

On the correct treatment of Inverted Earth

Gary Bartlett

Abstract. The Inverted Earth case has seen fierce debate between Ned Block, who says it defeats the causal-covariational brand of wide representationalism about qualia, and Michael Tye and Bill Lycan, who say it does not. The debate has generated more heat than light because of a failure to get clear on who is supposed to be proving what, and what premises can be deployed in doing so. I argue that a correct understanding of the case makes it clear that the causal covariation theory is in deeper trouble over Inverted Earth than is generally supposed even by the theory's detractors.

1. Introduction: Inverted Earth for dummies.

In his oft-cited 1990 paper 'Inverted Earth', Ned Block argues that conscious experience has intrinsic qualities – that the *qualia* (as I shall call them) or *phenomenal character* of our experiences cannot be analyzed in functional or intentional terms. Block explains that the paper began life as a response to Gilbert Harman (1990). Harman is a *wide representationalist*: an externalist about mental content who is a representationalist about qualia. (*Narrow representationalism*, which I do not discuss here, endorses representationalism but denies content externalism.) He holds that a perceptual state's qualia are exhausted by its wide representational content. For example, the color that your visual experience of the sky represents the sky as being, and thus the color qualia of that experience, Harman takes to constitutively depend on what color the sky *is* in your normal environment.

On such a view, moving to a new environment where the sky is a different color *must* sooner or later result in your experiences of the sky having a different color content, and therefore different color qualia. The Inverted Earth case compares the color experience of a person before and after such a move. On Inverted Earth, colors are inverted with respect to Earth colors: grass is

red, the sky is yellow (or orange, if you like), and so on. People on Inverted Earth, however, make the same linguistic utterances as we do: neighbors point out to each other how “blue” the sky is today, discuss whose lawn is a more pleasing shade of “green”, and so on. Block’s tale then goes like this:

A team of mad scientists knock you out. While you are out cold, they insert color-inverting lenses in your eyes, and change your body pigments so you don’t have a nasty shock when you wake up and look at your feet. They transport you to Inverted Earth, where you are substituted for a counterpart who has occupied a niche on Inverted Earth that corresponds exactly (except for colors of things) with your niche at home. You wake up, and since the inverting lenses cancel out the inverted colors, you notice no difference at all. (Block 1990, p. 63)

Block then argues that although the representational content of your visual experiences will eventually invert to match the colors in your new environment – your experiences of grass, for example, will come to have the content ‘red’ – nevertheless the *qualia* of your visual experiences will not change. Block thinks the case therefore shows that Harman’s wide representationalist view cannot be correct.

For the sake of clarity, I shall put the case in programmatic form as follows. It makes a simple comparison of two instances in the life of an Inverted Earth ‘abductee’:

Instance (1): The abductee at t_1 on Earth, prior to her abduction.

Instance (2): The abductee at t_2 on Inverted Earth, with her visual experiences’ color content inverted.

It is stipulated that time t_2 is long enough after the abduction that the color content of the abductee’s experiences would have shifted as a result of her new surroundings.¹ So assuming that the *content* shift has occurred, one of two judgments must be made concerning the abductee’s color *qualia* across the two instances. Those *qualia* would either remain the same (e.g., the sky would look just the same color on Inverted Earth as it did on Earth), or they would be different:

[SAME] “The abductee’s experience would have the same color *qualia* at t_2 as at t_1 .”

[DIFFERENT] “The abductee’s experience would have different color qualia at t_2 than at t_1 .”

Anyone who holds that qualia supervene partly on the subject’s normal environment must endorse [DIFFERENT], since on such a view the inverted colors on Inverted Earth would result in abductee’s color qualia inverting. But in his 1990 and in a string of papers since, Block (1994, 1996, 1998, 2003) has argued that [SAME] is correct, and has persistently pressed the Inverted Earth case against wide representationalists, especially Bill Lycan (1996) and Michael Tye (1995, 2000).²

A wide representationalist may endorse [SAME] if he adopts a theory on which it is not the *subject’s* environment that determines his perceptual content. Fred Dretske (1995) holds that an experience’s content, and thus its qualia, depends on what the experience indicates (that is, what information it carries) when the relevant sensory system is functioning as natural selection designed it to function. For him, what matters is the environment of the subject’s evolutionary ancestors – the environment in which the sensory system evolved into its present form. Since moving to Inverted Earth does not change the distribution of colors in the environment in which one’s sensory systems evolved, Dretske endorses [SAME]. I shall discuss only the wide representational theory that entails [DIFFERENT], which I dub the *causal covariation theory* or CC-theory for short – since it says that qualia covary with the features of the environment with which the subject normally interacts. This theory applies the work of Putnam and Burge on the content of propositional attitudes (e.g., Putnam 1975; Burge 1979, 1986) to conscious experiences.

On paper, neither Tye nor Lycan is presently a CC-theorist. Tye, in *Consciousness, Color, and Content* (2000), favors a mixed covariation-etiological theory that has him joining Dretske in endorsing [SAME]. Lycan, in *Conscious Experience* (1996) does not commit to a particular theory. (Henceforth, all citations of Lycan and Tye will be to these two books.) Nevertheless, I think the Inverted Earth case is worth discussing. There’s a general reason, and a specific reason.

The general reason is that we know that having once fallen from favor, a philosophical idea will often be taken up again later, whether in renewed hope for its success, despair over the al-

ternatives, or just lack of attention to the reasons it was dropped the first time. By getting clear now on how bad things look for the CC-theory on Inverted Earth, we may avoid having to go over the same ground again in the future.

