

INTENTIONALISM AND INTRANSITIVITY

ABSTRACT. I argue in this paper that the existence of ‘sorites series’ of color patches – series of color patches arranged so that the patches on each end look different in color though no two adjacent patches do – shows that the relation of *same phenomenal character as* is not a transitive relation. I then argue that the intransitivity of *same phenomenal character as* conflicts with certain versions of *intentionalism*, the view that an experience’s phenomenal character is exhausted, or fully determined by its intentional content. Lastly, I consider various objections to the arguments and reply to them.

1. INTRODUCTION

*Intentionalism*¹ is the view that a conscious experience’s phenomenal character – what it is like to undergo the experience – is *exhausted*, or *fully determined*, by its intentional content.² Intentionalism is popular these days and has been defended in print by a number of prominent philosophers of mind.³ Perhaps the most seductive reason for adopting an intentionalist theory of phenomenal character is that it appears to make consciousness safe for materialist metaphysics. If intentionalism is true, the problems of consciousness seem to reduce to one: the problem of intentionality. And once intentionality has been “naturalized”, as optimistic materialists assure us it will be, there will be nothing to fear from those once recalcitrant qualia.⁴ Unfortunately, as I will argue below, the main versions of intentionalism are false.

The anti-intentionalist arguments I will present have the virtue of applying to nearly every version of intentionalism. Intentionalists divide on whether the determinants of an experience’s intentional content are wholly internal. The vast majority of intentionalists deny that they are; most intentionalists are *externalist* intentionalists. Among externalist intentionalists there are further divisions depending on what sort of externalism is espoused. ‘Causal covariational’ externalist intentionalists claim that an experience represents the objects and properties that are its dominant causal source. ‘Teleological’ externalist intentionalists say that an experience represents whatever it is the (natural) function of that experience-type to

represent. My arguments apply to both kinds of externalist intentionalism and to internalist intentionalism to boot.⁵

‘Unrestricted’ intentionalists claim that intentionalism is true of experiences generally. ‘Restricted’ intentionalists reserve their intentionalism for perceptual experiences, allowing that the phenomenal characters of bodily sensations (for example) are determined otherwise than by their intentional contents. ‘Intramodal’ intentionalists maintain that intentionalism holds, but only *within* a given perceptual modality, so that, for example, phenomenal differences between visual experiences are to be accounted for intentionally, while phenomenal differences between visual and auditory experiences are to be accounted for non-intentionally. According to ‘intermodal’ intentionalists, *all* phenomenal differences can be accounted for intentionally.⁶ Again, these various divisions do not affect the scope of my arguments.

There is yet another dimension along which intentionalist theories of phenomenal character differ, and the differences along this dimension do matter to the scope of my arguments. In Section 2, I explain these differences and how they matter. In Section 3, I present my arguments, and in Section 4, I reply to potential objections.

2. GRADES OF INTENTIONALISM AND THE LIMITS OF THE COMING ARGUMENTS

I have formulated intentionalism as the view that an experience’s phenomenal character is exhausted, or fully determined, by its intentional content. This loose formulation is intended to capture several different *grades* of intentionalism. The strongest of these, call it ‘Strong Intentionalism’, is the view that an experience’s phenomenal character is *identical* to its intentional content. Strong Intentionalism is defended by Michael Tye, who claims that “[p]henomenal character (or what it is like) is one and the same as a certain sort of intentional content” (Tye 1995, 137), and by Fred Dretske, who claims that “[a]ll mental facts are representational facts” (Dretske 1995, xiii). A weaker grade of intentionalism, call it ‘Medium Intentionalism’, is the conjunction of the following two theses:

- (A) It is impossible for two experiences to share a phenomenal character while differing in intentional content.
- (B) It is impossible for two experiences to differ in phenomenal character while sharing an intentional content.

Strong Intentionalism implies Medium Intentionalism. And Medium Intentionalism implies a still weaker grade of intentionalism, call it ‘Weak Intentionalism’, which is simply the second conjunct of Medium Intentionalism, i.e., (B) from above – the thesis that it is impossible for two experiences to differ in phenomenal character while sharing an intentional content.

Weak Intentionalism is the minimal intentionalist view – the view, as it is sometimes put, that phenomenal character “supervenes” on intentional content. Some intentionalists who do not explicitly commit themselves to Strong Intentionalism explicitly commit themselves to only Weak Intentionalism (Harman 1990, 49; Lycan 1996, 160; Byrne 2001, 200) but some of these same intentionalists argue in ways that suggest that what they really mean to defend is something that either is or implies Medium Intentionalism (though, oddly, nobody *explicitly* commits themselves to Medium Intentionalism).⁷ Most of these intentionalists seem prepared to admit, for example, that Ned Block’s “Inverted Earth” thought experiment is *prima facie* a threat to their view (Block 1990).⁸ But Block’s thought experiment purports to generate a case in which two intentionally different experiences are phenomenally the same, and such cases are *compatible* with Weak Intentionalism.⁹ (Perhaps a single phenomenal character can be “multiply realized” by different intentional contents.)

I present two arguments in the next section. The first demonstrates that Strong Intentionalism is false. The second demonstrates that Medium Intentionalism is false.¹⁰ I do not have a conclusive argument against Weak Intentionalism. I would like to have one, but for two reasons I feel okay about not having one. Firstly, as I mentioned above, many intentionalists who are not quite Strong Intentionalists are something like Medium Intentionalists even though they do not explicitly say so. In other words, I think my arguments target the intentionalisms actually held by most professed intentionalists. Secondly, although I do not have a conclusive argument against Weak Intentionalism, I think that the considerations raised in the arguments below strongly suggest its falsity.

