

Spectrum inversion and intentionalism

PHIL 93507

Jeff Speaks

September 15, 2009

1	What is a spectrum inversion scenario?	1
2	Intentionalism is false because inverters could have in common	2
2.1	... their behavior and dispositions to behavior	2
2.2	... the functional roles of all of their states	3
2.3	... the fact that neither misrepresents the world	4
2.4	... the beliefs their experiences would non-inferentially justify	5
2.5	... the meanings of the color words in their language	5
3	Intrapersonal spectrum inversion	5

1 What is a spectrum inversion scenario?

A spectrum inversion scenario is one in which the phenomenal character of the experiences of a pair of perceivers systematically differ. (This is sometimes called ‘qualia inversion’, since what is being inverted is not the spectrum, whatever that might mean, but phenomenal character.) One can also consider inversion in aspects of phenomenal character other than color experience, but we will for the most part stick with inversions in the phenomenal character of color experience.

Spectrum inversion arguments are characteristically arguments against supervenience theses. It is argued that a pair of subjects can be spectrum inverted despite being the same with respect to some class of properties; it then follows that phenomenal character does not supervene on those properties.

Spectrum inversion scenarios differ in two ways: in what, exactly, is different about the phenomenal characters of the relevant experiences; and in what, exactly, is supposed to be held fixed between the relevant subjects.

2 Intentionalism is false because inverts could have in common ...

(For now, we will think of ‘intentionalism’ as ‘interpersonal intramodal local intentionalism.’ I will use ‘converse intentionalism’ as a name for the converse of this thesis. We’ve already discussed an argument for converse intentionalism, as well as a few arguments for intentionalism.)

Let’s use Invert and Nonvert as names for our pair of spectrum-inverted subjects. Then, since intentionalism is a thesis about the supervenience of phenomenal character on content, we will want Invert and Nonvert to be the same with respect to the content of their perceptual experiences.

This raises a problem: it is not easy to know how we can be sure that, in some imagined scenario, the content of a pair of experiences is the same. It is not obvious that this is the sort of thing which can simply be stated, or visualized. So the usual strategy is to come up with a scenario in which Invert and Nonvert share some class of base properties, on which content can be assumed to supervene, and yet differ in phenomenal character.

That is, the argument is some version of the following:

1. Possibly, there is a pair of experiences E1 of Invert and E2 of Nonvert such that (i) the phenomenal characters of E1 and E2 are inverted, and (ii) Invert and Nonvert are alike with respect to their *R* properties.
 2. Necessarily, if two perceivers are alike with respect to their *R* properties, then the contents of their experiences are the same.
-
- C. Possibly, there is a pair of experiences E1 of Invert and E2 of Nonvert which differ in phenomenal character but are the same in content.

which trivially entails the falsity of intentionalism.

The best way to proceed systematically is to imagine various candidates for the *R*-properties.

2.1 ...their behavior and dispositions to behavior

The most natural spectrum inversion scenario is one which identifies the properties held fixed with ‘behavior.’ This is the version that most often gets people thinking ‘How can I tell if apples look to you the way that grass looks to me?’ etc. If the contents of perceptual experiences supervened on behavior and behavioral dispositions, then it would suffice to refute intentionalism to imagine a scenario of this kind.

However, the claim that perceptual content supervenes on behavioral dispositions is, to put it mildly, controversial. A better version of the argument focuses on functional rather than behavioral duplicates (and the issues that arise in connection with

that version will also show what complications are in store for the hypothesis that behaviorally undetectable spectrum inversion is possible).

2.2 ...the functional roles of all of their states

So let's consider instead the hypothesis that Invert and Nonvert can be functional duplicates, and that the *R*-properties are the sum of their functional properties. Our first question is: is (P1) in the basic argument above true? Could functional duplicates be spectrum inverted?

To answer this question, we will need to get a bit clearer on what spectrum inversion involves. Construct a color circle (the Natural Color System hue circle) with yellow on the bottom and blue at the top, red and the left and green at the right. (These hues are chosen because they are unique, in the sense that they do not seem to be mixtures of the other hues.) When you look at this circle, there seems to be no problem with imagining it flipped: either blue/yellow flipped (as Locke imagined), red/green flipped, or both at the same time.