The specific reason is that Lycan and Tye both devote significant space to arguments responding to Block on the Inverted Earth case. So even if they do not officially hold the CC-theory, they clearly still think it has no trouble with anything Block says about Inverted Earth. Tye ultimately abandons the CC-theory because he finds that it cannot explain what color qualia the abductee's experience of (e.g.) the sky would have during the *transition* from blue to yellow qualia (pp. 134-135). In writing this paper I do not mean to imply that I am unconvinced by this argument; to the contrary, the problem it poses seems serious. However, I want to dispose of the idea that this is the *only* good argument against [DIFFERENT]. Before offering the argument, Tye spends nearly twelve pages expounding a "possible hard-line" defense of [DIFFERENT] (p. 123). This indicates his belief that the CC-theory *would* be viable if we had an account of the qualia transition. Lycan, meanwhile, finds the transition issue wholly untroubling (p. 129), so he presumably regards the CC-theory as a live option. So in case someone finds a way to finesse or defuse the qualia transition problem, it is beneficial to elucidate why the debate need not even get that far. There is no need to consider what quale might be associated with the content 'blue or yellow' if there is no reason to think that the abductee's qualia would shift in the first place. I shall argue that nothing Tye or Lycan say should make us think that [DIFFERENT] is true, and therefore that they have offered no reason to think that the CC-theory is a live option.

2. The dialectical role of the Inverted Earth case.

Both Lycan (p. 113) and Tye (p. 117) seem to construe Inverted Earth as a case devised by Block as an attempted counterexample to the CC-theory. In particular, they take it that the case comes with a built-in assumption that the abductee's qualia on Inverted Earth would be the same as on Earth. That Tye makes that construal is clear in the following: "Block is arguing that repre-

sentationalism fails since it cannot handle the Inverted Earth example. This example assumes that things now look phenomenally just as they used to look” (p. 129). Tye is wrong. Block does, of course, judge that things would continue to look the same color to the abductee; but that judgment is not built into the case itself. The judgment – i.e., [SAME] – is what Block *concludes* from his analysis of the case. The Inverted Earth case can be fully described without including the claim that the abductee’s qualia remain the same. One simply describes the abduction and the insertion of the lenses, and then adds the CC-theorists’ own stipulation that the content of visual experience constitutively depends on the environment in which the subject is embedded (and has been embedded for whatever period externalists deem sufficient for content to be affected). That’s all. The CC-theory entails that the abductee’s color qualia will shift, so CC-theorists are forced to defend that entailment. All that their opponents have to do is supply the set-up.

The key point is this: it is central to the dialectical force of the Inverted Earth case that the purported qualia inversion occurs *in virtue of* the inversion of the color content of the abductee’s visual experiences. Since the CC-theory dictates that the content inversion will (at some point) occur, it entails that the qualia inversion will occur too. The Inverted Earth case thus challenges the CC-theorist to explain why [DIFFERENT] is true, or at least why it is plausible. If [DIFFERENT] cannot be made plausible independently of the CC-theory, the theory is left without an evidential leg to stand on. The CC-theorist must do more than cast doubt on the argument for [SAME]; he must give us an argument for [DIFFERENT].³

The dialectical tangle over who needs to establish what might have been avoided if Block had more consistently advertised the origins of the Inverted Earth case. It was invented by Harman (1982). Block (1990) acknowledges this, but (perhaps oddly) does not say that Harman offered the case as an *illustration* of wide representationalism. Wider recognition of this fact would have done a lot to clarify where the burden of proof lies in the Inverted Earth debate. There is a perception that the burden of proof is on Block, but in fact it is on the representationalist.

The representationalist’s bottom line is that qualia will *track* phenomenal content: i.e., that a shift in the phenomenal content of a subject’s experiences will be mirrored by a shift in their

qualia. (By ‘phenomenal content’ I refer to the representational content of experiences.) Harman introduced Inverted Earth to illustrate exactly this sort of claim (as part of a wide conceptual role semantics; see also Harman 1988). Here is how he uses the case:

Consider someone on Inverted Earth who perfectly adapts to color inverting lenses.⁴ Looking at the sky of Inverted Earth, this person has an experience of color whose physical basis is the same as that of a normal person on Earth looking at Earth’s sky. But the sky looks orange to the person on Inverted Earth and blue to normal people on Earth. What makes an experience the experience of something looking the color it looks is not its intrinsic character and/or physical basis but rather its functional characteristics within an assumed normal context. (1982, p. 251)

For Harman, the Inverted Earth case illustrates the claim that a subject’s visual qualia can shift in accord with a shift in the content of his visual experiences, *even if there is no relevant change in his physical make-up*.

It is no accident that the immediately previous section of Harman (1982) is an exposition of Twin Earth cases. *Inverted Earth is a Twin Earth case*. Block has described it as such (1996, p. 41), but he still does not emphasize the effect of this fact on the case’s dialectical role.⁵

The point of a Twin Earth case is to be a counterexample to the internalist claim that mental states supervene on the intrinsic physical properties of the subject’s body. The counterexample is constructed by showing that two subjects, or one subject at two times, can be mentally distinct despite being physically identical (in the relevant ways). If the two subjects in a given Twin Earth case were in fact physically different in some relevant way, then the case would lose its force, for an internalist could then explain the mental differences in terms of the internal physical differences between the subjects. A good Twin Earth case does not permit an internalist explanation of the purported mental difference. An opening for such an explanation exists if the subjects differ in any intuitively relevant physical respects. Thus Putnam’s classic case stipulates that H₂O and XYZ are indistinguishable (barring chemical analysis), so that a subject who was instantaneously abducted to Twin Earth would not *notice* that the stuff coming out of faucets was

different than before, and so would continue acting just as if he had never been abducted at all. Therefore, there would be no relevant physical differences between the bodily movements of an Earth subject and those of his Twin Earth 'twin'. (The difference between H₂O-based physiology and XYZ-based physiology is not thought to be relevant to the subjects' mental lives.)