3. THE ARGUMENTS

Suppose we divide the spectrum from red to yellow into a series of adjacent patches, and that we divide it finely enough so that, for a normal human subject, each patch looks precisely the same in color as each patch adjacent to it. By so doing, we make it the case that a normal human subject’s visual, color experience of any particular patch has the same phenomenal character as that subject’s visual, color experience of any patch adjacent

to it.¹¹ However, the phenomenal character of a subject's experience of the first patch in the series is undeniably different from the phenomenal character of that subject's experience of the last patch in the series. The first patch, which is red, produces an experience with a "red-feeling" character, while the last patch, which is yellow, produces an experience with a "yellow-feeling" character.¹² It follows from what has been said so far that:

- (P1) The relation of *same phenomenal character as* is not transitive.¹³

If the relation of *same phenomenal character as* were transitive, we would be able to derive the conclusion that the phenomenal character of a subject's experience of the first patch is the same as the character of her experience of the last patch. Since the experiences of the patches in every pair of adjacent patches have the same phenomenal character, the experience of the first shares a character with the experience of the second. Likewise, the experience of the second shares a character with the experience of the third. If *same phenomenal character as* were transitive, it would follow that the experiences of the first and third patches share a phenomenal character. But since the experience of the fourth patch shares a character with the experience of the third, it would follow by transitivity that the experiences of the first and fourth patches share a character, and hence, repeating the reasoning, that the experiences of the first and fifth share a character, and so on, up until the last patch in the series. But the experiences of the first and last patches do *not* share a character, so (P1) is true. As you may have guessed, (P1) is the first premise of my first anti-intentionalist argument.

The second premise is that:

- (P2) The relation of *same intentional content as* is transitive.

Intuitively, (P2) is true. If my belief has the same intentional content as George's desire, and George's desire has the same content as Ralph's intention, then my belief has the same content as Ralph's intention. It is hard to fathom even an *apparent* counter-instance to the transitivity of *same intentional content as*. It seems that *same intentional content as* really is (while, given the argument for (P1) just presented, *same phenomenal character as* really is *not*) an *identity* relation, and identity is transitive.

The transitivity of *same intentional content as* seems hardly in need of argument, but to be on the safe side consider the following. There are dif-

fering accounts of intentional content, but despite the differences everyone can agree that:

Two content-bearing entities have the same intentional content just in case every sentence (of a given language, L) expressing the content of the first can be substituted by any sentence (of L) expressing the content of the second in any sentential context preserving truth (except quotational contexts).

Indeed, it is just this necessary and sufficient condition on sameness of intentional content that is used as a litmus test for accounts of intentional content. Call this condition the ‘Everywhere Substitutable’ condition, or ‘ES’ for short.

Now suppose that *same intentional content as* is not transitive and hence that there are (or are possibly) three content-bearing entities, a , b , and c , such that a has the same content as b , and b has the same content as c , but a does not have the same content as c . By ES, it follows that:

- (1) Every sentence that expresses the content of a can be substituted by any sentence that expresses the content of b in any sentential context preserving truth.

ES also entails that:

- (2) Every sentence that expresses the content of b can be substituted by any sentence that expresses the content of c in any sentential context preserving truth.

From (1) and (2) it clearly follows that:

- (3) Every sentence that expresses the content of a can be substituted by any sentence that expresses the content of c in any sentential context preserving truth.

But now from (3) and ES it follows that:

- (4) a and c have the same intentional content,

which contradicts the assumption that there are (or are possibly) three content-bearing entities, a , b , and c , such that a has the same content as b , and b has the same content as c , but a does not have the same content as c . In other words, given ES, *same intentional content as* cannot be intransitive – (P2) is true.

Recall the view I called ‘Strong Intentionalism’ in the Introduction, the intentionalist view espoused by Dretske (1995, 1996) and Tye (1995, 1998, 2000). Strong Intentionalism maintains that an experience’s phenomenal character is identical to its intentional content. Presumably, it is a consequence of Strong Intentionalism that, with respect to experiences, the *same phenomenal character as* relation just *is* the *same intentional content as* relation. But the conjunction of (P1) and (P2) conflicts with this consequence of Strong Intentionalism. The two relations have different logical properties. They *cannot* be one and the same. From (P1) and (P2) it therefore follows that:

(C1) Strong Intentionalism is false.

This completes the first promised argument.

The second promised argument, the argument against Medium Intentionalism, begins with the premise that *same intentional content as* is transitive, i.e., (P2) from the preceding argument against Strong Intentionalism. Its next premise is simply the consideration that I used to support (P1), namely that:

(P3) There can be a color-patch series of the sort described in the argument for (P1).

What I mean by (P3) is that we can indeed divide the spectrum from red to yellow into a series of adjacent patches finely enough so that, for a normal human subject, each patch looks precisely the same in color as each patch adjacent to it. And that by so doing we can indeed make it the case that a normal human subject’s visual, color experience of any particular patch has the same phenomenal character as that subject’s visual, color experience of any patch adjacent to it, even though the first and last patches produce experiences for that subject with distinctly different characters. It seems utterly straightforward that we can do these things and that there can be such a series.

The next premise of the argument against Medium Intentionalism is that:

(P4) If *same intentional content as* is transitive, but there can be a color-patch series of the sort described in the argument for (P1), then it is possible for two experiences to share a phenomenal character while differing in intentional content.

Consider again a color-patch series of the sort described in the argument for (P1). Such a series, remember, is such that there are no two adjacent

patches the experiences of which differ in phenomenal character. Another fact about such a series seems to be this: A subject's experience of the first patch in the series differs in intentional content from that subject's experience of the last patch in the series.¹⁴ Intentionalists can hardly deny this. The experiences of the first and last patches differ in character, and so on any grade of intentionalism they differ in content. It is, besides, intuitively the case that the experiences of the first and last patches differ in content. The experience of the first represents the patch as *being red*, while the experience of the last represents the patch as *being yellow*.

Given that the experiences of the first and last patches differ in content, if *same intentional content as* is transitive, it follows that there must be two patches in the series that are (a) adjacent to one another and (b) such that they give rise to experiences that differ in their intentional contents. How does this follow?

Either the first and second patches in the series give rise to experiences with the same intentional content, or they do not. If the first and second patches do give rise to experiences with the same content, then it cannot be that every *other* pair of adjacent patches gives rise to experiences with the same intentional content, since, given the transitivity of *same intentional content as*, it would then follow that the first and the last patches produce experiences with the same content, which they do not. But if the first and second patches do not produce experiences with the same content, then there are adjacent patches the experiences of which differ in their intentional contents, namely the first and second patches.

