Perceptual Content, Phenomenal Contrasts and Externalism

ABSTRACT: According to Sparse views of perceptual content, the phenomenal character of perceptual experience is exhausted by the experiential presentation of ‘low-level’ properties such as (in the case of vision) shapes and colours and textures. Whereas, according to Rich views of perceptual content, the phenomenal character of perceptual experience can also sometimes involve the experiencing of ‘high-level’ properties such as natural kinds, artefactual kinds, causal relations, linguistic meanings, moral properties. An important dialectical tool, which has frequently been employed in the debate between Rich vs. Sparse theorists, is the so-called ‘method of phenomenal contrast’. I explore how this method of phenomenal contrast interacts with the sort of content-externalism made familiar by Putnam. I show that the possibility of Twin Earth style cases places important restrictions on the range of properties that the method of phenomenal contrast could plausibly apply to. Moreover, these restrictions would apply to some paradigmatically low-level properties as well as to some of the frequently advanced high-level properties. I also draw some general lessons about the different ways one might conceive of the relation between phenomenal character and representational content.

1. Introduction: the Rich vs. Sparse debate

There is by now a flourishing literature devoted to the debate between Rich vs. Sparse (‘Thick’ vs. ‘Thin’) accounts of our conscious perceptual experience. We can think of this debate as starting from the following deceptively simple-looking question: which properties can we consciously perceive? So in the case of vision: which properties can we see? Or in the case of audition: which properties can we hear? These questions direct us towards drawing a division between the genuinely perceptible properties and all those other properties that we can at most recognise, judge or infer to be present in our surroundings, but which we cannot just see or hear or taste etc. At first pass, it might seem that at least some properties are pretty easy to place on one or other side of this divide. For example, almost everyone will agree that we can visually experience shapes and colours. Likewise, almost everyone will agree that we can hear the pitches and timbres and volumes of sounds. Conversely, it might seem obvious that humans cannot perceptually experience the radioactivity of a substance; that is a property that we are just not perceptually sensitive to, it is not something that we can see or hear or taste etc. – which is one reason why radiation is so dangerous! Of course an expert on radioactivity might know that certain substances are more radioactive than others and might be able to recognise these substances on sight. Still, prima facie, it might seem obvious that the radioactivity itself is not something that humans can consciously perceive. (However, we will come back to this example in a moment…) But there is a range of properties which are less easy to classify as perceptible or not. For example, when I see a tiger, as well as seeing its shape and colours do I also see its property of being an animal? When I see the tiger chase a gazelle, as well as seeing their physical movements from one location to another, do I also see the tiger’s movement as having the property of causing the gazelle to flee? And when I see the tiger’s teeth sink into the poor gazelle’s flesh, as well as seeing the physical convulsions of the gazelle’s body, do I also perceive its property of being in pain?

Sparse theorists deny that we really do consciously perceive these sorts of natural kind properties (such as being a tiger), or causal properties or emotional properties, and insist that we should restrict the range of consciously perceptible properties to so-called ‘low-level’ properties such as shape and colour, pitch and timbre, textures and tastes etc. Whereas Rich theorists propose that we can…

---

genuinely perceive some ‘high-level’ properties. As well as various natural kind properties, causal properties and emotional/mental properties, we might also add affordance or action properties (such as being climbable-by-me or being edible), artefactual kind properties (such as being a table or being an oscilloscope) and moral properties (such as being kind or being cruel) to the list of properties that some rich theorists have claimed we can consciously perceive.

Our question becomes more complex once we allow for the possibility of the ‘cognitive penetration’ of conscious experience. This is the alleged possibility that cognitive (i.e. non-perceptual) states such as beliefs and desires and hopes and fears (perhaps also biases and prejudices) can affect – in a certain direct way that needs some care to spell out – the phenomenal nature of perceptual experiences. Once we allow for the possibility that cognitive, non-perceptual, perhaps sub-personal mental states that are very far downstream and removed from the impacts on our sensory organs can nevertheless influence which properties ‘show up’ in our perceptual experiences, then we are no longer treating the visual system as modular. And so our initial intuition that at least some properties are clearly not visible may be undermined. For example, returning to the example of radioactivity, perhaps as a result of all her expert beliefs and knowledge ‘penetrating’ the nature of her perceptual experiences, when a radiation expert looks at a chunk of uranium-235 maybe she really can visually experience the property of being highly radioactive as well as visually experiencing its shape and colour properties. The idea here would be that, for the expert, it has a distinctively radioactive look that is over and above the typical visual appearance of the shape and colour and texture properties of uranium-235. We must bear in mind, however, that the reality of cognitive penetration is a contested issue. Some initially striking experimental results that seemed to support cognitive penetration having failed to replicate. Whilst other experiments have been criticized by a number of psychologists for failing to demonstrate the supposed top-down effect of cognitive states on perception.

So far we have not said anything about the different philosophical theories that exist concerning the metaphysics of perceptual experience. For it seems clear that our initial question: “which properties can be consciously perceived?” will arise for any/all of the standard competing theories – Representational, Naïve Realist, Sense-Data. However one conceives of the metaphysics of perception there will still be a question as to which kinds of features can be perceived (by humans) and which cannot. A Naïve realist might think of this in terms of which external properties we can be consciously acquainted with. A sense-data theorist might think of this in terms of which external properties have the right kind of direct, causal relation to the nature of the sense-data array we are acquainted with. A representational theorist will frame the question in terms of which external properties are represented in the content of perceptual experience. It is this last, representational way of framing the debate between Rich vs. Sparse accounts which is the standard framework in the