Parallel considerations apply to Inverted Earth. For the case to be a counterexample to internalism about qualia, the move to Inverted Earth cannot cause the abductee to physically change in any way that is relevant to her experience of colors. Here is why. If the abductee ever says anything like, "Hey! Everything has reversed in color!", then there has to be a physical explanation for her making of that remark, which we assume she would not have made had she remained on Earth. Since environmental conditions cannot cause an organism to perform certain actions, including verbal actions, except by affecting its physical body in some way, there must have been changes within the abductee's body that were caused by the move to Inverted Earth. Those physical changes will do nicely as a hook on which to hang an internalist explanation of the abductee's changed visual qualia. It is precisely to avoid this outcome that Harman stipulates that the abductee is given color-inverting lenses. The lenses have no analog in the original Twin Earth cases. They are introduced so that, as Harman says, the abductee "has an experience of color whose physical basis is the same as that of a normal person on Earth" (1982, p. 251).

The Inverted Earth case has a job to do for the CC-theory of experience that is analogous to the job Twin Earth cases do for Putnam/Burge-style externalism about thought content. Twin Earth cases are, to use Daniel Dennett's apt phrase, externalism's main 'intuition pumps': they illustrate the plausibility of the idea that a change merely in a subject's environment can entail a change in the subject's mental content. Were they not as compelling as they are (by which I just mean that the cases warrant serious consideration) externalism itself would not be compelling. Twin Earth is thus a key testing ground for externalism. A Scotch verdict on these cases will not do for the externalist. Twin Earth cases have to put some positive weight behind the claim that two subjects could be internally identical but mentally distinct. If that claim is not plausible in

those cases, devised by externalists themselves, then it is very likely not going to be plausible in any other cases.

The same is true of the Inverted Earth case and the CC-theory. For the CC-theory to be plausible, it must be possible for qualia to shift even while the subject remains physically unchanged in any relevant way. With the Inverted Earth case, Harman (1982) tried to illustrate just such a qualia shift. According to him, a shift in the abductee's environment drives a shift in her content and thus a shift in her qualia. So Inverted Earth is a key testing ground for the CC-theory of qualia, just as Twin Earth is for externalism about thought content. This is why the construal of the Inverted Earth case that I described earlier in this section – that the case is merely an attempted counterexample to the CC-theory – is wrong. That construal leads to the view that all that CC-theorists are obliged to do is undermine the counterexample by showing that Block's argument for [SAME] is unsound. Yet this is just as if the only thing an externalist were to say about a Twin Earth case were that the internalists' argument that the two subjects' thoughts have the same content is unsound. To say *only* this would be to completely misunderstand the dialectical role of Twin Earth cases. Externalists must hold that the cases give positive evidence that the subject's thoughts have *different* content. Likewise, it is not enough for CC-theorists to argue just that Block has not shown good grounds for adopting [SAME]. They must show us good grounds for adopting [DIFFERENT].

So CC-theorists face a problem. Very few philosophers, even other externalists, find [DIFFERENT] intuitively plausible. (Burge [2003, p. 405], for example, says that representationalism “has long seemed to me to resist belief”, and that he sees no good arguments for it.) Most observers agree with Block that [SAME] is correct. Further, that position can be supported by an argument that does not seem available in Twin Earth cases. Where there is little inclination to think that a subject's failure to notice or react to a shift in the content of his thoughts shows that the content of his thoughts hasn't changed, there is great readiness to think that a subject's failure to notice or react to a shift in the character of his experiences shows that the qualia of his experiences hasn't changed. If the only support for [SAME] were that the internal supervenience of qualia en-

tails it, then CC-theorists might have been able to declare a stalemate.⁶ But that isn't the case. Where regular externalists use Twin Earth cases as the intuitive motivation for their externalist theories, CC-theorists find that the Inverted Earth case doesn't pull off the same trick. They find themselves at risk of having to use their theory to motivate their interpretation of the Inverted Earth case. That is just the reverse of how things are supposed to go, and it is, of course, unacceptable. CC-theorists must find support for [DIFFERENT] elsewhere, or accept that Block wins the day.

In the rest of this paper I consider three arguments that Lycan and Tye offer in response to Block's claims concerning Inverted Earth. I shall show that none of the arguments succeeds in providing evidence for [DIFFERENT]. So, since the CC-theory says that [DIFFERENT] must be true, yet there is no reason to think it is true, we have no reason to believe that the CC-theory is true.

3. The externalist memory argument.

Both Tye (p. 123) and Lycan (p. 129) suggest that, although the abductee's color qualia would invert, she would fail to notice this fact because her memories of Earth colors would *also* have inverted – because (in turn) the content of memories supervenes widely. The idea is that the abductee would remember, for example, that tomatoes used to look the color she calls 'red'; and 'red' for her on Inverted Earth would have come to mean green, since the content of her thoughts would have inverted. Thus, her visual memory of Earth tomatoes would have the content that they were green. So even though her color qualia would have inverted – tomatoes looked red to her on Earth, but would look green to her on Inverted Earth – she would notice no difference, for her memory would wrongly inform her that tomatoes had always looked green. In short, the claim is that the inversion of the color qualia of the abductee's *memories* would prevent her from noticing the inversion of the qualia of her *present* experiences.

