The Tractability of the Debate on Relationalism

Penultimate Draft

Roberta Locatelli

Abstract

The debate between relationalism and representationalism in the philosophy of perception seems to have come to a standstill where opponents radically disagree on methodological principles or fundamental assumptions. According to Fish (this volume) this is because, not unlike Kuhnian scientific paradigms, the debate displays some elements of incommensurability. This diagnosis makes advancing the debate impossible. I argue that what is hindering progress is not a clash of research programmes, but a series of misunderstandings that can be avoided by disentangling the different questions each theory is invested in and by making explicit the hidden assumptions at play in the debate. One such central assumption is what I call the Superficiality Constraint. This is the idea that the phenomenal character of experience is superficial with respect to introspection. I argue that we can make progress in the debate by assessing to what extent and at what cost relationalism can accommodate this constraint.

Keywords

perception; disagreement; relationalism; representationalism; introspection

1. The Debate on Relationalism

One of the liveliest debates in the contemporary philosophy of perception concerns the way perception relates us to the world. The dominant alternatives in this debate are representationalism and relationalism. Representationalism, which has become
something of an orthodoxy since the second half of the twentieth century, claims that perception has a representational content, capable of being veridical or illusory, which represents the way things appear to be to the subject. Relationalism denies that perception is (fundamentally) representational: instead, it is a non-representational relation of acquaintance with mind-independent objects.

Much of this debate has focused on how the alternative views account for non-veridical experiences, in particular hallucinations. For representationalists hallucinations are the same kind of experience we have when we perceive veridically: the only difference from veridical perceptions is that their content typically incorrectly represents how things are in the world. But for the relationalist, the kind of experience we have when veridically perceiving (the obtaining of a relation of acquaintance with mind-independent objects) is not one you can have when hallucinating, where the relevant objects are absent. Therefore, most relationalists embrace disjunctivism, the view that perception and hallucination differ in some important way. Proponents of disjunctivism disagree as to where exactly this difference lies. For reasons I will discuss later on, here I will focus on a version of disjunctivism, often called ‘phenomenal disjunctivism’ (Haddock and Macpherson 2008), which locates the difference between perception and hallucination in their respective phenomenal character. For simplicity, I will simply use the term ‘disjunctivism’ to refer to this particular claim.

Disjunctivism has been met with strenuous resistance by many philosophers, for reasons that fall under two main lines of reasoning:

(1) The very notion of phenomenal character entails that whenever two phenomenal characters are indistinguishable, they are qualitatively identical, as there is no seem/is distinction when it comes to phenomenal character. This idea goes often under the name

(2) Disjunctivism is incompatible with vision science because the latter is committed to either of the following claims:

(a) a visual state ‘causally depend[s] only on proximal stimulation, internal input, and antecedent psychological condition’ (Burge 2005, 22) (same cause same effect principle) (see also Smith 2008, 2002; Foster 2000; Sollberger 2007);

(b) the phenomenal character of a visual state locally supervenes on the neural activity (Robinson 1994, 151–2).

Proponents of relationalism retort that both objections are ill-founded:

(1) Nothing—either in the notion of phenomenal character or in what we know about how introspection works—forces us to accept the Indistinguishability Principle (Hinton 1967b, 1967a; Martin 1997, 2003, 2006; Fish 2008, 2009).

(2) Disjunctivism is not incompatible with science because:

(a) The ‘same cause same effect principle’ begs the question against relationalism, as it ‘rules out the possibility that relational states of affairs or events can form part of a causal nexus where relational states of affairs may differ purely in their distal elements’ (Martin 2006, 368);

(b) Local supervenience is only a working assumption concerning the relation between neural and mental states, and is not better supported by empirical data than competing models (Fish 2009, 117–44).

---

1 These two claims are distinct, as the local supervenience claim doesn’t entail causal dependence, but they are very similar inasmuch as they both rule out the possibility that subjects being in the same neural state can have experiences with different phenomenal characters.
Opponents of relationalism, in turn, insist on the validity of the principles they rely on or insist that, somehow, disjunctivism is ‘a bitter if not impossible pill to swallow’ (Knight 2013, 3).

Each camp seems to think that the opponents grossly misunderstand their position, fail to see the inescapability of some principle (or, on the contrary, blindly accept principles that should not be taken for granted), or have an irreducibly different (possibly wrong-headed) way of thinking about the methods and purposes of philosophical reflection on perception.

