basis of mistaken perceptions are true. Gradually, he comes to associate yellow-feeling visual experiences with violet to such an extent that the transition between the feeling and the belief becomes effortless, automatic – non-inferential. But this way of non-inferentially coming to have a belief on the basis of perception is unusual. In a normal case, we do not come to believe that objects have the colours they have in virtue of associating an experience of a certain phenomenologically individuated type with the relevant colour. Rather, the colour-feel of an experience partly constitutes my grasp of what it is that my perceptual experience represents the relevant objects as having. The colour-feel of an experience just is what the object has, when the experience is veridical.

I submit that the sense of puzzlement over how Fred might non-inferentially arrive at true beliefs about colours on the basis of misperceptions is rooted in our thinking about his belief-forming process on the model of a normal case. In a normal case, one non-inferentially forms a belief about the colour of an object on the basis of a visual perception by tacitly endorsing it. To say that one tacitly endorses a visual perception is to say in part that one treats the phenomenal character of the experience as partly constituted by the colour of the seen object (together perhaps with other features of the perceived situation, e.g., lighting conditions). We can put this by saying that one typically *simply trusts* one's experience. It would be hardly coherent to suppose that a belief arrived at by simply trusting an experience in this sense could differ in content from the experience itself. But this is not

Fred's situation, even long after the inversion. Fred has simply come to associate effortlessly an object's seeming to be yellow with its actually being violet. But there is no incoherence in supposing that one can effortlessly associate one colour with experiences that perceptually represent another.

Earlier, I distinguished between evidential and perceptual representations of colour. So long as Fred is using his experience as evidence from which to infer the colour of objects, his experience at best evidentially represents violet objects as violet; it does not perceptually represent them as violet. On this point, I am thus in agreement with Shoemaker. But Fred's case suggests a further requirement which perceptual representation of colour must meet beyond the absence of inference. In normal cases, one's ability to know the colours of objects on the basis of experience does not derive from using one's experience as a mediating indicator at all – neither as an indicator from which one infers the colour of an object, nor as an indicator with which one automatically associates the colour. To represent evidentially an object as being a certain colour is to be fit to serve as a mediating indicator in one of these two senses. While I shall not offer a full-blown theory of perceptual representation here, it seems plausible to say that this involves a method of acquiring knowledge of the environment non-inferentially, and more broadly, not on the basis of mediating indicators.

It might be argued that I have mischaracterized the nature of the difference between Fred before the operation and Fred after he ceases basing his perceptual beliefs on inference. In each case, one might think, Fred perceptually represents violet objects as violet. What has changed is simply *how* he perceptually represents violet. Prior to the operation, he perceptually represents the objects as violet via a violet-feeling experience; long after the operation, he perceptually represents them via a yellow-feeling experience. On this view, phenomenal characters are *modes of presentation*.<sup>14</sup> I have said that Fred, long after the operation, uses his experience as a mediating indicator. But, to borrow a line from Gareth Evans, a way of perceiving colour is no more obliged to get in the way of perceiving colour, or to render perception of colour indirect, than is a way of dancing liable to get in the way of dancing, or to render dancing somehow indirect. If this rival characterization of Fred's development is correct, his experience long after the operation is not functioning as a mediating indicator, and there is thus no bar to viewing him as correctly perceiving the colours of objects.

There may indeed be something plausible about the idea that the phenomenal character of a visual experience is part of how colour is perceived,

<sup>14</sup> See D. Chalmers, 'The Representational Character of Experience', in B. Leiter (ed.), *The Future for Philosophy* (Oxford UP, 2004), pp. 153–81, for a defence of this sort of view.

and further that it is part of how colour is perceived without itself 'getting in the way' of our perception of colour. But the plausibility of this idea diminishes considerably when conjoined with the possibility of intentionalism-defeating inverted spectrum scenarios. If such scenarios are possible, then the phenomenal character of a visual experience *as such* makes no contribution to its representational content. Rather, a visual experience's being of yellow or of violet is a consequence of the fact that it happens normally to be caused by yellow or violet. Phenomenal character, on this view, is thus a matter only of *how* colours are represented, not at all a matter of *what* colours are represented. Given that there is nothing about a yellow-feeling experience as such that helps to put me in touch with the yellow it portrays the object as having, however, it is hard to see the phenomenal character of such an experience as constituting a direct mode of presenting yellow. If, on the contrary, phenomenal character as such helped to determine representational content, then we could better see it as directly presenting colour. But to see it in this way rules out the possibility of an intentionalism-defeating inverted spectrum.