At this point the exchange gets a bit messy. Block (having heard this sort of line at a 1995 conference) protests that the externalist memory argument begs the question, since one cannot

defend wide representationalism about qualia by bluntly invoking wide representationalism about the qualia of memories (1996, pp. 44-45). He is surely right about this. However, it is equally true that one cannot *attack* wide representationalism about qualia by bluntly *denying* wide representationalism about the qualia of memories; and I suspect it is this point that Lycan (p. 132) and Tye (p. 129) have in mind when, in counter-response to Block, they accuse *him* of begging the question. So the direction of the dialectic matters here. In invoking wide representationalism about the qualia of memories, are Lycan and Tye playing defense or offense? The tangled to-and-fro of the dispute as it has played out in print makes this question rather difficult to answer unequivocally, but fortunately we may side-step it by turning it into a dilemma. Whichever is the case, the CC-theory loses.

Suppose Lycan and Tye are, as at least Lycan seems to think, playing defense.⁷ That is, suppose they are just pointing out that Block, in arguing for [SAME], cannot assume that the qualia of memories supervene narrowly; they are *not* assuming on their own behalf, in arguing for [DIFFERENT], that the qualia of memories supervene widely. Well if this is the dialectical situation, then the CC-theory is being left undefended. As I argued in the previous section, CC-theorists *need* to play offense; if they aren't doing so, then they should be. It is not enough for them to say that we have insufficient grounds for [SAME]. That sort of line may work against typical inverted spectrum cases (see n. 5), since there it will do if the representationalist shows that there is no good evidence that spectrum inversions are possible. But he cannot take that line regarding Inverted Earth. As far as wide representationalists themselves are concerned, the Inverted Earth case is no less possible than the Twin Earth cases that externalists have been deploying for thirty years. Regarding Inverted Earth, then, CC-theorists have to play offense. They have to try to convince us that [DIFFERENT] is true.

So let's suppose that in offering the externalist memory argument Lycan and Tye are, in fact, playing offense, trying to establish [DIFFERENT]. Of course if that is the case, Block is right to charge them with begging the question in claiming that the qualia of the abductee's memories would invert. Lycan and Tye themselves seem to realize that this move on their part would be

less than kosher. Lycan insists (p. 131) that in putting forward the externalist memory argument he is invoking the wideness only of memory *contents* (which Block himself accepts), not of memory *qualia* – which implies that he knows he can't get away with invoking the latter. Tye, for his part, initially argues from the wideness of memory contents (p. 129), but then, presumably realizing that Block's objection is to the claim of the wideness of memory *qualia*, provides what I call the *reconstructive memory argument* (see my §5) in an attempt to give independent support to that claim.⁸

Lycan's claim to have invoked only the wideness of memory contents seems disingenuous. The claim that memory contents are wide will have no relevance to the Inverted Earth case unless it *does* entail that memory qualia are also wide, since what is at issue is the color qualia of the abductee's memories of Earth objects. The externalist memory argument clearly requires the premise that phenomenal memory is externalist, and Block is correct that the CC-theorist needs to defend that premise. Neither Block nor I have any special quarrel with the idea that declarative memory, for propositions such as 'the stuff in the lakes is water', is externalist (see Heal 1998 for someone who does). But the CC-theorist cannot just stipulate that if declarative memory is externalist, then so is *visual* memory (e.g., of the colors of objects). We need an independent argument for that entailment, and neither Lycan nor Tye provide one.

And actually, things are much worse than that. I have not yet mentioned the most serious problem with the externalist memory argument, though some readers may have spotted it already. It is this: even if the CC-theorist *can* produce an argument that the wideness of memory contents entails the wideness of memory qualia, he will still not have provided an argument for [DIFFERENT] itself.

I wouldn't blame anyone who was a bit lost at this point. As I said, this exchange has gotten messy. So here's a précis of where we are. For the Inverted Earth case to be any use to the CC-theorist in showing that qualia supervene widely, he needs to show independent reason to think [DIFFERENT] is true. So we are considering the externalist memory argument, which says that an inversion of the abductee's qualia could occur without the abductee's awareness because her

memory of Earth colors would mislead her. A problem with the argument is its assumption that memory qualia are wide.

What would a defense of this assumption be like? Well, it would either show that *all* qualia (of both present and remembered experiences) are wide, or it would show only that *memory* qualia are wide. If the former, then the defense would beg the question again, since the claim that qualia are wide is what the Inverted Earth case is meant to show. So the defense of the assumption that memory qualia are wide would have to be a defense of just that specific assumption. Now maybe such a defense can be given (Tye offers an argument that he seems to think constitutes such a defense, but I argue in §5 that it is far off-target).⁹ But – and here is the crucial point – even if it could be given, it would not show that the abductee’s color qualia on Inverted Earth would differ from those she had on Earth. For [DIFFERENT] to be true, the qualia of the visual experiences the abductee would have while on Inverted Earth must be inverted with respect to the qualia of the experiences she had while on Earth – e.g., ripe tomatoes on Inverted Earth would have to look green to her. The externalist memory argument, supported with a defense of its assumption that memory qualia are wide, does not show this. All it shows is that *if* tomatoes on Inverted Earth would look green to her, she would not notice the shift in the way they look (because she would remember tomatoes on Earth as also looking green, because of the wideness of phenomenal memory). *The antecedent remains unsupported*: we have no argument that the Inverted Earth tomatoes *would* look green to her.