This fuels scepticism about the possibility of ever adjudicating on its own terms a debate which seems to rest on a disagreement over fundamental assumptions and methodological approaches. In a paper from 2005 and in his contribution to this volume, William Fish voices this sort of scepticism and likens the debate over disjunctivism to Kuhn’s incommensurable competing paradigms.

In the next section I will argue that, despite Fish’s attempts to give it a positive spin, his diagnosis of the debate at hand leaves us without any way of adjudicating it (and possibly other philosophical debates). This outcome is certainly something most philosophers would like to avoid.

The good news is that we do not have to be resigned to such defeatism. Contrary to Fish, I will argue that the current lack of progress in the debate is not due to an insurmountable structural feature of this debate but to the failure to make explicit and address an implicit assumption which motivates the widespread scepticism towards disjunctivism. In order to identify this implicit assumption, in section 3 I will start by surveying how the current debate originated from the attempt to reconcile two apparently incompatible, but equally compelling, intuitions about perception. This suggests that adjudicating the debate depends on assessing the respective costs and
benefits of competing views and that, most likely, no solution will come without a certain price to pay. However, not every price is equal, and one of philosophy's tasks is to guide us in weighing up which price is worth paying.

In section 4 I will argue that the failure to address an implicit assumption that operates in the background of the debate has hindered the efforts to assess the real costs of accepting relationalism. The assumption is that there is a basic constraint to any meaningful account of phenomenal character, which I dub the ‘superficiality constraint’. This is the idea that the phenomenal character of experience is tied to introspection in a way that limits the pattern of possible introspective mistakes we can introspectively make about phenomenal character. In section 5 I will argue that the superficiality constraint is effectively embedded in the way the notion of phenomenal character is used in the debate, so all parties need to accept it. In section 6 I will formulate an argument against relationalism based on this assumption, and I will briefly discuss what relationalists could say to address it. This will shed a new light on the debate and offer a roadmap to advancing it: whether the cost of accepting relationalism is one that is worth paying will depend on how satisfactory their responses to the argument from superficiality are.

2. Philosophical Disagreement and Incommensurable Paradigms

In his contribution to this volume, Fish focuses on the heated dispute between Burge (2005) and McDowell (2011). He suggests that it ‘displays enough of the classic features of a clash of Kuhnian paradigms—just with philosophical theories, rather than scientific theories, playing the role of paradigms’ (Fish this volume: xx).
According to Kuhn, science undergoes alternating phases of normal science and revolutions. In periods of normal science, there is only one prevailing scientific paradigm—a set of beliefs, instruments, metaphysical and methodological assumptions, and models to carry out scientific enquiries. When problematic experimental findings build up and the paradigm encounters increasing difficulties in solving the new puzzles that arise, science enters a phase of revolution, where competing paradigms might emerge. The competing paradigms are, according to Kuhn, incommensurable: there is no objective way (i.e. neutral with respect to the assumptions and methods of each paradigm) of comparing and evaluating them.

Fish notices elements of incommensurability in the debate between Burge and McDowell, inasmuch as ‘each finds different ways of speaking to be “natural”, and attempts to translate between the two “languages” are fraught with difficulty and the potential for misunderstanding’ (Fish this volume)\(^2\) and each has their own distinctive ‘list of problems that any candidate for paradigm must resolve’ (Kuhn 1962, 148,

\(^2\) An example of this is McDowell’s understanding of veridical perception in terms of a subject’s ‘openness to the layout of reality’ (McDowell 1994, 26. Quoted by Fish this volume xx), an expression that reveals his commitment to disjunctivism (if veridical perception is openness to the reality, hallucination is something else completely, which offers the mere appearance of such openness). While he finds this idea ‘completely natural and intuitive’ (2010, 245), Burge finds it rather unintuitive and confusing—nothing more than a misplaced metaphor—and is keen to translate it into his own non-disjunctivist vocabulary: an experience is ‘open to the layout of reality’ if its content is true and it thereby accurately represents the world (Burge 2005, 3. See Fish this volume xx).
quoted in Fish this volume). Because of this incommensurability—Fish argues—competing philosophical views, much like Kuhnian competing scientific paradigms, cannot do anything to convince the opponent. Any argument to this effect will inevitably rely on the assumptions associated with one paradigm, and thus seem compelling only to those who already accept them, but will fail to persuade the opponents who reject those assumptions.