Furthermore, as Shoemaker himself has pointed out, the view that phenomenal character concerns not the 'what' but only the 'how' of colour perception seems to get the phenomenology wrong. The phenomenal character of a visual experience seems to be inextricable from the properties it represents the world as having. Because of this, once we allow the possibility of an intentionalism-defeating inverted spectrum, it cannot help but seem that what visual experiences represent immediately is not colour but rather some other kind of property.<sup>15</sup> It is, then, by seeing this other kind of property that we see colours. Colours themselves remain wholly external to perception, and as such are at best only indirectly present to the perceiver.

The no-inference argument fails. It fails because genuine perceptual representation of colour requires not just that subjects form beliefs about the colours of objects non-inferentially on the basis of their experience, but that they do so, more generally, without using their experience as a mediating indicator.

## V. THE BEST THEORY OF REPRESENTATION ARGUMENT

*Best theory of representation argument.* According to our best theories of representation, a perceptual content that is normally caused by chartreuse is

<sup>15</sup> This is Shoemaker's view: see 'Introspection and Phenomenal Character'.

thereby chartreuse-representing, regardless of its feel. Hence (a) and (b) hold in  $(I_n)$  and  $(2_n)$ .<sup>16</sup>

Even if this argument were a good one, it would still not show that our leading accounts of representation, of which the causal–correlational and the teleological are perhaps the most prominent examples, support the anti-intentionalist reading of inverted-spectrum/inverted-earth scenarios. Michael Tye, who holds a version of the causal–correlational account of representation, has used it to argue *against* the possibility of an intentionalism-defeating inverted spectrum. No matter how long Jean remains on inverted earth, he holds (*Consciousness, Color, and Content*, pp. 136–40), her chartreuse-feeling experience will continue to represent chartreuse, in virtue, roughly, of the fact that her chartreuse-feeling experience will arise in response to chartreuse in the environment when conditions are optimal or normal. (Teleological accounts of representation could also be recruited to defend intentionalism along these lines.)

It is still true that according to such theories of representation, a perceptual content which is *normally* caused by chartreuse is thereby chartreuse-representing. But intentionalist advocates of such a theory will simply deny that Jean's situation is ever normal, so long as she remains on inverted earth, and also deny that Fred's situation, post-operation, is ever normal. In that sense, this strategy amounts to an explanation of why Jean's case and Fred's case are not properly described by  $(2_n)$  and  $(1_n)$  respectively. As Tye would understand 'normally', there cannot be a case in which crimson objects normally prompt a subject to have experiences phenomenologically indistinguishable from ones that I would have when confronted with chartreuse objects. Thus Tye must deny the conceivability of  $(1_n)$  and  $(2_n)$ .

Given that I do not rest my argument on denying the conceivability of  $(1_n)$  and  $(2_n)$ , and given that I also do not assume any specific theory of representation, I shall limit my attention in this section to those causal–correlational accounts of representation according to which content is based on features of a subject's recent environment. These accounts, in so far as they allow for the conceivability of  $(1_n)$  and  $(2_n)$ , pose the most serious threats to intentionalism. An anti-intentionalist can indeed use such an account to argue that Jean's chartreuse-feeling experience will eventually come to represent crimson. But I shall argue that even these accounts do not secure the possibility of (a) and (b). For the intentionalist can accept  $(1_n)$  and  $(2_n)$ , and allow only that Jean's chartreuse-feeling experience represents crimson *evidentially*. It does not represent crimson *perceptually*: that is, crimson is not a constituent of its perceptual content. So even if Tye's theory of

<sup>16</sup> See Block, 'Inverted Earth', p. 683, for an explicit statement of this argument.

representation is wrong, and  $(1_n)$  and  $(2_n)$  are genuinely conceivable, (a) and (b) do not follow.

That Jean's chartreuse-feeling experience can serve as evidence from which one might correctly infer the presence of crimson should be obvious from my discussion in the previous section. Suppose Frank is a newly blind denizen of inverted earth, familiar with the details of Jean's history. His brain is hooked up to a contraption that produces in him experiences that have the same visual phenomenology as Jean's. He can thereby use Jean's chartreuse-feeling perceptual experiences as a sure guide to the occurrence of crimson in her environment. If her experience is chartreuse-feeling, he can infer on this basis that there is a crimson object in the area. And he can use the relevant theory of representation to justify his claim to know that there is a crimson object in the area. His knowledge is underwritten by the fact that crimson objects are connected to chartreuse-feeling experiences in Jean in the manner specified by the relevant causal–correlational account of representation, underwritten by the fact, that is, that her chartreuse-feeling experiences do (evidentially) represent crimson. And if Jean is informed of the details of her own history, she also can treat her experience in the way Frank does, as supplying evidence concerning the colours of objects. She also can thereby come to know the colours of objects on the basis of perception.