In sum, the externalist memory argument either begs the question, or it cannot show that the abductee’s qualia would invert in the first place, which is what the CC-theorist actually needs to show. It seems that Lycan and Tye become so focused on attacking Block’s reasons for [SAME] that they omit to defend [DIFFERENT]. I think they make this omission because they are under the spell of the false interpretation of the dialectical role of the Inverted Earth case that I described in §2: i.e., that the case is simply a counterexample to the CC-theory, and thus that all they need to do is cast doubt on the argument for [SAME]. The externalist memory argument, suitably bolstered as described, certainly would weaken the argument for [SAME], which relies on the prem-

ise that the abductee would surely notice if her qualia had inverted. But the externalist memory argument is simply not an argument for [DIFFERENT]. And what the CC-theorist needs is an argument for [DIFFERENT].

4. The gradual shift argument.

It is ironic that both Tye (p. 129) and Lycan (pp. 130ff) accuse Block of confusion over who should be arguing for what concerning Inverted Earth. To be sure, Block has not been optimally clear about the dialectic surrounding the Inverted Earth case (see §2); but Tye and Lycan lose sight of the goalposts in a much more extensive way. They offer two further arguments that, as well as having the same flaw as the externalist memory argument (i.e., the flaw of not actually being an argument for [DIFFERENT]), have a second flaw: the arguments say that the abductee *might not* notice an inversion of her color qualia, when what is needed is an argument that she *could not* notice such an inversion. So Tye and Lycan appear to believe that it can help the CC-theorist's case to argue that the abductee *might not* notice her qualia inversion. It cannot.

Lycan suggests that a gradual content shift might be matched by a gradual qualia shift, which would go unnoticed by the abductee just because it would happen gradually. This is the *gradual shift argument*. Both Tye and Lycan admit that for the CC-theory to be plausible, the inversion must be gradual, to allow an explanation of why the abductee does not notice it. (David Chalmers [1996, pp. 266-269] argues cogently that the idea of phenomenal colors rapidly switching without the subject's awareness is incoherent.) Moreover, as Tye notes (p. 135), no one wants to hold that the abductee's visual state could *instantaneously* shift from covarying with blue to covarying with yellow.¹⁰ Lycan's idea, then, is that the abductee's color qualia might very gradually invert, and therefore that it would be beyond the limited discriminative capacities of the human visual system to pick up on the inversion. (On another reading, the idea is that the inversion would be beyond the discriminative capacities of all but the most attentive and visually

perspicacious of humans. It's not clear which Lycan intends. But since this reading is the weaker of the two, any response that succeeds against the first reading will succeed against this one.)

The first point to note about the gradual shift argument is that its conclusion is that *if* [DIFFERENT] is true, the abductee still might not notice that her color qualia had inverted. So in making it, Lycan again seems to be laboring under the misconceived dialectical view of the Inverted Earth case that I described in §2. The second point is that the argument's conclusion is only that *if* [DIFFERENT] is true, the abductee *might* not notice that her color qualia had inverted; and this conclusion does nothing to support the CC-theory. Let me explain this second point in detail.

Suppose that before abducting a given person, Kay, we surgically upgrade her visual system so that she *can* make the discriminations necessary for her to notice a gradual inversion in her color qualia. Assume that the inversion happens (i.e., that [DIFFERENT] is true of Kay). Kay is able to spot the inversion and remark on it: at some point she says that "tomatoes look a different color than they used to". Now the explanation for this utterance must advert to some physical changes within her body (see my discussion of the methodology of Twin Earth cases in §2). That means the CC-theorist is in trouble. Of course, *normal* humans, if abducted, would still not notice that *their* qualia had inverted. But the *reason* they wouldn't notice this is that their visual systems haven't been upgraded. And since this reason appeals to facts about human neurophysiology, the internalist may use it to say one of two things in explanation of how it is that Kay's color qualia are now different than they used to be.

Suppose first that *only* Kay would undergo the inversion: that is, suppose that [DIFFERENT] is true of Kay, but [SAME] is true of normal abductees. This is fine for the internalist, since the explanation of the difference is an internalist one: i.e., Kay's upgraded visual system physically differs from that of a normal human. Alternatively, suppose that although *any* human abductee would undergo the inversion – i.e., [DIFFERENT] would be true of all such abductees – only Kay can *notice* it. This too is fine for the internalist, who may reason as follows. The inversion itself has to be the same kind of phenomenon in all abductees, for it's caused by the same event: the move to Inverted Earth. So it must receive the same explanation in all abductees. Well, as we

said in the previous paragraph, our explanation of Kay's case will advert to physical changes in her body. So the same changes must occur in all abductees, causing the inversion in all of them. The only difference with Kay is that, because of her upgrade, her visual system can pass on information enabling her to *notice* her inversion, which is not the case with other abductees.

Lycan might try arguing that *no possible creature* could notice the inversion. But we would be entitled to ask why not; and his answer could not be that the inversion is a purely relational phenomenon, for that is just the claim that the Inverted Earth case (*qua* Twin Earth case) is supposed to make plausible. Nor may he object that only normal humans can be abductees. Externalism is not meant to be true only of creatures with certain kinds of psychological mechanisms. It's meant to be a metaphysical truth about *all* minded creatures (or at least, all minded creatures in causal interaction with an environment). For the CC-theorist, the reason that an abductee's qualia inversion would not be noticed must be that it would be a purely relational change – one that could occur without any change in the abductee's physical state. This must be true of all possible abductees, human or otherwise. If some possible creature (such as Kay with her upgrade), when abducted to the Inverted version of its home, could notice the resulting inversion of its color qualia, then the externalist explanation for the inversion can be pre-empted by an internalist one.

Thus where the CC-theorist needs to support the proposition that the abductee *would not* notice that her color qualia had inverted, the gradual shift argument can only show that the abductee *might not* notice that her color qualia had inverted. But this latter claim permits an internalist explanation of how it is that the abductee's qualia have shifted, and thus provides no support for the CC-theory.