---

2 E.g. Siegel 2006, Bayne 2009
3 E.g. Siegel 2009, Butterfill 2009, Michotte 1963
4 E.g. Siegel 2010, Butterfill 2015, Helton 2018
5 E.g. Nanay 2011, though he would, I think, reject the label ‘affordance’.
6 E.g. Van Gulick 1994
7 E.g. Cullison 2010, Wisnewski, 2015
8 For example, the results of the ‘classic’ experiments by Bruner & Goodman (1947), which seemed to show that a subject’s belief about the monetary value of a coin affects the subject’s perception of its size, failed to replicate in a number of subsequent experiments. Likewise the well-known results of Delk & Fillenbaum (1965), which purport to show that ‘past association of color and form does in some way influence perceived color’ (p.293), have failed to replicate. More recently, a famous experiment by Bhalla and Proffitt (1999), which seemed to show that wearing a heavily loaded backpack influences distance perception, also failed to replicate in subsequent experiments. See Machery (2015) for a very useful overview of these and other replication failures.
9 For example, see the criticisms made by Francis (2012) and Firestone & Scholl (2014a, 2014b, 2016) of design flaws in many recent experimental studies that purport to show the effects of cognitive penetration. Again, see Machery (2015) for an excellent discussion of these methodological concerns.
literature, reflecting the fact that representational theories of experience are the current philosophical orthodoxy. But notice, once we shift to considering whether high-level or only low-level properties can figure in the representational contents of perceptual experiences, we are no longer straightforwardly addressing exactly the same kind of question we started with – which properties can we see? which properties can we hear? This is because it could be that properties can figure in the content of perceptual experience without being perceived (or perceptible). One way this could be possible is if perceptual experiences systematically (mis-)represent properties that are not in fact present in the external scene. It may also be possible that certain high-level properties can enter into the content of veridical perceptual experience – so this content could be accurate rather than illusory – and yet nevertheless these properties are not perceived because the causal conditions on perceiving something (whatever those conditions are exactly) are not met.

Whilst there are various empirical results that are clearly relevant to the Rich vs. Sparse debate, we currently lack any kind of agreed upon theory as to when activity or information in the brain ‘enters into’ or ‘show up’ in the stream of consciousness – and so we are not currently in a position to experimentally determine which properties are consciously perceived. Given the current scientific intractability of deciding the Rich vs. Sparse debate, philosophers have had recourse to introspective or phenomenological considerations. Perhaps the main, most frequently discussed such tool is the Method of Phenomenal Contrast. The basic idea is to consider cases involving a pair of experiences that are almost identical except for some one key respect in which they contrast – and then to ask: what is the best explanation for this phenomenal contrast? Must we appeal to the perceptual representation of Rich, high-level content in order to explain the phenomenal contrast, or can the contrast be explained simply by appeal to Sparse, Low-level properties? Employing this method is in principle open to all the different theories of perceptual experience – Naïve-Realist, Sense-Data, Representational. But again, discussion of the method of phenomenal contrast has in fact almost always been pursued within a broadly representational framework. And indeed the issue I will be focusing on in this paper is specific to representational accounts of experience, for I will be considering how the method of phenomenal contrast and the Rich vs. Sparse debate interacts with the thesis of externalism about mental content. I will suggest that this debate about Rich vs. Sparse

---

10 For example, an eliminativist about colours thinks that there are really no such things as colours, at least qua features of physical surfaces and volumes. Nevertheless, an eliminativist can still allow that the inaccurate representation of physical colour properties is part of the content of the experience and is something that ‘shows up’ in the phenomenal character of the experience.

11 For example, if we accept that cognitive colour properties are part of the content of the experience and is something that ‘enters into’ the phenomenal character of the experience.

12 It is a relatively straightforward empirical question as to what stimuli our sensory transducers are sensitive to – what the cells in the retina or the cochlea are sensitive to, etc. And it is also an experimentally tractable question as to which properties in the environment reliably co-vary with certain patterns of neural activity in the areas of the brain that are immediately downstream from our sensory transducers – e.g. populations of neurons in V1 often function as ‘edge detectors’, whilst neurons in V5 are sometimes ‘tuned’ to the direction of motion of visual stimuli. See e.g. Hubel & Wiesel (1959, 1962), Shapley & Tollhurst (1973), Dubner & Zeki (1971), Maunsell & Van Essen (1983).

13 Another experimental approach that bears on the Rich vs. Sparse debate is to investigate adaptation to stimulus effects. It is very well-known that sensory systems adapt to a constant level of stimulus. Experimental results due to Kaping, Bilson, & Webster (2002) suggest that this sort of adaptation effect can also occur with high-level properties: when subjects fixated on a face with very masculine features and then looked at a new face, this second face tended to look more feminine than it otherwise would. This has been adduced as evidence in favour of the thesis that we genuinely perceive high-level properties such as masculinity or femininity, since adaptation effects are often thought to be a distinctly perceptual phenomenon. See also: Javadi & Wee, 2012, Block 2014.
Theories and Phenomenal Contrasts has tended to ignore\textsuperscript{14}, to its detriment, the issue of externalism about content. Both Rich theorists and Sparse theorists have tended to freely talk about perceptual experiences possessing contents about properties that are ‘Twin-Earth-able’. But assuming that there is some kind of explanatory/determination relation between perceptual content and perceptual phenomenology, I will argue that we face a choice: either our account of phenomenal contrast cases ought only appeal to content that cannot be Twin-Earth-ed, or we must accept that externalist factors outside of the subject’s body are (at least partly) determining the phenomenal character of experience. Neither of these alternative commitments has generally been acknowledged in discussion of phenomenal contrasts or Rich vs. Sparse content – and indeed either option might well be considered an unattractive theoretical cost by at least some philosophers.

The plan for this paper then will be as follows: in section 2 I briefly clarify the method of phenomenal contrast and how it bears on the Rich vs. Sparse debate. In section 3 I discuss the different possible relationships that might be thought to hold between perceptual phenomenology and perceptual content. In section 4 I discuss how externalism about content bears on cases of phenomenal contrast and the choice point that this creates. Finally, in section 5 I sum up and draw some more general morals about the relation between content and phenomenal character.

2. The Method of Phenomenal Contrast

So what is the method of phenomenal contrast?

Here is how it figures in the Rich vs. Sparse debate: the method is to identify a pair of cases involving a conscious subject, or subjects, that are very similar in most respects but which exhibit a phenomenal difference or contrast for the subject or subjects in question – and then to ask: what is the (best) explanation for this phenomenal contrast? The Rich theorist will then claim that the best explanation will somehow or other involve the presence/absence of some High-level representational content in the experience. Whereas of course the Sparse theorist will deny this and insist that the best/correct explanation appeals only to low-level perceptual content – or perhaps to other non-perceptual states or non-representational factors (see below). It is worth briefly clarifying that appeals to cases of phenomenal contrast also occur in the literature on cognitive phenomenology. However, this is a different use for phenomenally contrasting cases in a different debate. There the central question is not about perceptual experience; it is about whether thought itself can have a proprietary kind of cognitive phenomenology or whether any phenomenology is due to the phenomenal character of accompanying sensory, imaginative or emotional (etc.) states rather than to the thought itself. In what follows I will be focusing exclusively on the method of phenomenal contrast as employed in the Rich vs. Sparse debate and ignoring phenomenal contrasts as they are employed in discussions about cognitive phenomenology\textsuperscript{15}.