So, if there can be a color-patch series of the sort described in the argument for (P1), then, if *same intentional content as* is transitive, there can be a pair of patches the experiences of which differ in content but are the same in character, since there will be at least one pair of adjacent patches in the series the experiences of which will differ in content. But then (P4) is true: If *same intentional content as* is transitive, but there can be a color-patch series of the sort described in the argument for (P1), then it is possible for two experiences to share a phenomenal character while differing in intentional content.

The next premise is simply the consequent of (P4), namely that:

- (P5) It is possible for two experiences to share a phenomenal character while differing in intentional content.

(P5) follows by *modus ponens* from (P4) and the conjunction of (P2) and (P3).

The conclusion of the argument is that:

- (C2) Medium Intentionalism is false.

Medium Intentionalism, remember, is the conjunction of the following two theses:

- (A) It is impossible for two experiences to share a phenomenal character while differing in intentional content.
- (B) It is impossible for two experiences to differ in phenomenal character while sharing an intentional content.

The falsity of Medium Intentionalism therefore follows from the truth of (P5), since (P5) contradicts (A).

4. OBJECTIONS AND REPLIES

In full, the arguments from the last section are as follows.

The Argument Against Strong Intentionalism

- (P1) The relation of *same phenomenal character as* is not transitive.
- (P2) The relation of *same intentional content as* is transitive.
- (C1) Strong Intentionalism is false.

The Argument Against Medium Intentionalism

- (P2) The relation of *same intentional content as* is transitive.
- (P3) There can be a color-patch series of the sort described in the argument for (P1).
- (P4) If *same intentional content as* is transitive, but there can be a color-patch series of the sort described in the argument for (P1), then it is possible for two experiences to share a phenomenal character while differing in intentional content.
- (P5) It is possible for two experiences to share a phenomenal character while differing in intentional content.
- (C2) Medium Intentionalism is false.

Both arguments are clearly valid. I think they are also sound, but I can imagine various objections to (P1), (P2), and (P3). In the present section I will defend (P1), (P2), and (P3) from some of these objections.¹⁵

There is a sense in which any objection to (P1) is an objection to (P3) and *vice versa*. (P3), the premise that there can be a color patch series of the sort described in the argument for (P1), *entails* (P1), the premise that *same phenomenal character as* is not transitive, so any objection to (P1) is equally an objection to (P3). But (P3) is the main ground for holding (P1), so any objection to (P3) undercuts the support for (P1). For these reasons, objections 1–4 below can be viewed as objections to both (P1) and (P3). Objection 5 below targets (P2).

Objection 1

It does not make sense to claim that *same phenomenal character as* is not transitive. There is that word in there, ‘*same*’. Its presence shows that we are dealing with an identity relation, and identity is transitive. The *same phenomenal character as* relation is the same relation as the *is identical in phenomenal character to* relation. If the conceptual confusion inherent in sincerely uttering ‘*Same phenomenal character as* is not transitive’ is not manifest, surely it is manifest in sincerely uttering ‘*Is identical in phenomenal character to* is not transitive’.

Reply to Objection 1

If the argument given for (P1) is sound, then *same phenomenal character as* is not transitive despite the fact that it might sound odd to say so. It is neither impossible nor incoherent that an intransitive relation should be referred to by a relation name with ‘*same*’ or ‘*identical*’ as a syntactic part.

Objection 2

The argument for (P1) simply assumes that if two patches look the same in color, then the experiences of those patches share a phenomenal character. Why should this be accepted? Perhaps patches that look the same can nevertheless produce experiences with different phenomenal characters.

Reply to Objection 2

It seems rather central to the notion of phenomenal character that when two things look (smell, sound, etc.) the same, those things produce phenomenally identical experiences. What it is like to perceptually experience something that is perceptually indistinguishable from something else is just what it is like to perceptually experience that something else. Our notion of phenomenal character seems essentially tied to our notions of *appearing the same as*, or *being perceptually indistinguishable from*. Things

that are perceptually indistinguishable cannot give rise to phenomenally different perceptual experiences – this ought to strike one as a conceptual truth.

The assumption that adjacent patches produce experiences with the same phenomenal character can be further supported by citing the *introspective* indistinguishability of the experiences in question. *Being introspectively indistinguishable from* is yet another relation that is intimately connected to the relation of *same phenomenal character as*, introspection being the means by which we acquire knowledge of the phenomenal characters of our experiences. Notice too that all of these relations – *appears the same as*, *is perceptually indistinguishable from*, and *is introspectively indistinguishable from* – are themselves intransitive relations as consideration of the color-patch series and the visual experiences produced by it reveals. Given the close connection between these relations and *same phenomenal character as*, this is yet more evidence for the intransitivity of *same phenomenal character as*.

In any case, the move from the fact that adjacent patches look the same to the claim that they produce phenomenally identical experiences is not strictly needed. That the experiences of adjacent patches have the same phenomenal character is given directly in experience. Any subject can tell just by looking at a series of color patches laid out in the way specified in the argument for (P1), and introspecting his or her experiences, that each pair of adjacent patches is such that the patches in that pair produce phenomenally identical experiences. There is no need to *infer* this from the holding of some other relation between the patches themselves (like perceptual indistinguishability) or the experiences produced by them (like introspective indistinguishability), even though, as I have argued, it *can* be so inferred.

Objection 3

It must be assumed that the color-patch series described in the argument for (P1) is such that the experiences of adjacent patches have the same phenomenal character for a single subject *at a single time*. That is, the series must be taken to be such that the experiences (for that single subject) of patches 1 and 2 share a character at a certain time *t*, and that the experiences of patches 2 and 3 share a character at that *same* time *t*, and that the experiences of patches 3 and 4 also share a character at *t*, and so on. Only then would it follow from the fact that the experiences of the first and last patches do not share a character at *t* that *same phenomenal character as* is not transitive. The experiences of a single patch might differ in phenomenal character at different times. And hence, even if *same*

phenomenal character as were transitive, that would not guarantee that if the experience of patch 1 has the same character as the experience of patch 2 at t and the experience of patch 2 has the same character as the experience of patch 3 at t_1 , then the experience of patch 1 has, at t , t_1 , or any time, the same phenomenal character as the experience of patch 3. So there is no threat to the transitivity of *same phenomenal character as* just so long as there cannot be a series of color patches such that the experiences of adjacent patches have the same phenomenal character *at a single time, t* , while the experiences of the first and last patches differ in character at t . But what reason do we have to suppose that there can be such a series?