My point here is not to suggest that Jean's chartreuse-feeling experience only represents crimson because she uses it as evidence for crimson. This gets things backwards. She can only use it (successfully) as evidence for crimson because it already represents it. To represent crimson evidentially is to be fit to serve as the evidence for a claim to know that crimson is present. But the fact that her experience evidentially represents the correct colours does not entail that her experience perceptually represents the correct colours. Indeed, there are the same reasons to deny the latter claim in Jean's case as there were to deny it in Fred's. Even if Jean eventually forms true beliefs about the colours of objects non-inferentially on the basis of her perception, it still does not follow that her perception represents the colours of objects correctly.

The case for Jean's correctly perceiving colours is even worse than for Fred's correctly perceiving colours, post-operation. In Ned Block's original telling of the inverted earth scenario, Jean is not informed of her changed environment. Nevertheless, it is alleged that unwitting Jean eventually perceives the colours of objects in her environment correctly. Block's reasons (in 'Inverted Earth') appear to include something like both the symmetry argument and the best theory of representation argument. I have pointed out, however, that the symmetry argument is no good, and that an intentionalist



can hold that Jean's chartreuse-feeling experiences do in one sense represent crimson. But apart from any such reason, the thought-experiment all by itself hardly compels one towards Block's conclusion.

How implausible it is to think that Jean, in the original telling, gets colour right emerges from the contrast with Fred. Part of the reason why it seemed at first plausible that Fred eventually perceives correctly is that even immediately after the inversion he has true beliefs about the colours of objects. After he begins to form these true beliefs non-inferentially, there is a pull to the notion that he sees colours correctly (albeit a pull that should be resisted). Two factors explain this pull. The first I have already considered: he arrives at his beliefs non-inferentially. And perception is plausibly a non-inferential means of acquiring beliefs about one's environment. Secondly, his beliefs are *true*. That Fred's true colour-beliefs are arrived at non-inferentially on the basis of perceptions suggests that his perceptions are true as well. But if he were very stupid, and became convinced immediately after the inversion that all of the objects in the universe had suddenly changed colour, then he would thereafter have false beliefs, non-inferentially based on false perception. And unless he changed his mind, it is, I submit, natural to think that Fred would go on continuing to believe that grass was now red, etc. Given that his beliefs are false, their being based non-inferentially on perception provides no reason for thinking that his perception is veridical. According to Block, it is plausible to think that unwitting Jean, who gets colours wrong initially, eventually gets them right. But in unwitting Jean's case, we cannot see her perceptions as mechanically falling into line, as it were, with her true non-inferentially formed beliefs, as we can in the case of Fred. For Jean starts by getting colours *wrong*. In Fred's case, his perceptions follow the lead of his beliefs; but Jean's beliefs provide no such lead. The case for spectrum inversion based on unwitting Jean is thus far weaker.

This difference between Jean and Fred also makes possible the following development. At some point long after the time during which Jean's perceptions (are alleged to) make the transition from false to true, the kidnapping is revealed to her. Jean, we can imagine, is startled. The colours of objects in the world around her, it would be natural for her to think, are not what they seemed to be. She asks for the colour-inverting lenses to be removed so that she can, she thinks, see inverted earth as it actually is. According to Block, when she does have them removed, and exclaims 'So *this* is what the world I've lived in for the last twenty years is *really* like!', she is mistaken. For she has now again begun to misperceive the world.

I find this interpretation of her condition implausible. I do not claim that its implausibility is decisive. But there is nothing in the thought-experiment itself that pushes us in the direction of Block's interpretation. What pushes

us in that direction, if anything, is externalist considerations, and our leading theories of representation. But, as I hope to have shown in §III, one can both reject the imaginability of the intentionalism-defeating inverted spectrum and accept externalism. And, as I hope to have shown in this section, we can also accommodate our best theories of representation without accepting the imaginability of the intentionalism-defeating inverted spectrum. We cannot, to be sure, allow that these accounts can serve as good theories of *perceptual* representation, but we can allow that perceptual experiences do (evidentially) represent just what these accounts say they do. And, if resistance to the imaginability of intentionalism-defeating inverted spectra is any guide, this is all we should allow.