As an illustration, consider Susanna Siegel’s (2006, 2010) well-known example of a phenomenal contrast. In the first case, suppose we have a conscious subject who is not an expert about species of trees and who is staring at a stretch of forest. In the second, contrasting case, the conscious subject is almost exactly the same, mentally speaking, staring at exactly the same patch of forest in the same lighting conditions etc., except that now the subject has expert knowledge about different kinds of trees and how to recognize them. The supposed phenomenal datum is that there is a phenomenal contrast between the two cases; the scene looks different to the expert compared with the novice. And so then a Rich theorist, such as Siegel, will claim that the explanation for this phenomenal contrast between the two otherwise very similar situations, somehow or other involves the presence

\textsuperscript{14} Though see Ashby (forthcoming), who provides an excellent discussion of some similar issues, though the focus of his discussion is different in various ways.

\textsuperscript{15} Supporters of cognitive phenomenology who have also appealed to phenomenal contrasts include: Strawson 1994, Siewert 1998 2011, Pitt 2004, Chudnoff 2015, Kriegel 2015. For an interesting recent argument that use of the phenomenal contrast method in the cognitive domain can only yield much more restricted results than these authors suggest, see Jorba & Vicente 2020.
in the subject’s perceptual experience of content representing high-level properties such as being a pine tree, being an elm tree, etc. In support of this Rich thesis, the idea here is that: the trees don’t look any different regarding their shapes or their colours or their spatial arrangements etc. So the phenomenal difference – the difference in look – does not involve any change in the perceptual representation of their low-level properties. Rather they now look pine-ish and elm-ish. So the change should be somehow explained in terms of the experience representing these high-level properties. Of course, this line of thought will be rejected by a Sparse theorist. A Sparse theorist might simply insist that the phenomenal change does involve a change to the low-level properties that the experiences represent. E.g. in Siegel’s case of looking at the trees, perhaps the distinctive shapes and colours of an elm tree’s leaves are perceptually represented in more detail – with more precision or determinacy – after one becomes an expert about elm trees. Or a Sparse theorist could try to argue that the phenomenal contrast is explained by something other than a change of content in the perceptual experience itself. A number of authors have made important criticisms of the method phenomenal contrast. Koksvik (2015) argues that both the ‘richness’ (i.e. complexity) and the ‘flux’ (the fleeting and constantly changing nature) of one’s total stream of consciousness means that there will always be far too many possible explanations for the overall phenomenal contrast between two cases. Thus, he claims, the method is effectively toothless for resolving disagreements about the nature or content of phenomenal consciousness. And Fürst (2017) argues that our introspective judgements about phenomenal contrast cases (actual or hypothetical) are ‘regimented’ by our prior theoretical commitments about conscious experience. In other words, introspection is theory-laden in such a way that appealing to cases of phenomenal contrast will not be able to resolve theoretical disputes about consciousness. We might also recall here the sort of general concerns with the reliability of introspection that Schwitzgebel (2008, 2011) forcefully presents. I have much sympathy with these kinds of worries concerning the difficulties with effectively employing the method of phenomenal contrast. However, in what follows (see section 4, below), I will be making a different point; even if we grant that the method could be effectively employed, it won’t straightforwardly deliver the sort of results that many of its exponents on either side of the Rich vs. Sparse debate seem to think it can deliver.

3. How might representational content and phenomenal character be related?

When describing the Method of Phenomenal Contrast I used the somewhat evasive phrase ‘explanation which somehow or other involves’ a change in representational content. The reason for this inelegant choice of words is that there is a whole range of different possible views on the relationship between content and phenomenal character. Let’s suppose for the moment that a given phenomenal contrast does somehow or other involve a difference in the perceptual experience’s representational content. One natural way for philosophers to give a more precise formulation of the relation between phenomenal changes and representational changes is in terms of supervenience or entailment:

- Phenomenal Character supervenes on Representational Content
- (Necc.) Phenomenal Contrast → Representational Contrast
- Representational Content supervenes on Phenomenal Character
- (Necc.) Representational Contrast → Phenomenal Contrast

16 For example: perhaps the phenomenal change is explained by some other non-perceptual/cognitive state, perhaps it is explained by differences in attention, perhaps it is explained by differences in non-representational qualia, etc.
17 Note that Fürst’s discussion focuses on the use of phenomenal contrast cases in the debate about cognitive phenomenology, rather than the Rich vs. Sparse debate. However, concerns about the theory-laden-ness of introspection could presumably also apply to this latter debate.
• 2-way Supervenience
• (Necc.) Phenomenal Contrast ↔ Representational Contrast

But of course relations of supervenience or entailment leave open which (if either) of the relata explains or grounds or depends on the other. Two theorists could both accept all of these supervenience/entailment theses and yet they might still disagree about the explanatory priority of representational content vs. phenomenal character. It will be useful then to briefly to sketch the different possible positions one might hold concerning the relation of metaphysical explanation or dependence (if any) that obtains between content and character. Let’s begin with the following highly non-committal, non-specific thesis:

• EXPLANATORY LINK: The difference in phenomenal character between two experiences, E1 and E2, has some kind of important metaphysical explanatory or dependence relation to a difference in the representational contents of E1 and E2.

It has become very common to discuss relations of non-causal, metaphysical explanation or dependence in terms of grounding. And that is certainly the sort of thing I have in mind as a ‘metaphysical explanatory or dependence relation’. Many different views on the exact nature of grounding exist. For present purposes, these differences will not matter. I want to leave the exact nature of the explanatory/dependence relation as open as possible. So one could think in terms of grounding – i.e. representational content grounds phenomenal character, or vice-versa. But some theorists might prefer to think of the relation between phenomenal character and content in terms of, say, constitution or composition. And perhaps other theorists would prefer some other kind of ‘in virtue of’ relation, such as truth-making or ontological dependence, etc. Indeed, I want to leave open that the relation is simply identity – a view endorsed by Dretske (1995), Carruthers (2000) and Smythies (2019). In what follows I will use the phrases ‘determines’ and ‘determination relation’ as neutral, catch-all terms for any of these kinds of relations. It is worth differentiating EXPLANATORY LINK from an importantly different claim about cases of phenomenal contrast:

• EVIDENTIAL LINK: When two experiences contrast in phenomenal character this is evidence that there is also a difference in content.