Not only is the assumption that there can be such a series unjustified, but it looks as though we have good reason to reject it. For if the assumption is correct, then *same phenomenal character as* is not transitive, and so it is possible for three experiences, a , b , and c , to be such that:

- (I) a has the same character as b at t .
- (II) b has the same character as c at t .
- (III) a does not have the same character as c at t .

But it follows from (I)–(III) that b has two distinct phenomenal characters *at one and the same time*. From (I), it follows that there is a phenomenal character, C , such that a and b both have C at t , and from (II) it follows that there is a phenomenal character C^* , such that b and c both have C^* at t . And from (III), it follows that *there is no* character such that a and c both have that character at t . But by combining these consequences we get that b has both C and C^* at t , but that $C \neq C^*$, and that, surely, is impossible.¹⁶

Reply to Objection 3

It is true that in order for the intransitivity of *same phenomenal character as* to follow from the possible existence of a series of color patches of the sort described in the argument for (P1), the series must be such that when the patches are viewed together at a single time by a single subject, there are no two adjacent patches the experiences of which differ in phenomenal character even though the experiences of the first and last patches do. But series of this sort are not only possible, they are actual. I know because I have seen them. I have seen a series of color patches arranged so that while the patches on each end looked different in color at a certain time, t , no two adjacent patches looked different in color at t . And I have seen a series of color patches arranged so that while the patches on each end produced (color) experiences with different phenomenal characters at a certain time, t , no two adjacent patches did so at t . (Each series of the former kind is a series of the latter kind, so I could not help but see a series of the latter kind

when I saw a series of the former kind.) I am a “normal human subject” (I hope) and so I have no reason to suspect that the series I saw would strike any other normal human subject any differently.

Someone seriously pressing *Objection 3* would likely say that this “evidence” for the existence of the relevant sort of series begs the question, and that all I am entitled to claim is that certain series of color patches would elicit a *judgement* from normal human subjects to the effect that the experiences of no two adjacent patches differ in phenomenal character at some single time, *t*, while the experiences of the first and last patches do. What is needed, the objector might say, is an argument for the truth of such judgments. (The demand for an argument for the truth of such judgments is somewhat strange, however. What better evidence is there for the truth of judgments about *how things look* or *what certain experiences are like* for a given subject than the fact that the subject sincerely makes those judgments? But I will not press this issue since the demand for an argument can be met.)

Presumably, our powers of discrimination are – in general, and with respect to differences in color in particular – *finite*. For example, there is some amount of color difference, *n*, such that if two color patches differ from one another by *n* or less, they will look the same in color. But if so, then constructing a series of the sort described in the argument for (P1) is straightforward. One starts with a red patch, and then one selects as the second patch a patch that is more yellow than the red one, but only by *n* or less. Next, one selects as the third patch a patch that is more yellow than the second, but again only by *n* or less, and so on – repeating the process until one ends up selecting a yellow patch.¹⁷ A series of color patches constructed in this way cannot help but be such that no two adjacent patches look different in color, regardless of whether the patches are viewed two by two or all at once at a single time (and it cannot be seriously doubted that the patches *can* be viewed all at once). Any series of color patches such that there are no two adjacent patches that look different in color *is* a series of color patches such that there are no two adjacent patches the experiences of which differ in phenomenal character.¹⁸ So, the argument that there can be a series of the latter sort is just this: *there are coherent instructions for constructing one*, namely the instructions just presented.

Given that there can be series of the relevant sort, and given that this implies that *same phenomenal character as* is not transitive, what are we to make of the (I)–(III) argument above that if *same phenomenal character as* is not transitive, then there can be experiences with two distinct phenomenal characters at a single time? One option is to accept its conclusion. What, exactly, is the reason for thinking that no single experience can

have two different phenomenal characters at a single time? The argument assumes that such a thing will strike us as an obvious impossibility. But it is not conceptually impossible, or not clearly so at least.¹⁹ Another option is to deny that the argument can be made valid. As it stands, the argument is not valid. For example, from (I) it does not follow that there is some character *C*, such that *a* has *C* and *b* has *C*. That follows from (I) only given the assumption that when *x* bears the *same phenomenal character as* relation to *y*, then *there is* a phenomenal character such that *x* has that character and *y* has that character. Such an assumption seems harmless enough, but given that it leads, in conjunction with other equally harmless seeming assumptions, to the conclusion that a single experience can have two distinct phenomenal characters at a single time, perhaps we should reject it. In any case, given these various options against the (I)–(III) argument, and given the plausibility of the ‘coherent instructions’ argument for possibility of a series of color patches that displays the intransitivity of *same phenomenal character as*, the (I)–(III) argument can hardly be taken to demonstrate that *same phenomenal character as* must be transitive.

Objection 4

Graff (2001) maintains that *looks the same as* (in a certain respect) is transitive and she offers an argument that purports to show that it must be.²⁰ Her argument can be adapted in support of the conclusion that *same phenomenal character as* must be transitive as well, and thus that (P1) is false.

Graff argues that if *looks the same as* is not transitive, we are forced to reject what she calls “same-appearance claims” like, ‘If two color patches look the same in color, then if one looks red, so does the other’. But Graff thinks that same-appearance claims are truisms and thus that to reject them is not only to make a mistake but to betray a conceptual confusion. How does the intransitivity of *looks the same as* force us to reject same-appearance claims?

A same-appearance claim serves as the inductive premise in instances of the so-called ‘phenomenal sorites’. If we assume that there can be a series of color patches such that the first looks red, the last looks yellow, and every patch looks the same in color as its immediate neighbors, then the following argument is valid:

- (i) The first patch looks red.
- (ii) If two color patches look the same in color, then if one looks red, so does the other.
- (iii) Therefore, the last patch looks red.