5. The reconstructive memory argument.

The final argument I wish to discuss is offered by Tye, ostensibly in support of the claim (essential to the externalist memory argument – see §3) that the qualia of remembered experiences are wide.

Tye (pp. 129-130) correctly points out that empirical evidence from the study of memory indicates that recalling an image is a *reconstructive* process. It is more like making a sketch than pulling up a stored photograph. So, Tye infers, we should expect our present concepts to influence the content of recalled images, thus impeding our ability to correctly judge whether the colors of things have changed. The color qualia of the abductee's memories of Earth could in fact invert as a result of the reconstructive process, with the further result that she would then not be able to tell if the qualia of her *present* experiences have inverted.

Again, like the gradual shift argument, this *reconstructive memory argument* has the flaw that it can only support the conditional claim that *if* [DIFFERENT] is true, the abductee still might not notice that her color qualia had inverted. It is not an argument for [DIFFERENT] itself, which is what CC-theorists really need. So the argument's usefulness to the CC-theorist is extremely dubious.

But again like the gradual shift argument, the reconstructive memory argument has a further flaw. It does not even entail that it is metaphysically impossible for the abductee to notice the inversion of her qualia. It only entails that she *might not* notice it, if her visual memories of Earth are reconstructed using color concepts influenced by her Inverted Earth environs. It is not metaphysically mandated that this will happen; it is only empirically possible or probable.

The empirical theory of reconstructive memory cannot support the metaphysical claim that memory qualia are externalist. It is strange that Tye offers the reconstructive memory argument, for he himself stresses that the CC-theory entails that an abductee's comparison of her present and past experiences "*must* be mistaken" (p. 129, orig. emphasis): i.e., that her Earth memories *must* (eventually) feature inverted color qualia. So invoking empirical laws about memory cannot

help to bolster the CC-theory. The reconstructive account of memory allows that some abductees might well recall Earth colors correctly, and so would be able to notice the difference between colors now and colors before. (Or imagine again surgically upgrading an abductee, this time so that her memories are photographic rather than reconstructive.) And as I explained in the previous section, if even one abductee is able to notice her qualia inversion, then the inversion itself must be such that it can be explained in internalist terms.

So again, like the gradual shift argument, the reconstructive memory argument can only show that the abductee *might not* notice that her color qualia had inverted. Yet the CC-theorist needs to support the proposition that the abductee *would not* notice that her color qualia had inverted. The reconstructive memory argument simply cannot support this proposition.

6. Conclusion.

My purpose has been simply to show that the CC-theory is much further behind the eightball on Inverted Earth than the discussion that the case has received would indicate. I have argued that CC-theorists have a dialectical obligation to defend [DIFFERENT] rather than just to oppose the arguments for [SAME]. Yet none of the three arguments that Lycan and Tye offer are capable of the former; indeed two of them do not even support the claim that, if the abductee's color qualia *were* to invert, the abductee would not notice the inversion. In sum, the Inverted Earth case is more than just a counterexample to the CC-theory. The CC-theorist must defend the claim that [DIFFERENT] would be true of all Inverted Earth abductees, but we have been given no independent reason to believe that this is so. This means we have no independent reason to believe that the CC-theory is correct.¹¹

Department of Philosophy and Religious Studies

Central Washington University

Notes.

¹ In his 1996 Block makes a small change to avoid the complaint that if the abductee were unaware of having changed location, then since she could not have decided to defer to the linguistic customs of her new community, her content would remain as it was. The amended story is that instead of being kidnapped, the abductee knows she has traveled to Inverted Earth (and had the lenses implanted), and decides to adopt the concepts and language of her new community. The reader may imagine I am dealing with the amended story if he or she prefers.

² Tye (2003, p. 7) and McLaughlin (2003, p. 134n) say that the Inverted Earth case is not a counterexample to *weak representationalism* (WR), the thesis that qualia supervene on phenomenal content. WR says only that a shift in qualia must be accompanied by a shift in phenomenal content; yet Block's claim is precisely that the abductee's qualia do *not* shift as a result of the move to Inverted Earth. Hence, say Tye and McLaughlin, the case is no use against WR; logically, it cuts only against *strong representationalism*, the thesis that qualia are *identical* with phenomenal content. (McLaughlin, though, is no representationalist; he goes on to argue against WR with inverted spectrum cases.) However, Macpherson (2005, p. 131) points out that as well as comparing the abductee's experience of something on Earth with his experience of the analogous thing on Inverted Earth (e.g., Earth grass versus Inverted Earth grass), we can compare his experiences of non-analogous things, such as Earth tomatoes and Inverted Earth grass. The abductee's experiences of the latter two things would purportedly have the same color *content* (i.e., red – since Inverted Earth grass is red) but, according to Block, different *qualia* (Inverted Earth grass would purportedly look green, since the lenses would cancel out the inverted colors). And that is just what WR says cannot happen. So if Tye and McLaughlin mean only to talk about 'analogous-object' comparisons, they are strictly correct that the Inverted Earth case is irrelevant to WR; but the point is inconsequential in view of the fact that the Inverted Earth proponent may easily switch to a non-analogous object comparison.