This no longer talks about any explanatory or dependence relation, but just treats a phenomenal contrast as a sign or indication that there has also been a change in content. Clearly one might take a phenomenal contrast to be a useful guide to changes in content even if one denies that there is any relation of metaphysical determination.

Now, whenever we are considering whether/how one thing, A, metaphysically determines another, B, there seem pretty clearly to be four main options: A determines B, B determines A, Mutual Determination, No Determination. In terms of phenomenal character and representational content then we get the following four basic options:

(1) The Phenomenal Character of Experience Determines its Representational Content

This is perhaps the simplest version of the cluster of views known under the label “phenomenal intentionality”.

(2) The Representational Content of Experience Determines its Phenomenal Character.

This is the view that is sometimes called ‘Representationalism’, sometimes “strong intentionalism”, or sometimes just “intentionalism” – e.g. Tye 2000, Byrne 2001.

---

18 The label ‘phenomenal intentionality’ appears first, I believe, in Horgan & Tienson 2002, though the general idea that intentionality is grounded in phenomenal consciousness goes back at least to Brentano and Husserl.
(3) Some kind of Mutual Determination or No priority view.

Whether this even makes sense as a possibility will depend on the details of one’s specific views about the determination/explanation/grounding relation in question. E.g. many accounts of metaphysical grounding treat the relation as anti-symmetric. A ‘no priority’ view will make sense, of course, if phenomenal character is simply identical to representational content19.

(4) Neither one determines the other.

One clear example of a theory that endorses thesis (4) is David Papineau’s (2021) recent ‘pure mental paint’ account, according to which the phenomenal character of experience has no essential or intrinsic representational nature whatever. But as a matter of contingent psychological fact, in us humans this ‘mental paint’ does (normally) serve, according to Papineau, to represent various environmental features. Options (1), (2) and (3) are different ways of fleshing out EXPLANATORY LINK and so any theorist endorsing any version of these options might well be interested in using the method of phenomenal contrast to uncover the content of experience. But versions of option (4), such as Papineau’s, could hold that a phenomenal change does in fact, for us humans, generally indicate that there has been a change of content20. And so a pure mental paint theorist, though she denies EXPLANATORY LINK, might still be interested in the method of phenomenal contrast because she endorses EVIDENTIAL LINK. However, notice that taking a phenomenal change to be a mere sign that some mental content has also changed is not the same as saying that that content in question is part of what we consciously perceive. For the phenomenal change might be a sign or indication of a change in entirely unconscious, sub-personal content.

We could complicate and sub-divide these four basic options by allowing for the possibility of partial determination, and also for the possibility that just some (but not all) of the phenomenal character and/or the representational content might be involved in the determination relation. For example: perhaps All of the representational content determines Some of the phenomenal character, whilst other phenomenal aspects/features are determined otherwise – e.g. by qualia or non-representational sensational properties. This position is sometimes labelled ‘weak intentionalism’21. Or perhaps: Phenomenal Character partially determines Representational Content. So the phenomenal character of experience contributes to determining the representational content but other factors – e.g. perhaps externalist, causal-historical factors – also contribute22. If I understand correctly, this is the position of a number of recent theorists in the ‘phenomenal intentionality’ camp: e.g. Farkas (2013), Mendelovic (2018), Kreigel (2013).

But more importantly, cross-cutting these four options is the question of whether the determination relation is one-to-one or many-to-one. In other words: can there be a kind of multiple realizability? So, for example, even if we accept thesis (1), that phenomenal character determines representational content, still there is the question whether two or more different (total) phenomenal characters might determine the one same representational content. E.g. perhaps my colour phenomenology represents a certain physical property of surfaces and a pigeon’s different, alien colour phenomenology also represents the exact same physical property. Or perhaps two different sensory modalities, such as

19 Again, see e.g. Dretske 1995, Carruthers 2000, Smythies 2019.
20 So this sort of position could understand the conditional or biconditional claims listed at the start of this section in terms of merely nomic necessity – a contingent law about human psychology rather than a necessary truth of metaphysics.
22 On this view, the phenomenal character of experience is inherently intentional, pointing beyond itself to the external world, but it needs extra help from other factors (external factors) for this pointing to actually amount to a determinate representational content. Notice that this is different from the Papineau ‘pure mental paint’ position, discussed above under Option (4), which treats the phenomenal character as offering no constraint at all on content – i.e. the phenomenal ‘paint’ could in theory stand for anything.
vision and touch with their different phenomenal characters, might sometimes represent the very same content. And even if we accept thesis (2), that representational content determines phenomenal character, still there is the question whether two or more different representational contents might determine the one same phenomenal character. E.g. perhaps representational content that there is a circle before me, viewed head-on, determines a certain phenomenal character. But perhaps also content that there is an ellipse before me, viewed from a specific angle, determines the very same phenomenal character.