But (iii) is false – the last patch looks yellow, not red. According to Graff, we cannot reject (ii). It is a truism. So the only thing left to reject is our initial supposition that there can be a series of color patches chained together by *looks the same as* with patches that look different from each other on each end. In other words, we must give up the assumption that *looks the same as* is not transitive.²¹

The assumption that *same phenomenal character as* is not transitive leads to a closely related paradox that I will call the ‘phenomenal-character sorites’. If we assume that there can be a series of color patches such that the first patch produces an experience with a “red-feeling” character, the last patch produces an experience with a “yellow-feeling” character, and every patch produces an experience with the same phenomenal character as the experiences produced by its immediate neighbors, then the following argument is valid:

- (i*) The first patch produces an experience with a red-feeling character.
- (ii*) If two color patches produce experiences with the same phenomenal character, then if one produces an experience with a red-feeling character, then so does the other.
- (iii*) Therefore, the last patch produces an experience with a red-feeling character.

Here again the conclusion is false. But it seems that (ii*) is just as much a (seeming) truism as (ii). Should we then insist, as Graff does with respect to (ii) and the case of the phenomenal sorites, that (ii*) *cannot* be rejected and thus that the only thing left to reject is the supposition that there can be a series of color-patches of the relevant sort (i.e., one that demonstrates the intransitivity of *same phenomenal character as*)?

Reply to Objection 4

Insisting on such a thing is a bit like insisting that there cannot be a series of men, the first not bald, the last bald, each one differing from the next by the subtraction of a single hair, because the assumption that there can be such a series of men, in conjunction with the truism that a difference of a single hair cannot make the difference between a man who is not bald and one who is, leads to the conclusion that the last man in the series is not bald. But it is absurd to insist that there cannot be such a series of men.

Graff would say that the difference between the phenomenal sorites and (I would guess) the phenomenal-character sorites on the one hand, and the sorites for *bald* on the other, is that the first two paradoxes have truisms as their inductive premises. But as I already suggested, that a difference of a

single hair cannot make the difference between man who is not bald and one who is, seems to be a truism as well, and that claim is the inductive premise of the sorites for *bald*.²²

One striking fact about the series appealed to in the phenomenal and phenomenal-character sorites is that they not only can exist but *do* exist. As I said in my *Reply to Objection 3*, I have seen them. Graff describes claims like this on the part of those who maintain that *looks the same as* is not transitive as “tendentious” (Graff 2001, 916). But her only reason for describing them as such is that their truth is sorites-paradoxical, and that is not reason enough.²³ We have two options when it comes to the phenomenal and phenomenal-character sorites: Deny that there can be series of the relevant kinds or concede that the “truisms” that serve as the inductive premises in the respective paradoxes are not truisms at all. Neither option is particularly attractive, but given that I have *seen* series of the relevant kinds, I think that we must opt for the latter. In any case, regardless of how the solutions to the phenomenal and phenomenal-character sorites ought to go, there is no decisive argument from the existence of the paradoxes to the conclusion that *looks the same as* or *same phenomenal character as* must be transitive.

Objection 5

If there can be a color-patch series of the sort described in the argument for (P1), then *same intentional content as* is *not* transitive – (P2) is false. In such a series, each patch in each pair of adjacent patches is admitted to look precisely the same (in color) as the other patch in that pair. But if so, then the (visual, color) experiences produced by adjacent patches must share an intentional content. And hence, if *same intentional content as* is transitive, we can derive the conclusion that the experiences of the first and last patches share an intentional content, which they do not. So, if there can be such a series, *same intentional content as* is not transitive.

Reply to Objection 5

In form, Objection 5 mimics the argument I gave in support of (P1). And besides the unobjectionable claim that the experiences of the first and last patches differ in their intentional contents, it depends only on the claim that experiences of patches that look the same must share an intentional content. Happily, this latter claim is false. One argument for its falsity has

in effect already been given.²⁴ In my argument for (P2), I argued that (P2) is true if and only if ES, the ‘Everywhere Substitutable’ condition, is true:

Two content-bearing entities have the same intentional content just in case every sentence (of a given language, L) expressing the content of the first can be substituted by any sentence (of L) expressing the content of the second in any sentential context preserving truth (except quotational contexts).

But ES is not merely true, it is *a priori*, necessarily true. The argument against (P2) we are currently considering is therefore unsound, and since its only suspicious claim is that patches that look the same must produce experiences with the same intentional content, that claim must be false.

There are good grounds for denying that experiences of patches that look the same must share an intentional content. Are there any grounds at all for asserting it? In my *Reply to Objection 2*, I claimed there to be a conceptual connection between visual indistinguishability and sameness of visual phenomenal character. If *a* and *b* are visually indistinguishable for *S*, it is incoherent to suppose that what it is like for *S* to visually experience *a* is different from what it is like for *S* to visually experience *b*. It seems plain that there is not a similar conceptual connection between visual indistinguishability and sameness of intentional content for visual experiences. In fact, it is widely held, especially among intentionalists,²⁵ that our perceptual experiences have *non-conceptual* intentional contents, and that these contents are often far richer than we can cognitively discern. One consequence of such a view is that two experiences might differ in their contents without this fact being known, or even knowable, to the subject of those experiences. So not only is the claim that patches which look the same must give rise to experiences with a shared content not supported by the going account of content for experiences, but its negation is entirely consistent with that account.

5. CONCLUSION

In the last section, I showed that some of the main objections to the anti-intentionalist arguments of Section 3 fail. There are undoubtedly objections that I have not considered, but I think it is a safe bet that the arguments are sound and that Strong Intentionalism and Medium Intentionalism are therefore false.

I have not argued against Weak Intentionalism, but I think that the arguments given against Strong Intentionalism and Medium Intentionalism

suggest that it too is false. Those arguments show that an experience's phenomenal character is not identical to its intentional content, and that two experiences might differ in their intentional contents while not differing in their phenomenal characters. Generalizing a bit, what the arguments show is that content and character are largely *independent* of one another. Given the degree of independence they have been shown to have here, it would be surprising, to say the least, if Weak Intentionalism were true – if, that is, there could not be differences in phenomenal character without corresponding differences in intentional content. When this consideration is added to the considerations that have already been adduced in the literature against Weak Intentionalism, I think we have more reasons to reject Weak Intentionalism than we have to accept it.²⁶

It is widely recognized that, from a materialist perspective, phenomenal character and consciousness in general are deeply puzzling.²⁷ I hope to have shown here that even if the puzzle would be less of one if intentionalism were true, it is in fact as much of a puzzle as ever.²⁸

NOTES

¹ Intentionalism goes by many other names including 'representationism' in Block (forthcoming), 'representationalism' in Block (1996) and Tye (2000), 'the hegemony of representation' in Lycan (1996), and 'the Representational Thesis' in Dretske (1995). 'Intentionalism' is its newest name however, appearing as such in Byrne (2001).