³ Things would be different if the argument for [SAME] were question-begging, as Lycan (p. 115) has claimed. He accuses Block (1990) of *assuming* that qualia supervene internally. I do not think the accusation can be grounded in what Block says. But in any case, by 1994 Block was explicit that he did not intend (or no longer intended) to rely on that assumption: "What justifies the idea that your qualia are the same is that you notice no difference... after the switch: no appeal to supervenience [of qualia on the brain] is required" (p. 518). Stalnaker (1996, pp. 102-104) similarly notes that the key premise is that the abductee's qualia after the move remain introspectively indistinguishable from her qualia before the move: if told that she had been relocated to Inverted Earth, she would be unable to say when it had happened. Lycan himself admits (p. 126) that this argument does not assume that qualia supervene internally. It says only that if the abductee would not *notice* any change in her experience, and would rebuff attempts to persuade her that such a change had occurred, we may infer that the move *makes* no difference to her experience. The key premise is not that qualia supervene internally, but that big shifts in the character of experience – such as an inversion of color qualia – are noticeable by the subject. So Block's argument for [SAME] does not contain the claim

that qualia supervene internally. The argument's conclusion – that the abductee's qualia remain the same despite the change in her environment – might be taken as evidence for that claim, but that's all.

⁴ Harman imagines giving the lenses to an Inverted Earth inhabitant rather than a person from Earth. Actual subjects who don *spatial*-inverting lenses that turn the world 'upside-down' take a while to adapt to them, but then become so well-adapted that things seem 'normal' again – indeed, removing the lenses then causes another period of disorientation. Harman assumes that donning color-inverting lenses would cause similar initial disorientation, and thus would require a period of adaptation. But Block's story eliminates the period of adaptation to the lenses, since the inversion they cause is cancelled out by the inversion of colors in the environment.

⁵ In his 1994 Block calls the Inverted Earth case a “variant of the inverted spectrum thought-experiment” (p. 517). This is a bit misleading. His 1990 description was more accurate: there he bills it as “the ‘converse’ of the usual inverted spectrum case” (p. 682), and notes that where the typical intrapersonal inverted spectrum case inverts the subject's qualia while holding her functional and intentional state fixed, the Inverted Earth case keeps the abductee's qualia the same while manipulating the functional or intentional description of her visual experiences.

In the typical inverted spectrum case used against a representational theory of qualia, the spectrum inversion is imposed in direct contradiction of the targeted theory, which says that qualia cannot shift unless content does. It is part of the set-up of the case that the subject's qualia invert with respect to their content. So representationalists must deny that the inversion is even possible, for if they allow its possibility, the game is up. Unsurprisingly, this is just what they do (cf. Harman 1990, p. 48). Now if we say that the Inverted Earth case is merely a *variant* of the typical inverted spectrum case, we make it seem that all the CC-theorist needs to do in regard to the former is pretty much what he needs to do in regard to the latter: deny that the case is possible. Yet as we have seen, this is hardly likely to go well for the CC-theorist, given that the Inverted Earth case was *invented* by a representational theorist to *defend* a representational theory. And as I'm about to explain, such a denial would amount to dialectical suicide.

⁶ Actually, I'm not sure that even this is true. A little-noted fact favors internalism in the event that neither side provides a sound argument for its position. Since no one in this debate is a substance dualist, we may say that the *default* extent of the supervenience base of a person's qualia is his body. That is, internalists and externalists agree that a minimal supervenience base for qualia will *at least* include certain intrinsic states of the body. That is (again), both parties agree that changes in the physical state of a person's body can entail changes in his qualia. So the burden of proof sits a bit more heavily on externalists than on internalists. To be convinced of (some version of) externalism about qualia, we need an argument for the claim that any minimal supervenience base for qualia will include factors extrinsic to the subject's body. If no such argument is offered, then we have no reason to adopt externalism; and since qualia must (barring substance dualism) supervene *somewhere* in the physical realm, internalism is surely the reasonable default position – even if we have seen no positive argument for it. In sum, in the absence of evidence for either externalism or internalism, if one is committed to the mind's having *some* material basis (whatever one's views on whether mental properties can be *reduced* to the physical properties of that basis), then one should at least provisionally accept that that material basis is (within) the subject's body.

⁷ Lycan clearly thinks it is Block who is arguing from premises about the qualia of the abductee's memories. In a section titled 'Block's memory argument' (p. 129), he explains that it was Block who first raised the issue of the abductee's memories. This is true: in response to Lycan's 'gradual shift argument' (see my §4), which suggests that a gradual qualia inversion would go unnoticed, Block (1996, p. 44) says we have no reason to be suspicious of the abductee's long-term memories of Earth, and so no reason to think she would fail to notice that (e.g.) grass that now looks red (because of the qualia inversion) used to look green. Thus Lycan's stance appears to be that in arguing for [SAME], Block assumes that the qualia of the abductee's memories supervene internally:

Either Block's argument presumes some *narrow* notion of memory content and that notion accurately tracks a narrow kind of qualitative content, or it simply assumes that the [abductee] has some other kind of access to an unchanging qualitative content over time. But the notion of narrow memory content is controversial all by itself, and that of a narrow and unchanging qualitative content is just the point at issue. So if I have begged the question, I am not alone. (p. 132)

In the main text I argue that even if Lycan is right, it only shows that he is leaving the CC-theory undefended. But I also think the internalist does not beg the question – or at least, he certainly doesn't need to. As I said in n. 3, the internalist argument for [SAME] need not assume that qualia supervene internally; ergo, it need not assume that the qualia of the abductee's memories supervene internally. The fact that Block was the first to *mention* the abductee's memories is hardly the point.

⁸ Oddly, despite his apparent recognition that the real issue is memory *qualia*, Tye says (p. 129) that Block (1996) claims it is question-begging for the representationalist to say that memory *contents* are wide. Yet Block says no such thing: he says only that "trotting in externalist representationism about memory of phenomenal character to defend [externalist representationism about phenomenal character] seems a bit pathetic" (p. 44). Tye may be relying on an earlier draft of Block's paper, which Lycan (p. 130) quotes as saying that "trotting in an externalist representationism about memory to defend [externalist representationism] is futile."