A distinction is often drawn between the phenomenal content and the non-phenomenal content of an experience – see e.g. Chalmers (2004, 2006) or Speaks (2009) who uses the labels ‘phenomenology affecting content’ vs. ‘phenomenology silent’. Both Speaks and Chalmers illustrate the distinction between phenomenal and non-phenomenal content by appeal to the idea that experiences can have singular contents that represent specific, particular objects. So, for example, my perceptual experience might represent of this specific tennis ball, B1, that it is spherical and yellow and located in front of me. But it is very plausible that I, or perhaps someone else, could have had a phenomenally identical (or at least indiscriminable) experience of a different, though visually indistinguishable, specific tennis ball, B2. And that the content of this phenomenally identical experience would be that B2 is spherical and yellow. We have then a difference in representational content between these two experiences without any change (or at least without any noticeable change) in phenomenal character. And so, according to both Speaks and Chalmers, these singular contents specifically about B1 or B2 are not part of the phenomenal content of either experience. What is important for present purposes is that there is a way of resisting this line of thought that says: singular, object-dependent content cannot be determining or ‘shaping’ phenomenal character, since the one same phenomenal character might represent different particular objects on different occasions. One could resist this line of thought if one accepted that two different contents can both determine the one same phenomenal character – i.e. if there can be a kind of multiple realizability of phenomenal character by different contents. In which case these two different singular contents would be determining phenomenal character, they just both happen to determine it in exactly the same way. Of course, there is nothing especially radical about the idea that a specific phenomenal character could be multiply realised by two different underlying neural states. But what will presumably strike many people as radical about the idea that two different singular contents can realize or determine the one same phenomenal character is that this means that factors outside of the subject’s body are (at least partially) determining the subject’s phenomenology. Which brings us neatly onto the issue of externalism…

4. Twin-Earth-able Contents

So now let’s turn to considering Twin Earth type cases.

Here on Earth a subject looks at some trees at t1 – they look a certain way. Later, at t2, the subject has gained expertise about elms and how to spot them and now looking at the exact same scene some of the trees look different; their elmy-ness seems salient and ‘pops out’. Over on Twin Earth, an internal duplicate looks at some visually indiscriminable trees at t1 – they look a certain way. Later, at t2, the twin subject has gained expertise about “schmelms” – trees that are visually indistinguishable from elms but which are a different species with different DNA. And now at t2, looking at the exact same scene some of the trees look different – their schmelmy-ness seems salient and pops out. Throughout the course of the respective phenomenal changes the two subjects remain internally identical to each other at all times. They both undergo exactly the same internal changes from t1 to t2, but on earth it is elms that are looking different at t2 whereas on Twin Earth it is Schmelms that are looking different at t2. It seems clear then that in the latter, expert state, when
both twins are still internal duplicates, if their visual experiences really do represent the natural kinds of trees, then one twin’s experience represents elmhood and the other’s schmelmhood.23

There is now a choice for those who take these phenomenal changes to somehow or other involve a change in perceptual content – i.e. for anyone who accepts any version of theses (1), (2) or (3).

- EITHER: phenomenal character supervenes only on the internal state of subject.
- OR: external factors, outside of the subject’s brain and body, at least partially determine phenomenal character.

If one goes for the former option, then representational contents about properties that are Twin-Earth-able, hence which supervene on external factors, cannot be determining/determined-by the phenomenal contrast. So, for example, Siegel’s case of a phenomenal contrast could not be explained by, or be explaining, the representation of either elm-hood or schmelm-hood. On this option, anyone who employs the method of phenomenal contrast should not appeal to Twin-Earth-able mental representations of properties as playing a determining/determined-by role vis-à-vis some specific phenomenal contrast.

The latter option is to embrace some form of phenomenal externalism. Whilst this is these days a respectable philosophical position with a number of prominent defenders, it remains very much a minority view. Although it is commonly (but by no means universally) allowed that factors outside of our heads can (partially) determine representational content, the idea that such externalist factors can (partially) determine our conscious phenomenology is much less popular. One kind of phenomenal externalism simply denies that internal duplicates are always phenomenal duplicates – their respective experiences can have different phenomenal characters even though they are internally identical. Of course, Naïve Realists – such as Martin (2002, 2004), Brewer (2011), Fish (2009), Logue (2012) – accept this kind of phenomenal externalism, as they think that phenomenal character is partially constituted by external features. It is worth briefly noting that there is also an important school of thought based around the idea that perceptual experience, especially spatial experience, somehow constitutively depends on action – or at least on dispositions to act, or on knowledge of how actions affect appearances – where the actions in question are to be specified in externalist, environment-involving terms rather than just as internal efferent signals or muscle-activations, etc. Something like this idea can be found in the work of ‘Enactivist’ theorists such as Hurley (1998), Noe (2004) and O’Regan (2011). But it can also be found in some seminal remarks by Gareth Evans (1982), later developed in detail by Grush (2007, 2009) and Briscoe (2009, 2014). On these sorts of action-based theories, the phenomenal spatial content of experience is constitutively linked to the subject’s possible actions, characterized in externalist, egocentric terms – e.g. Evans provides the examples: ‘extending the arm’, ‘walking forward two feet’.24 It is a somewhat delicate question whether these sorts of positions should really count as instances of phenomenal externalism, in the sense of allowing that internal duplicates could have phenomenally different experiences. On the one hand, the idea is that the phenomenal character of experience is somehow essentially tied to its role in explaining actions, characterized in wide, environment-involving terms.25 On the other hand, if what is really doing the explanatory work is something like dispositions to act/behave, or knowledge of how actions would affect sensory appearances, it becomes less clear that such theories should really count as phenomenally externalist.26 In any case,

23 I assume that there can be no plausible basis for saying that just one twin’s experience veridically represents the natural kind of tree in his/her local environment but the other twin’s does not. Ashby (forthcoming) explicitly labels this assumption ‘symmetry’.

24 See Evans 1982, p156.

25 Noe (2008, p.703) writes: ‘My view is as naïve realist as one can hope to be’, which would seem to support interpreting his brand of Enactivism as a kind of phenomenal externalism.

26 Grush (2007), states that so long as the relevant areas in the brain can still encode the right connections between sensory signals and behavioural signals then the actual production of behaviour/action (or lack thereof) is irrelevant to the subject’s spatial experience. Likewise, in reply to Ned Block, Kevin O’Regan
there is a different possible kind of phenomenal externalism which accepts that internal duplicates are also phenomenal duplicates. Siegel, without endorsing or disavowing it, briefly notes the possibility of this kind of position – though the phenomenal externalism may not be immediately apparent on first reading:

“When the Rich Content View is made true by visual experience representing natural kind properties, such as the property of being a pine tree. And suppose one accepts externalism. It is open to someone who accepts both of these claims to hold that physical duplicates whose environments differ (where only one includes pine trees) have the same visual phenomenology. This would entail that the property of having that visual phenomenology is not identical with the property of representing the property of being a pine tree in experience. But it is compatible with the view that that visual phenomenology supervenes on the contents of visual experience.” (Siegel, 2006, 501)

In other words: theorists who accept content externalism but who nevertheless want internal duplicates to be phenomenal duplicates could maintain that even though the internal duplicates are indeed phenomenal duplicates, in fact this common phenomenal character supervenes on different externalistically individuated contents in each case. So the suggestion is that a specific phenomenal character might be multiply realizable by different sets of externalist factors (as discussed at the end of section 3, above).