² This formulation of intentionalism in terms of "exhaustion" or "full determination" is common in the literature (Block, forthcoming) but too loose to be very useful. I say more precisely what it means to be an intentionalist in Section 2.

³ These include Byrne (2001), Dretske (1995, 1996), Harman (1990, 1996), Lycan (1996, 1998, forthcoming), and Tye (1995, 1998, 2000).

⁴ Dretske (1995) and Lycan (1996) are especially attracted by the promise intentionalism seemingly holds for the prospects of "naturalizing the mind" and solving the problems of consciousness. Lycan, for example, writes:

I am concerned to maintain a weak version of Brentano's doctrine that the mental and the intentional are one and the same. It would follow that once representation itself is (eventually) understood, then not only consciousness in our present sense but subjectivity, qualia, 'what it's like', and every other aspect of the mental will be explicable in terms of representation with the underlying functionally organized neurophysiology, without our positing any other ingredient not already well understood from the naturalistic point of view. I do not think there will be any 'problem of consciousness' left. (Lycan 1996, 11)

For a healthy dose of skepticism about the benefits of intentionalism for materialism see the excellent (and short) Warfield (1999).

⁵ Externalist intentionalists include Dretske (1995, 1996), Lycan (1996, 1998, forthcoming), Tye (1995, 1998, 2000), and as far as I can tell, Harman (1990, 1996). Dretske and Lycan hold teleological versions of externalist intentionalism and Tye, adding a few bells

and whistles (Tye 2000, chap. 6), holds a causal covariational version of externalist intentionalism. It is unclear which externalist version, if either, Harman endorses. Internalist intentionalists are rare, perhaps at this stage merely possible, although Rey (1998) comes close.

⁶ Byrne, Dretske, and Tye are unrestricted, intermodal intentionalists. Lycan is an unrestricted intramodal intentionalist. There are perhaps no longer any restricted intentionalists, though Tye used to be one (Tye 1992). I have borrowed the ‘restricted/unrestricted’, ‘intramodal/intermodal’ terminology from Byrne (2001, 201–202).

⁷ Something like Medium Intentionalism seems to be implied by an intentionalist theory which holds that phenomenal character is *fully explainable in terms of*, but not identical to, intentional content. For if phenomenal character is fully explainable in terms of intentional content, then one would expect phenomenal *sameness* to be fully explainable in terms of intentional sameness, and it would then be impossible for two intentionally different experiences to be phenomenally the same, just as Medium Intentionalism asserts. I suspect that most intentionalists believe that character is fully explainable in terms of content. This is one reason I claim that most intentionalists are committed to something like Medium Intentionalism. Another reason follows in the main text.

⁸ This is true of Byrne who references Block’s Inverted Earth papers and describes them as containing objections to intentionalism in Byrne (2001, 218). And it is true also of Lycan who makes an extensive reply to Block’s alleged counterexample in Lycan (1996, 113–118).

⁹ My thanks to Norva Lo for drawing my attention to this.

¹⁰ Why not skip immediately to the argument against Medium Intentionalism given that Strong Intentionalism implies Medium Intentionalism? Because the falsity of Strong Intentionalism follows directly from a pair of fairly intuitive premises. The argument against Medium Intentionalism takes a bit more setting up.

¹¹ One is meant to suppose here and throughout the rest of the paper that our normal human subject is looking, under normal conditions, at the entire series of color patches at once. The relevant experiences of such a subject are the *visual, color* experiences of the color patches, and the relevant phenomenal characters are the *visual* phenomenal characters of those *color* experiences. For ease of exposition, I will usually drop explicit references to a subject and the kinds of experiences in question, but a description like ‘the experience of the last patch’, for example, is meant to be short for ‘the visual, color experience had by a normal human subject when looking under normal conditions at the last patch in the series’. When I say that two patches look the same (for a normal subject, under normal conditions) it is to be understood that I mean they look the same in color.

¹² “Has a red (yellow, etc.)-feeling character” is an abbreviation for “has the character of the type of experience normally produced by red (yellow, etc.) objects under normal conditions”. I have adopted this convention from Byrne and Hilbert (1997).

¹³ Goodman (1951) can be read as suggesting that since *looks the same as* is not transitive, two objects looking the same as one another cannot suffice for the visual experiences of the two objects to be the same in phenomenal character. Goodman proposes, in effect, that the visual experiences of two objects have the same phenomenal character just in case the two objects look the same as each other *and look the same as all the same third parties as well*. My view is that two objects looking the same as one another *does* suffice for the visual experiences of the two objects to be the same in phenomenal character, but since *looks the same as* is not transitive, neither is “identity” of phenomenal character. I think my view is partly justified by the fact that on Goodman’s proposal one can never

tell just by introspecting two experiences that they share a phenomenal character. I deal more thoroughly with the idea that *looking the same as* does not suffice for producing experiences with the same visual phenomenal character in Section 3. (See especially my *Reply to Objection 2* in Section 3.)

¹⁴ The (visual, color) experiences of the first and last patches differ in a trivial intentional way – the experience of the first represents the *first* patch as having a certain color, while the experience of the *last* patch represents the last patch as having a certain color. What I mean to deny by denying that they have the same intentional content is that they represent the patches as having the same (maximally specific) color.

¹⁵ The argument for (P4) is straightforward, and (P5) is a simple consequence of (P3) and (P4), so I will not be considering objections to (P4) or (P5).

¹⁶ This objection is adapted from an objection raised in Jackson and Pinkerton (1973) against Armstrong's (1968) argument against "sense-data" theories of perception. Armstrong's argument relies on the premise that *looks the same (in color) as* is not transitive, and Jackson and Pinkerton object to this premise in a way analogous to the way in which *Objection 3* objects to (P1). I was helped in constructing the adaptation by Graff's discussion of Jackson and Pinkerton's objection (Graff 2001, 913–916.)

¹⁷ This is meant to be an elaboration of the idea that the spectrum from red to yellow can be divided into patches *finely enough* so that no two adjacent patches look the same in color.