⁹ Bernard Molyneux has suggested to me the following argument. 'Since stored memories are not conscious, and the notion of a nonconscious experience is of dubious coherence, we may suppose it's metaphysically impossible to encode experiences in memory *as experiences*. So they must be encoded propositionally. If so, then externalism about *declarative* memory would entail externalism about *visual* memory, since an image's color qualia would invert in virtue of its propositional encoding.' This is a nice try. However, I'm skeptical about the assumption that the only alternative to encoding an experience as an experience is encoding it propositionally, so I'm not ready to say the argument is sound. And as I'm in the midst of explaining in the main text, the entire externalist memory argument suffers from a bigger problem that Molyneux's argument won't solve.

¹⁰ In my view, this is the first sign that the CC-theory is in trouble. If the shift is just meant to be a matter of the abductee's visual experiences starting to covary with a different color, there should be no problem with its being instantaneous, and the resulting qualia inversion should be unnoticeable no matter *how* quickly it happens. Once CC-

theorists are reduced to having to find ways to explain why the inversion goes unnoticed (such as that it happens gradually), they have given the game away.

¹¹ For advice and discussion of this paper at various stages of its development, I thank Brian McLaughlin, J'aime Wells and especially Bernard Molyneux. They should not, however, be held responsible for any of it.

References.

- Block, N. (1990). "Inverted Earth," in J. E. Tomberlin (ed.), *Philosophical Perspectives, 4: Action Theory and Philosophy of Mind*. Atascadero, CA: Ridgeview, pp. 52-79.
- Block, N. (1994). "Qualia," in S. Guttenplan (ed.), *A Companion to the Philosophy of Mind*. Oxford: Blackwell, pp. 514-520.
- Block, N. (1996). "Mental Paint and Mental Latex," in E. Villanueva (ed.), *Philosophical Issues, 7: Perception*. Atascadero, CA: Ridgeview, pp. 19-49.
- Block, N. (1998). "Is Experiencing Just Representing?" *Philosophy and Phenomenological Research*, 58, pp. 663-670.
- Block, N. (2003). "Mental Paint," in M. Hahn and B. T. Ramberg (eds.), *Reflections and Replies: Essays on the Philosophy of Tyler Burge*. Cambridge, MA: MIT Press, pp. 165-200.
- Burge, T. (1979). "Individualism and the Mental," in P. A. French, T. E. Uehling, and H. K. Wettstein (eds.), *Midwest Studies in Philosophy IV: Studies in Metaphysics*. Minneapolis: University of Minnesota Press, pp. 73-121.
- Burge, T. (1986). "Individualism and Psychology," *The Philosophical Review*, 95, pp. 3-45.
- Burge, T. (2003). "Qualia and Intentional Content: Reply to Block," in M. Hahn and B. T. Ramberg (eds.), *Reflections and Replies: Essays on the Philosophy of Tyler Burge*. Cambridge, MA: MIT Press, pp. 405-415.
- Chalmers, D. J. (1996). *The Conscious Mind: In Search of a Fundamental Theory*. Oxford: Oxford University Press.
- Dretske, F. (1995). *Naturalizing the Mind*. Cambridge, MA: MIT Press.

- Harman, G. (1982). "Conceptual Role Semantics," *Notre Dame Journal of Formal Logic*, 23, pp. 242-256.
- Harman, G. (1988). "Wide Functionalism," in S. Schiffer and S. Steele (eds.), *Cognition and Representation*. Boulder: Westview, pp. 11-20.
- Harman, G. (1990). "The Intrinsic Quality of Experience," in J. Tomberlin (ed.), *Philosophical Perspectives, 4: Action Theory and Philosophy of Mind*. Atascadero: Ridgeview, pp. 31-52.
- Heal, J. (1998). "Externalism and Memory," *Proceedings of the Aristotelian Society (Supplementary Volume)*, 72, pp. 95-109.
- Lycan, W. G. (1996). *Consciousness and Experience*. Cambridge, MA: MIT Press.
- Macpherson, F. (2005). "Colour Inversion Problems for Representationalism," *Philosophy and Phenomenological Research*, 70, pp. 127-152.
- McLaughlin, B. P. (2003). "Colour, Consciousness, and Colour Consciousness," in Q. Smith and A. Jokic (eds.), *Consciousness: New Philosophical Perspectives*. Oxford: Clarendon Press, pp. 97-154.
- Putnam, H. (1975). "The Meaning of 'Meaning'," in K. Gunderson (ed.), *Minnesota Studies in the Philosophy of Science VII: Language, Mind, and Knowledge*. Minneapolis: University of Minnesota Press, pp. 131-193.
- Stalnaker, R. (1996). "On a Defense of the Hegemony of Representation," in E. Villanueva (ed.), *Philosophical Issues, 7: Perception*. Atascadero, CA: Ridgeview, pp. 101-108.
- Tye, M. (1995). *Ten Problems of Consciousness: A Representational Theory of the Phenomenal Mind*. Cambridge, MA: MIT Press.
- Tye, M. (2000). *Consciousness, Color, and Content*. Cambridge, MA: MIT Press.
- Tye, M. (2003). "Blurry Images, Double Vision, and Other Oddities: New Problems for Representationalism?" in Q. Smith and A. Jokic (eds.), *Consciousness: New Philosophical Perspectives*. Oxford: Clarendon Press, pp. 7-32.