I certainly do not wish to dismiss any of these varieties of phenomenal externalism. (For what it’s worth I am actually pretty sympathetic27.) The point is just that no theorists that I am aware of on either side of the Rich vs. Sparse debate have explicitly allowed that their position involves a commitment to phenomenal externalism. And for at least many philosophers, such a commitment will seem to be an unattractive theoretical cost.

Now, the choice outlined above might initially appear just to be an issue for some of the Rich Theorists who have employed the method of phenomenal contrast, since mental representations of natural kind properties, like being an elm or being uranium 235, are clearly Twin-Earth-able (if anything is). But in fact this issue is very plausibly also a worry for Sparse theorists, since there can very plausibly be Twin Earth style scenarios involving low-level properties like shape and colour. E.g. there could be ‘El Greco worlds’ which are just like earth except everything is systematically stretched in one direction28. But if an El Greco world subject has been the life-long wearer of distorting goggles which perfectly counteract the stretching on this world, then there could be an internal duplicate of me, enjoying phenomenally identical or at least indistinguishable experience, but where the duplicate’s conscious mental states represent different, stretched-out shape properties. And of course there is Block’s (1990) classic inverted earth scenario, which seems to show that it is possible for an internal duplicate of me, enjoying the same phenomenal characters as me, to have mental states which systematically represent different external colour properties. Or we could follow Chalmers (2012) and think in terms of Matrix style scenarios instead of Twin Earth style scenarios – the upshot is apparently the same: there can be an internal duplicate of me whose mental states are reliably caused by and hence represent very different properties – Matrix squareness and Matrix roundness, Matrix red and Matrix blue, which are totally different to physical squareness and physical roundness, redness and blueness. But whichever style of externalist thought experiment we prefer, the point here is simply that both sides in the Rich vs. Sparse debate have tended to freely appeal to perceptual contents that are Twin-Earth-able (or Matrix-able).

If one needs to avoid appealing to any Twin-Earth-able representations of properties in the content of an experience when giving an account of a phenomenal contrast, this naturally leads to the question:

27 See Raleigh 2014, 2021
which mental representations of properties are Twin-Earth-able and which are not? At first pass, a subject’s mental representation is Twin-Earth-able if there can be a duplicate or twin of the subject in a different external context whose corresponding mental state would have a different extension – i.e. picks out a different property (or perhaps even fails to refer to any property whatsoever). And so, very roughly speaking, it seems that we can create Twin Earth style cases for representations of properties that have something like a further hidden nature or essence that goes beyond their manifest nature or appearance. Whereas, conversely, representations of properties that do not have any such further hidden nature or essence are not Twin-Earth-able. However, this characterization leaves room for significant philosophical disagreement. Firstly, there are different ways of understanding what a ‘duplicate’ or ‘twin’ should be in a Twin Earth scenario. The standard way of presenting Twin Earth cases appeals to intrinsic duplicates – i.e. two subjects who are atom-for-atom physically identical, from the skin inwards. But both Farkas (2003, 2008) and Chalmers (2012) argue that the crucial notion is actually of a phenomenal duplicate, i.e. two subjects who have exactly the same (or at least subjectively indistinguishable) conscious experiences. Secondly, it is not always just obvious whether or not a duplicate’s mental state would or would not have a different extension; philosophers’ intuitions/reactions to purported Twin Earth cases might reasonably differ. Chalmers gives the following list of concepts that he suggests are plausibly not Twin-Earth-able:

“Putnam-style arguments support Twin-Earth-ability for natural kind terms, names, and the like, but on the face of it they do not extend to expressions such as ‘bachelor’, ‘friend’, ‘action’, ‘conscious’, ‘cause’, ‘part’, ‘fundamental’, ‘two’, and ‘and’ (or at least, they do not extend to certain core uses of these expressions). (Chalmers, 2012)

Even if we grant that Chalmers is correct about these concepts, there are various concepts/properties for which it might be harder to say whether or not they are Twin-Earth-able. For example: consider an artefactual property, such as ‘being a chair’. Prima facie, there is nothing more to being a chair than being the sort of thing that is or can be sat on – there is no further hidden nature or essence to chair-hood that we could discover. But if we start to think about Matrix-style thought-experiments, one might begin to wonder if the mental representation of a life-long inhabitant of the Matrix would really be referring to chairs? Or, for that matter, in a Matrix scenario would there really be anything that counts as sitting? And so would a twin subject, native to the Matrix have a mental representation that picked-out sitting or would it rather refer to some other strange Matrix phenomenon? Would a life-long inhabitant of the Matrix have mental representations that picked out the same moral properties as our moral concepts do? Indeed, once we start to think about Matrix-style scenarios, it may begin to seem a little less clear than Chalmers suggests that, say, the concept ‘cause’ cannot be Twin-Earth-able. For you might think that within the Matrix simulation every apparent instance of A causing B would actually be an instance of both A and B being jointly caused by a third factor, C – the computer or machine on which the simulation is being run. So would a life-long inhabitant of the Matrix have a mental representation that has the same extension as our concept of ‘causation’? Or would the Matrix-dweller’s concept refer to cases of (what we would think of as) mere epiphenomenal succession? Interesting, or annoying, as it might be to speculate on these questions, I certainly have no intention of trying to settle them here! The point is just that it might well be tricky to decide for at least some high-level properties – artefactual properties, action/affordance properties, moral properties, causal properties – whether or not their mental representations would be Twin-Earth-able. And of course whilst I suggested above that there are also plausible Twin Earth style

29 Though, as is very well-known, Oscar and Twin Oscar in Putnam’s (1973) original case were not perfect physical internal duplicates, since Oscar’s body would contain H2O, whereas Twin Oscar’s body would contain XYZ.
30 In fact, Chalmers ultimately prefers the notion of a functional-phenomenal duplicate.
31 Horgan & Timmons (1991) discuss the possibility of a ‘Moral Twin Earth’, but argue that we should reject the idea that the extensions of our moral terms, or of our twins’ moral terms, are determined by externalist-causal factors. Moral Twin Earth has since been much discussed – see Laurence, Margolis & Dawson (1999) for one influential reply to Horgan & Timmons.
scenarios for low-level properties such as shape and colour, perhaps some philosophers will be unpersuaded by these cases.\(^{32}\)