¹⁸ I argued for this in my *Reply to Objection 2* above.

¹⁹ A related claim that does seem conceptually impossible is that a single experience should have two different and *incompatible* phenomenal characters at a single time, *t*; that, for example, a single visual experience could be both red-feeling and yellow-feeling at once. But this claim is *not* a consequence of the (I)–(III) argument.

²⁰ She is very much in the minority however. All of Armstrong (1968), Clark (1993), Dummett (1975), Everett (1996), Goodman (1951), Parikh (1983), Tappenden (1993), Travis (1985), and Wright (1975) say or imply that looks the same as is not transitive. Indeed, most of these philosophers take it to be obvious that it is not.

²¹ The argument occurs in Graff (2001, 905–909).

²² Graff explains that she views an ordinary sorites, like the sorites for *tall*, differently than she views the phenomenal sorites because:

If two men differ in height even by one-hundredth of an inch, then they differ in a respect that is relevant for the applicability of 'tall'. But if two color patches look the same (not just similar, but the *same*) in respect of color, then they do not differ, on the face of it at least, in any respect relevant for the applicability of 'looks red' (Graff 2001, 908).

But, *on the face of it*, two men who differ by one-hundredth of an inch do not differ in a respect relevant for the applicability of 'tall'. Since, on the face of it, a difference that slight cannot *make* a difference. On the other hand, even if two color patches that look the same do not differ, on the face of it, in any respect relevant for the applicability of 'looks red', it seems that we know despite this that they can so differ, since we know that indiscriminable differences between color patches might "add up" to a difference in (color) look.

²³ This is slightly unfair. Graff also describes such claims as tendentious because she thinks that there are no sound arguments for their truth (Graff 2001, 916–931.) As I said in my *Reply to Objection 3*, the demand for an *argument* for the claim that there are such series, or that such series have been seen strikes me as a bit odd. (I feel a little like Moore facing the skeptic: How do I *know* that there are such series? Well, *here* is one ... see?) But if

pressed I would argue, as I did in my *Reply to Objection 3*, that there can be (and are) such series because our powers of discrimination are finite: There are some changes so slight that they cannot be perceived. To take the example from my *Reply to Objection 3*, there is some amount of change in color such that if two color patches differ by that amount or less, then they look the same in color. Graff discusses this interpretation of what it means to say that our powers of discrimination with respect to the colors of color patches are finite, and she admits that if those powers were finite in this sense, then *looks the same (in color) as* would indeed be intransitive, but she denies that our powers of discrimination are finite in this sense. If they were, then, Graff says, we would be immune to “certain very mild hallucinations” (Graff 2001, 932): Every time some pair of color patches differed in color by *no* amount, they would look that way to us; we would see them as they are. But since it sometimes happens that color samenesses *look* like color differences to us, our powers of discrimination are not finite in the proposed sense, and there is no danger to the transitivity of *looks the same as*.

But surely the claim that our powers of discrimination are finite with respect to the colors of color patches is only true *ceteris paribus*. With its *ceteris paribus* clause attached, the claim is something roughly like: *If* the perceiver is normal and gets a long enough look, the patches are not too far away, the light is good, etc., *then*, if the patches differ in color by less than the threshold amount, they will look the same in color. This way of expressing the limits on our discriminatory powers allows for *both* the possibility of the mild hallucinations Graff rightly says we must admit *and* the existence of the ‘sorites series’ of color patches. So Graff’s complaints about the “powers of discrimination” argument for the existence of a sorites series of color patches – an argument she admits is valid – can be answered (Graff 2001, 916–920.)

²⁴ Actually, two arguments for its falsity have been given: the one about to be presented in the main text, and the argument I gave for (P4), which shows that there must be at least two patches that look precisely the same in color which produce visual, color experiences that differ in their intentional contents. However, the latter argument assumes that *same intentional content as* is transitive and so is of no use in the present context.

²⁵ For example, Dretske (1995) and Tye (2000).

²⁶ See Block (1996 and forthcoming), Levine (1997), Peacocke (1983), Robinson (1998), and Wager (1999) for a variety of criticisms of Weak Intentionalism.

²⁷ Phenomenal character is deeply puzzling independently of the question of materialism. It is puzzling, in part, precisely because the *same phenomenal character as* relation and the clearly related *looks (smells, sounds, etc.) the same as* relation are not transitive relations. I have argued here that these relations really are intransitive, but I do not deny that their being so is puzzling. One reason it is puzzling is that some philosophers have argued fairly powerfully from the intransitivity of these relations to the conclusion that there are no phenomenal properties, phenomenal characters, or qualia (Armstrong 1968; Dummett 1975; Everett 1996.) Of course, if their arguments are sound, then the fact that intentionalist theories of phenomenal character fall afoul of the intransitivity of *same phenomenal character as* is not so important, since such theories would be mistaken for a far more radical reason. I do not have the space to discuss these arguments here, but my general view of them is that they *must* contain some flaw since the very existence of phenomenal character is just not something we are in a position to doubt. There is no *argument* that could convince me that there is not “something it is like” to be in pain, for example.

²⁸ I would like to thank Harry Deutsch, Laurence Goldstein, Joe Lau, Ron Mallon, and an anonymous referee for helpful comments and criticism.