While Fish sees elements of incommensurability in the philosophical debate on perception, he is also aware of the limitations of comparing philosophical disputes to Kuhn’s scientific paradigms. The most notable difference he sees lies in the mostly diachronic nature of the Kuhnian model. For Kuhn, it is only in periods of ‘crisis’ that paradigms coexist, while in times of ‘normal science’ all research happens within a single prevailing paradigm. This doesn’t seem to be the case in philosophy, where competing theories are the norm rather than an exception. This leads him to turn to the work of another philosopher of science, Lakatos, whose competing research

---

3 According to Fish, McDowell and Burge disagree on whether a philosophical account of perception should be constrained by epistemological considerations on how perception makes knowledge of the world available: McDowell’s account is primarily driven by those, while Burge doesn’t think an account of perception should be constrained by them. They also disagree, Fish claims, on what philosophy and science have to say respectively about consciousness. For McDowell a central task of a philosophical theory of perception to explain consciousness. For Burge consciousness is ‘almost surely […] not an essential feature of all perception’ (Burge 2011, 79, quoted by Fish, this volume) therefore it should not be the primary concern of the philosophical investigation of perception. As a result, McDowell and Burge have different standards for what counts as a satisfactory account of perception.
programmes bear some resemblance to Kuhn’s paradigms. For Lakatos, a research programme is successful if the successive changes to the peripheral claims of the theory made to accommodate new data yields novel predictions and some of these predictions are corroborated. Thus, for Lakatos, competing paradigms are not incommensurable, but they can be evaluated only with the benefit of hindsight.

While this seems more optimistic than suggesting that competing philosophical views of perception are incommensurable, applying Lakatos’s model of scientific progress to philosophy might prove difficult. As Fish himself notices, philosophy doesn’t always yield testable predictions in the same way that science does, and while some philosophers (Fish mentions for instance Dennett) conceive of their theories as yielding predictions, many philosophers would emphatically deny that philosophy should do such a thing. Most importantly, even if we accept that at least some philosophical theories yield predictions, ‘there may well not yet be any non-contentious way to test those predictions’ (Fish, this volume: xx). Even if there was a way to derive predictions from philosophical theories and test them, ‘philosophers themselves—happiest, as they are, in more conceptual spaces—may neither be best placed nor particularly inclined’ to do so (Fish, this volume: xx). It will be up to scientists to derive predictions from philosophical theories, test them, and report back to philosophers which one of the views they have debated is more successful. Not many philosophers would like to think that deferring to science is ultimately the only way to vindicate their philosophical view.

Whichever way you look at it, Fish’s diagnosis of the state of the philosophical debate about perception is disheartening. If we embrace the Kuhnian model, advancing the debate is structurally impossible. If we opt for Lakatos’s model, it will be practically impossible to do so, at least within the realm of philosophy alone. But we
needn’t adopt such a defeatist diagnosis. I will argue that a more promising route forward is available. If we divide the overall debate into ‘sub-debates’, we may well find that what is hindering progress is not a clash of research programmes, but a series of misunderstandings that might be avoided if philosophers were more careful about disentangling the different questions they are invested in.

Some misunderstandings originate in the way the stand-off between relationalism and representationalism is often framed in terms of a disagreement about ‘the essential metaphysical structure’ (Genone 2016), or ‘the most fundamental characterization’ (Brewer 2011, 94) of perception. This is problematic, because philosophers often have different views about precisely what is essential or fundamental to perceptual experiences. For some, this is whatever accounts for the conscious aspect of perception—that is, its phenomenal character. For others, it’s whatever explains how perception generates knowledge of the world, or justifies beliefs about it, or motivates and prompts behaviours. For yet others, a fundamental characterization is one that explains all these phenomenological, epistemological, and behavioural explanations at once. This lack of clarity offers, as Fish remarks, ample opportunities for misunderstanding and disagreement.

But we can go past the unhelpful language often used to frame the debate and avoid talking past each other by making sure we understand the restricted aim of each proposal. Interestingly, most relationalists have made it quite clear that they are primarily concerned with offering an account of the phenomenal character of perception. They think that the non-representational relation plays a fundamental role in explaining the phenomenal character of perception, which is partly constituted by
the mind-independent objects one perceives. As such, they oppose representationalism to the extent that it claims that the representational content of experience can explain the phenomenal character of experience.

In what follows, I will focus on this more confined debate about the phenomenal character of perception. Here, the scope for disagreement about which questions are worth asking is significantly reduced, and it becomes possible to see how we can advance the debate between relationalists and representationalists.