In any case there are at least some mental contents that seem pretty clearly to be not Twin-Earthable. Firstly there could be mental content about appearance properties. Shoemaker (1994) suggests that experience represents appearance properties as well as more familiar properties like shape and colour etc. E.g. the property of looking circular, or the property of looking like an elm. These are properties that can perfectly well figure in the content of my twin’s mental state on Twin Earth – and indeed they are properties that can be instantiated by the objects on Twin Earth. A schmelm just as much as an elm can instantiate a certain kind of appearance property, which we might equally call ‘looking elmish’ or ‘looking schmelmish’. And so it is an open option to explain the before and after phenomenal contrast in terms of the latter experience representing this appearance property rather than its representing elm-hood or schmelm-hood. Siegel herself allows that worries about externalism might push one to accept that the property that is represented by the expert’s experience ‘is not the property of being a pine tree but a more general kind property that both pine trees and superficially similar trees share.’ (Siegel, 2006, 501-2). But once we start to think of ‘trees’ in El Greco world and on Inverted Earth and in the Matrix etc. it seems that the only suitably ‘general’ property that all these cases could share would be something like an appearance property.

In a similar vein, another possible kind of content that would not be Twin-Earth-able is what Chalmers (2006) calls ‘Edenic Content’, a kind of narrow content that refers to ‘Edenic’ properties such as Edenic red-ness or Edenic square-ness, properties which would be exactly as they appear, but which are never actually instantiated anywhere in our universe. Likewise, it seems clear that if experiences have Fregean contents, then these Fregean contents will supervene only on the intrinsic nature of the subject – i.e. they would be narrow contents – and so would not be Twin-Earth-able. A Fregean sense is roughly a condition on reference. And a standard way of glossing the supposed Fregean content of some experience which apparently presents some external property is something like: the property (whatever it is) that is the normal cause of this kind of phenomenal experience. So the Fregean content of my experience when I see something red would be representing an external colour property via the condition: whatever property it is that is the normal cause of phenomenal-red experiences. This narrow, Fregean content is something that my experience on Earth and the experience of my twin on Twin Earth and the experience of a phenomenal duplicate in the Matrix could all share – even if these three experiences would all have different wide contents, which refer to the three different properties that are in fact the normal causes of the experiences in these three different contexts. And so one could maintain that an experience can represent or refer to a Twin-Earth-able property (in a context) via a Fregean content that is not Twin-Earth-able. This sort of appeal to Fregean content is the route taken by Tim Bayne (2009), one of the very few theorists in the Rich vs. Sparse debate who not only employs the method of phenomenal contrast but who also considers the issue of externalism about mental content. Bayne is himself a Rich theorist. But it is worth noting, given that Twin Earth cases very plausibly impose constraints on Sparse theorists too, that a Sparse theorist could equally well appeal to Fregean contents in defence of the idea that experience represents low-level (but Twin-Earth-able) properties such as shape or colour. So the debate between Rich vs. Sparse theories can still arise on the assumption that perceptual experience has a narrow Fregean content. The debate then becomes whether these Fregean contents can sometimes serve to pick out different high-level properties (such as the natural kinds Elm or Schmelm) in different contexts (such as Earth vs. Twin Earth), or whether these Fregean contents

\(^{32}\) For example, Jeff Speaks writes: “Where, one might ask, should we draw the line between contents which are ‘externalist’ and those which are ‘internalist’—i.e., between those contents which are such that having a perceptual experience involving them does not supervene on intrinsic properties, and those which are such that having a perceptual experience involving them does so supervene? We should look to externalist arguments, and see what conclusions they support. It is very easy to generate externalist arguments in the case of perceptual representations of external particulars or natural kinds, but…these arguments do not clearly carry over to the cases of color and shape properties. A reasonable view therefore seems to be that particulars and natural kinds are externalist, and that ‘sensible qualities’ like color and shape properties are internalist.” (Speaks, 2009, 566)
only pick out different low-level properties (such as different physical shape properties) in different contexts (such as Earth vs. El Greco World). Of course, the idea that perceptual experience really does have a narrow, Fregean content is highly controversial. For example, Speaks (2009) argues that Fregean contents for experience would be in tension with the Transparency of experience and so could not be determining phenomenal character. More generally, many philosophers have simply denied that there is any such thing as narrow content in the first place\textsuperscript{33}.

In summary then, the moral to draw here is that for at least most of the theorists who have been using the method of phenomenal contrast, \textit{whether they are rich or sparse theorists}, it is not going to be able to do the work it has been supposed to do unless they take on one of two very substantial and potentially controversial commitments. For anyone who thinks there is some kind of determination relation between phenomenal character and representational content faces the choice (to repeat):

- EITHER: appeal only to content that cannot be Twin-Earth-ed.
- OR: accept that factors external to the subject are playing some determining role fixing the phenomenal character

No theorists, as far as I’m aware, have explicitly acknowledged that their position on the Rich vs. Sparse debate involves a commitment to phenomenal externalism. But equally, very few of the theorists in this debate have limited themselves to discussing only non-Twin-Earth-able contents. Unless phenomenal externalism is embraced, then neither the representation of most high-level properties (like natural kind properties) nor the representation of most low-level properties (shape and colour) will be suitable to determine/be determined by the phenomenal change in phenomenal contrast cases. These contrasts would instead have to be explained by appeal to non-Twin-Earth-able contents: e.g. representing appearance properties, or Edenic properties, or a Fregean modes of presentation. But this is not what either side in the debate has generally done.