REFERENCES

- Armstrong, D.: 1968, *A Materialist Theory of the Mind*, Routledge and Kegan Paul, London.
- Block, N.: 1990, 'Inverted Earth', in James E. Tomberlin (ed.), *Philosophical Perspectives*, vol. 4, Ridgeview, Atascadero, CA, pp. 53–79.
- Block, N.: 1995, 'On a Confusion About a Function of Consciousness', *Behavioral and Brain Sciences* **18**, 225–247.
- Block, N.: 1996, 'Mental Paint and Mental Latex', in Eriqve Villanueva (ed.), *Philosophical Issues*, vol. 7, Ridgeview, Atascadero, CA, pp. 19–49.
- Block, N.: Forthcoming, 'Mental Paint', in *Others on Burge: 10 Essays With Responses from Tyler Burge*, Cambridge, MA: MIT. (Available online at www.nyu.gsas/dept/philo/faculty/block/papers/mentalphaint.pdf.)
- Byrne, A.: 2001, 'Intentionalism Defended', *Philosophical Review* **110**, 199–240.
- Byrne, A. and D. Hilbert: 1997, 'Colors and Reflectances', in Alex Byrne and David Hilbert (eds), *Readings on Color: Volume 1: The Philosophy of Color*, MIT Press, Cambridge, MA.
- Clark, A.: 1993, *Sensory Qualities*, Oxford University Press, Oxford.
- Dretske, F.: 1995, *Naturalizing the Mind*, MIT Press, Cambridge, MA.
- Dretske, F.: 1996, 'Phenomenal Externalism', in Eriqve Villanueva (ed.), *Philosophical Issues*, vol. 7, Ridgeview, Atascadero, CA.
- Dummett, M.: 1975, 'Wang's Paradox', *Synthese* **30**, 301–324.
- Everett, A.: 1996, 'Qualia and Vagueness', *Synthese* **106**, 205–226.
- Goodman, N.: 1951, *The Structure of Appearance*, First edition, Harvard University Press, Cambridge, MA.
- Graff, D.: 2001, 'Phenomenal Continua and the Sorites', *Mind* **110**, 905–935.
- Harman, G.: 1990, 'The Intrinsic Quality of Experience', *Philosophical Perspectives* **4**, 31–52.
- Harman, G.: 1996, 'Qualia and Color Concepts', in Eriqve Villanueva (ed.), *Philosophical Issues*, vol. 7, Ridgeview, Atascadero, CA, pp. 7579.
- Jackson, F. and R. J. Pinkerton: 1973, 'On an Argument Against Sensory Items', *Mind* **82**, 269–272.
- Lalor, B.: 1999, 'Intentionality and Qualia', *Synthese* **121**, 249–289.
- Langsam, H.: 2000, 'Experiences, Thoughts, and Qualia', *Philosophical Studies* **99**, 269–295.
- Levine, J.: 1997, 'Are Qualia Just Representations? A Critical Notice of Michael Tye's Ten Problems of Consciousness', *Mind and Language* **12**, 101–113.
- Levine, J.: 2001, *Purple Haze*, Oxford University Press, Oxford.
- Lycan, W.: 1996, *Consciousness and Experience*, MIT Press, Cambridge, MA.
- Lycan, W.: 1998, 'In Defense of the Representational Theory of Qualia', in James E. Tomberlin (ed.), *Philosophical Perspectives*, vol. 12, Ridgeview, Atascadero, CA.
- Lycan, W.: 2000, 'Representational Theories of Consciousness', in Edward N. Zalta (ed.), *The Stanford Encyclopedia of Philosophy*, URL=<http://plato.stanford.edu/archives/fall1999/entries/consciousness-representational/>.
- Lycan, W.: Forthcoming, 'The Case for Phenomenal Externalism', in *Philosophical Perspectives*. (Available online at www.unc.edu/~ujanel/CaPhExt.htm)
- McGinn, C.: 1988, 'Consciousness and Content', *Proceedings of the British Academy* **74**, 219–239. Reprinted in McGinn, *The Problem of Consciousness*, Blackwell, Oxford, 1991).

- McGinn, C.: 1989, *Mental Content*, Basil Blackwell, Oxford.
- Parikh, R.: 1983, 'The Problem of Vague Predicates', in *Boston Studies in the Philosophy of Science*, vol. 31, D. Reidel, Dordrecht, pp. 241–261.
- Peacocke, C.: 1983, *Sense and Content*, Clarendon Press, Oxford.
- Rey, G.: 1998, 'A Narrow Representationalist Account of Qualitative Character', in James E. Tomberlin (ed.), *Philosophical Perspectives*, vol. 12, Ridgeview, Atascadero, CA, pp. 435–457.
- Robinson, W. S.: 1998, 'Intrinsic Qualities of Experience: Surviving Harman's Critique', *Erkenntnis* **47**, 285–309.
- Shoemaker, S.: 1982, 'The Inverted Spectrum', *Journal of Philosophy* **79**, 357–381.
- Shoemaker, S.: 1994, 'Phenomenal Character', *Nous* **28**, 21–38.
- Stalnaker, R.: 1995, 'On a Defense of the Hegemony of Representation', in James E. Tomberlin (ed.), *Philosophical Issues*, vol. 7, Ridgeview, Atascadero, CA, pp. 101–108.
- Tappenden, J.: 1993, 'The Liar and Sorites Paradoxes: Toward a Unified Treatment', *Journal of Philosophy* **90**, 551–577.
- Travis, C.: 1985, 'Vagueness, Observation, and Sorites', *Mind* **94**, 345–366.
- Tye, M.: 1992, 'Visual Qualia and Visual Content', in Tim Crane (ed.), *The Contents of Experience*, Cambridge University Press, Cambridge, pp. 158–176.
- Tye, M.: 1994, 'Qualia, Content, and the Inverted Spectrum', *Nous* **28**, 159–183.
- Tye, M.: 1995, *Ten Problems of Consciousness*, MIT, Cambridge, MA.
- Tye, M.: 1998, 'Inverted Earth, Swampman, and Representationism', in James E. Tomberlin (ed.), *Philosophical Perspectives*, vol. 12, Ridgeview, Atascadero, CA, pp. 459–477.
- Tye, M.: 2000, *Consciousness, Color, and Content*, MIT Press, Cambridge, MA.
- Van Gulick, R.: 1995, 'How Should We Understand the Relation Between Intentionality and Phenomenal Consciousness?', in James E. Tomberlin (ed.), *Philosophical Perspectives*, vol. 9, Ridgeview, Atascadero, CA, pp. 271–289.
- Wager, A.: 1999, 'The Extra Qualia Problem: Synaesthesia and Representationism', *Philosophical Psychology* **12**, 263–281.
- Warfield, T.: 1999, 'Against Representational Theories of Consciousness', *Journal of Consciousness Studies* **6**, 66–69.
- Wright, C.: 1975, 'On the Coherence of Vague Predicates', *Synthese* **30**, 325–365.

The University of Hong Kong
 Department of Philosophy
 Room 301 Main Building
 Pokfulam Road
 Hong Kong, China
 E-mail: maxdeutsch@hkusua.hku.hk