## VI. A FINAL DIFFICULTY

There is a final difficulty for anyone who denies that one *just is* imagining (a) and (b) in imagining ( $1_n$ ) and ( $2_n$ ). To state it, I need elements from both the no-inference argument and the best theory of representation argument. In §V, I argued that the case for thinking that unwitting Jean is the victim of an intentionalism-defeating spectrum inversion is far weaker than for thinking that Fred is such a victim. My reason was that Jean (after the kidnapping), unlike Fred (after the operation), starts with false beliefs about the colours of objects in her environment. But, one might counter, even if unwitting Jean *initially* has false beliefs about the colours of objects on inverted earth, *eventually* she has true beliefs.

In support of this idea, one might reason as follows: our best theories of the representational properties of individual utterances tell us that they are determined in part by the meanings of the uttered words in the language of the utterance. Jean, upon her arrival at inverted earth, is a speaker of English. But it seems quite plausible that after enough time has past, she will unwittingly have become a speaker of English, the inverted earth version of English. In English, ‘chartreuse’ refers to crimson and ‘crimson’ refers to chartreuse. Thus after enough time has passed, Jean calls the colours in her environment by their right names. Furthermore, when she (correctly) states the colours of objects, she is expressing her beliefs about the colours of objects. If the calls are right, then so are the beliefs which the calls express.

If this line of reasoning stands, then intentionalism is in trouble. For it is then committed to the following three jointly implausible propositions. First, unwitting Jean forever misperceives the colours of objects in her environment. Secondly, she *simply trusts* her experience. (In this, she is unlike Fred from §IV, who comes automatically to associate apparently perceived

colours with the occurrence in the environment of their complements.) Thirdly, she has true beliefs about the colours of objects in her environment. These propositions are jointly implausible, because it is totally unclear how Jean could form *true beliefs* on the basis of simply trusting her misperceptions. Simply trusting her experience, after all, means that she treats the phenomenal character of the experience as partly constituted by the colour of the seen object. But since the phenomenal character of Jean's colour experiences is always wrong (according to intentionalism), treating the phenomenal character of the experience as to be partly constituted by the colour of the seen object should lead to false beliefs.

Fortunately, intentionalists have a way out. They can reject, with some plausibility, the view that unwitting Jean's utterances accurately express her beliefs about colour, and can argue that since Jean has unwittingly become a speaker of a different language, she has unwittingly come to misexpress her false beliefs with true assertions. For an analogy, imagine that Thomas Hobbes never died. With a few exceptions, he has kept abreast of changes in the English language, and takes himself to be, and is in fact, a speaker of modern English. But it has miraculously escaped his notice that 'diffident' now means the same as 'shy' rather than 'mistrustful'. So he can still be heard occasionally to remark that humans are 'diffident' by nature. I submit that it would be wrong to view Hobbes as expressing a belief that people are by nature shy. Rather, we should view him as misexpressing a belief that people are by nature mistrustful. Because he is unaware that the meaning of the word 'diffident' has changed, he is liable to this sort of misexpression. Similarly, because Jean is unaware that the meanings of colour words have changed, she too misexpresses herself.

The unwitting change in Jean's language thus renders true most of her utterances concerning the colours of objects only at the expense of knocking her beliefs out of whack with her utterances. More generally, such an unwitting change creates a mess. It is a consequence of this change that while her beliefs about the world are false, her beliefs about how the world appears to her are true. For although she believes falsely that inverted-earth oranges are orange, she believes correctly that they *appear* orange to her. Conversely, while her utterances about the world are true, her utterances about how the world appears to her are false. For she (falsely) states that oranges appear 'orange' to her, when they in fact appear orange (and yet are 'orange'). But this messiness does not count against the intentionalist. For it is just what one should expect in the bizarre scenario envisaged, a *scenario in which one comes to speak a foreign language without knowing it*. The difficulty sketched in this section thus does not, I contend, tell in favour of the imaginability of the intentionalism-defeating inverted spectrum.

Intentionalism-defeating inverted spectrum scenarios are not genuinely imaginable. Still, I have shown why many philosophers have been drawn to the idea that they are. There may well be imaginable scenarios in which perceivers regularly have visual experience which is phenomenally colour-inverted relative to what ours would be in the same environment. And there are a variety of reasons for taking these scenarios to be impossible if intentionalism is true. But I have shown that these reasons collapse under close examination.<sup>17</sup>

*Auburn University, Alabama*

<sup>17</sup> Thanks to Kelly Jolley, Ram Neta, Guy Rohrbaugh, Michael Watkins, participants in the University of Chicago's Wittgenstein Workshop in Winter 2003, and two anonymous referees.