If one holds a version of option (4), such as Papineau’s pure mental paint position, then one denies that there is any kind of metaphysical determination relation between content and phenomenal character – i.e. one denies EXPLANATORY LINK. So of course \textit{a fortiori} one denies that Twin-Earth-able content depends on or explains the phenomenal contrast. However, recall that such a theorist could still perfectly well treat a phenomenal contrast as a \textit{sign or evidence} that in fact there has also been a change in content – i.e. one could endorse EVIDENTIAL LINK. Indeed it could be a contingent law of human psychology that a change in phenomenal character goes together with a change in content. Now, in one way a pure mental paint theory can be thought of as a \textit{maximally rich} position, for in theory any property could be the normal reliable cause\textsuperscript{34} of a certain configuration of mental paint and that is all that is required for the subject to count as consciously perceiving that property. So the change in phenomenal character – e.g. after one becomes an expert at identifying elms – could well be a sign that the experience now represents elm-hood, even though that new content does not actually determine or explain the change in phenomenal character (nor vice-versa). But on the other hand, this sort of pure mental paint position could also be thought of as a \textit{maximally sparse} position. For on this sort of view no external properties whatsoever, neither High-level nor Low-level, are ever really \textit{presented} in perceptual consciousness. We are not literally directly conscious of, say, square-ness, we are only ever really consciously aware of a certain configuration of our own ‘mental paint’, which is not itself square, but which happens to correlate or co-vary with physical square-ness. And so on this sort of view there really is no very deep/important distinction between those external properties in the environment that are visible and those that are not, since any external property could in theory be visible in the sense of reliably causing some mental paint configuration and so being what that configuration of mental paint (contingently) represents. So if this kind of pure mental paint theorist is interested in the question of which properties we can actually perceive – high level or just low level properties – using the method of phenomenal contrast

\textsuperscript{33} See e.g. Sawyer 2007, Burge 2010, Yli-Vakkuri and Hawthorne 2018.

\textsuperscript{34} One might plausibly need a further requirement that the property play the ‘right kind’ of causal role to deal with familiar worries about ‘deviant’ causal chains.
would not seem to be a natural way of proceeding. For what really settles this question, according to the pure mental paint theorist, are external facts about causal-historical links between the subject’s sensory systems and their environment. Carefully introspecting your own experience after some phenomenal contrast has occurred is not going to reveal whether it is a rich, high-level property or a sparse, low-level property that the new phenomenal character represents. For on this view the new phenomenal character is in itself just an arbitrary symbol that could in principle represent anything whatever (or nothing whatever). It is only if you think that external properties are in some sense presented to the mind in perceptual experience that it would make sense to attend carefully to the new phenomenal character of one’s experience and think, e.g. ‘Aha! this new way of looking is a distinctively elm-y way of looking; it is not just a new shape or colour appearance’ and then encourage others to introspect their own experiences in the hope that they will agree with you.

5. Conclusion

The foregoing discussion focused on the topic of Rich vs. Sparse views of perception and the method of phenomenal contrast. The moral I drew is that cases of phenomenal contrast will only bear on the richness or sparseness of perceptual content once one or other substantial and potentially controversial commitment is endorsed. But a more general moral here concerns the idea that content and phenomenology stand in some kind of metaphysical determination or explanatory relation. Anyone who endorses some version of this idea – options (1)-(3) in section 3, above – faces an issue with Twin Earth scenarios. If the content of a mental state is Twin-Earth-able then there can be an internal duplicate whose mental states have different contents. What this points to is a tension between the following three theses, all of which will be individually attractive to at least many theorists:

• Internalism about the Phenomenal Character of Experience. (I.e. Phenomenal character supervenes only on internal factors.)
• Externalism about the Representational Content of Experience. (I.e. Some mental content supervenes on external factors.)
• There is some relation of metaphysical determination between Representational Content and Phenomenal Character.

These three ideas will be straightforwardly inconsistent if one’s preferred relation of metaphysical determination commits one to the following: (Necc.) Representational Contrast $\rightarrow$ Phenomenal Contrast. I.e. if one is committed to content supervening on phenomenal character. The three options in response would then be to:

• Accept Phenomenal Externalism. (I.e. some phenomenal character supervenes on external factors.)
• Accept that Experience only has narrow content. (I.e. perceptual content supervenes only on internal factors.)
• Deny that there is a metaphysical connection between content and character which commits one to: (Necc.) Representational Contrast $\rightarrow$ Phenomenal Contrast.

One version of the last response would simply deny that there is any metaphysical connection at all – i.e. embrace option (4) from section 3, above. Alternatively, one could retreat to endorsing only the following: (Necc.) Narrow Representational Contrast $\rightarrow$ Phenomenal Contrast. This would allow one to endorse all three of the attractive theses, but only by accepting that whilst experience may possess wide content, this wide content plays no role in the determination relation with phenomenal character. Likewise, if one is only committed to the conditional: Phenomenal Contrast $\rightarrow$ Representational Contrast (i.e. Phenomenal character supervenes on content), this allows that some representational differences might make no phenomenal difference. So again, the tension could be defused by accepting that whilst experience does possess wide content, this wide content is not determining or determined by phenomenal character. Only narrow, non-Twin-Earth-able content is
involved in the determination relation. If it is correct that contents about familiar low-level properties, like shape and colour and distance etc., are Twin-Earth-able, then this places much more of a restriction on what sort of content might determine, or be determined by, phenomenal character. I am not the first person to note a tension between the three theses above (see e.g. Ellis 2010). But I do claim that this tension has not been sufficiently appreciated in the literature on Rich vs. Sparse accounts of perceptual content\(^{35}\) or by those who would employ the method of phenomenal contrast\(^{36}\).

References

- Ashby, B. (forthcoming) “The Price of Twin Earth”, *Philosophical Quarterly*
- Byrne, A. 2009 ‘Experience and Content’, *Philosophical Quarterly*, 59(236): 429–51

---

35 Though, to repeat, see Ashby (forthcoming) and also Papineau (2021).
36 Earlier versions of this paper were presented at the Ruhr University Bochum, the Norwegian University of Science & Technology (NTNU) and as one of a series of online talks organized by Rice University. I am grateful to the audiences on all those occasion for their questions and comments. Not for the first time, I owe a special debt of thanks to Anders Nes, whose typically insightful comments saved me from at least one embarrassing mistake. Many thanks also to: Brandon Ashby, Peter Brössel, David Chalmers, Laura Gow, Jonathan Knowles, Uriah Kriegel, Giulia Martina, David Papineau, Susanna Siegel, Jeff Speaks, Filippo Vindrola, and Simon Wimmer. Finally, I am grateful to an anonymous referee for this journal whose report substantially improved this paper.