The main problem which arises with imagining functional duplicates who are spectrum inverted in this way arises from certain asymmetries in the NCS color space:

- purple appears to be a distinct hue, whereas yellowish green does not.
- the most saturated yellow is lighter than the most saturated blue.
- dark yellow is seen as a distinct hue, whereas dark blue is not.
- pink in purple are seen as distinct hues, whereas saturated and de-saturated greenish-yellow are not.
- 'hot' and 'cold' colors.

Two natural replies to these sorts of worries: (1) try to show that these asymmetries are avoided in creatures with limited color vision (e.g., just black/white vision, and brightness inversion) or (2) claim that there could be creatures with a symmetrical color space.

A problem with the idea of brightness inversion: when brightness increases, there is in addition to the obvious phenomenal difference (and corresponding representational difference that the scene seems more well-lit), an increase in the ability to discriminate between colors. So if we are sticking to the way that black/white inversion would actually work, inverts would be functionally discriminable: with the lights off, one would be able to discriminate colors that the other would not.

This pushes us again toward option (2): the claim that even if our color space is asymmetrical, we can imagine creatures who have a sense modality with a symmetrical quality space. Behaviorally undetectable inversion would be possible for those

creatures (whether or not it was spectrum inversion), and this could be used to raise all of the same problems.

If this really is possible, then it looks like functionalism and intentionalism are inconsistent.

A problem with the idea of ‘functional duplicates.’

A further worry about this style of argument: why doesn’t this just beg the question against functionalism, by assuming that a pair of people can be functionally identical yet differ in phenomenal character? Isn’t this a rather hefty premise to use in an argument against the functionalist intentionalist?

2.3 ...the fact that neither misrepresents the world

Locke’s presentation of the inverted spectrum argument relied heavily on the idea that neither would be misrepresenting the color of the objects in their environment. This view – that neither misrepresents the colors – might in turn be bolstered by various claims about what the spectrum inverts have in common:

- They’re equally good at getting around in their environment.
- They’re on par: there could be no reason for thinking that one was misrepresenting which was not also a reason for thinking that the other was misrepresenting.
- The idea that one is misrepresenting opens the door to radical skepticism about our own color experience.

One might reasonably wonder whether any of these is sufficient grounds for thinking that neither misrepresents.

It is also worth noting two things about this kind of argument. First, the parity version of the argument does not show that neither misrepresents, but only that one misrepresents iff the other does; so it leaves open the possibility that both misrepresent the world. So it is no problem for an intentionalist who is also an eliminativist about color.

Second, and more important, intentionalism does not strictly imply that one of the inverts must be misrepresenting, but only that there must be some difference in their representations. There are at least two ways in which one might hold the latter without the former: one might be a relativist (relationist) about the colors, and hold that objects have different colors relative to different observers; or one might think that the difference in content comes in something other than color representation. (For example, representation of “appearance properties.”)

2.4 ... the beliefs their experiences would non-inferentially justify

Even if we say that Invert perceptually misrepresents the colors of objects, Invert's beliefs about the world are surely true. (Various ways of arguing for this.) But then Invert tries to form a belief with the same content as the content of some perceptual experience, and fails to do so; but this is surely incoherent. We have first-person access to the relations between the contents of our perceptual experiences and the contents of our perceptual beliefs.

2.5 ... the meanings of the color words in their language

Suppose that some inverts are part of the same language community; then presumably both will say 'that's red' when confronted with an apple. Sure, since they are a part of the same language community, they must mean the same thing by their words. But then, according to the intentionalist, they must either be misrepresenting their own beliefs about the apple or their beliefs must fail to match in content the contents of the experiences on the basis of which they formed the belief. But either is implausible.

3 Intrapersonal spectrum inversion

One way to try to get around some of the worries with the above versions of the inverted spectrum argument against intentionalism is to construct an intrapersonal version of the case. This, as Block (1990), says, could go in three stages:

1. Immediate (and thus noticeable) spectrum inversion.
2. Complete semantic adaptation.
3. Complete amnesia, so that one acquires just the behavioral dispositions/functional architecture one had before stage 1.

Further, one might think that if a person can differ at two times in this way, then there can be no in principle reason to resist spectrum inversion between distinct functional duplicates.

It seems that there's no room for the functionalist to insist that there is no phenomenal difference between and after step 1 — after all, the person notices the difference.

Is it plausible for the functionalist to say that the phenomenal character of color experience switches back after stage 3 to what it was prior to stage 1?

None of this helps with the worries about asymmetry mentioned above.

References

Ned Block, 1990. Inverted Earth. *Philosophical Perspectives* 4:52–79.