---


5 This leaves it open that relationalism (understood as an account of phenomenal character) might be compatible with allowing some other explanatory role (e.g. accounting for how perception brings about and motivates beliefs and actions) for the representational content of perception. It should also be noted that representationalism and relationalism are not the only options in this debate. Sense-data theorists claim that perception consists in a direct relation to immaterial objects called sense-data (Moore 1903; Russell 1912; Broad 1925; Price 1932; Robinson 1994; Foster 2000). Another family of theories includes views that account for the phenomenal character of perception in terms of ways of perceiving (adverbalism: see Ducasse 1942; Tye 1975, 1984; Coates 2007); for some of these views, the phenomenal character is determined by intrinsic properties of the experience (the qualia or mental paint view: see Chalmers 1996, 2004; Block 1996, 2003, 2010).
3. Perception, Hallucinations, and Bullets to Bite

One of the most effective ways of presenting what motivates the recalcitrant disagreement between relationalists and their opponents is offered by Valberg (1992).

According to Valberg, at the heart of the debate in philosophy of perception lies an antinomy between two equally reasonable attitudes towards perception. If we are ‘open to our experience’ and how it strikes us through introspection, we are led to think of experience as a relation of awareness with mind-independent objects, as per relationalism. On the other hand, however, reflection on the possibility of (certain kinds of) hallucinations compels us to deny that perception amounts to a relation to the mind-independent objects that we ordinarily take ourselves to be perceiving. For Valberg, these two attitudes form an antinomy: they are equally compelling, but mutually incompatible, ways of thinking about perception. Once this antinomy becomes manifest, one can either ‘demonstrate that the conflict in which it consists is merely apparent’ (Valberg 1992, 42) or accept it and explain why it ‘really is a conflict’ (Valberg 1992, 42). Disjunctivists go for the first option: they stick to the relationalist intuition and argue that there is a fallacy (or several) in the argument from hallucination. Opponents of disjunctivism might well agree that relationalism would be preferable, but they contend that such an option is simply not available, because the argument from hallucination offers a reductio ad absurdum of relationalism.

The argument originates from noticing that we can have hallucinations—i.e. experiences that seem to relate us to mind-independent objects in the absence of any suitable object and arguing that, because we are often unable to tell them apart from veridical perceptions, the same non-relational account that applies to hallucination must be true for perception too. From here, it has developed into many versions, roughly
corresponding to subsequent refinements in response to the relationalist’s responses to it.

Disjunctivism arises from the relationalist’s attempt to reconcile their view with the possibility of hallucinations: there is no need—they claim—to accept that perception doesn’t involve a relation to a mind-independent object just because hallucinations don’t, and sometimes we take one mistake for the other (see Austin 1962; Hinton 1967). Proponents of the argument from hallucination bring to our attention new counterexamples and thought experiments to challenge this move. Recent discussions have focused on the following thought experiment. Suppose you see a lemon in front of you. Then the lemon is removed, while a scientist activates, through transcranial stimulation, exactly the same areas of the brain that were firing when you were seeing the lemon. You now have a hallucinatory experience as of a lemon which is indistinguishable from the veridical perception of the lemon, but in this case the experience is brought about only by the firing of certain neurons, while there is no lemon in front of you. It is more difficult for the relationalist to deny that a hallucination which is both indistinguishable from a given perception and caused by the same brain activity deserves the same account as a veridical perception.

The back and forth of this debate is characteristic of what Casati (2011) calls ‘conceptual negotiation’, a practice that he deems to be at the heart of any philosophical activity. According to Casati, the main aim of conceptual negotiation is to evaluate costs and benefits of competing views, by unpacking implicit assumptions in descriptions, theories, and arguments, and exploring their unexpected consequences. The need for conceptual negotiation arises—Casati claims—when we encounter a tension between competing intuitions. Such a tension often follows the discovery of new facts, either through scientific discoveries, changes in our society, or simply by
new facts being brought to our attention. In our case, the tension arises from the realization that hallucinations are possible, and that this challenges the relationalist intuition.

Valberg’s antinomy suggests that taking any position in this debate comes at a cost. Part of the job of conceptual negotiation is to assess how big of a bullet one has to bite when embracing a view. However, as Casati points out, the final decision as to which bullet to bite is extra-philosophical. At some point, philosophy’s job of unpacking the respective costs of competing views comes to an end, and people will still disagree on whether a certain price is worth paying. Here Casati echoes Lewis:

Whether or not it would be nice to knock disagreeing philosophers down by sheer force of argument, it cannot be done. Philosophical theories are never refuted conclusively. (Or hardly ever. Gödel and Gettier may have done it.) The theory survives its refutation—at a price. Argle has said what we accomplish in philosophical argument: we measure the price. Perhaps that is something we can settle more or less conclusively. But when all is said and done, and all the tricky arguments and distinctions and counterexamples have been discovered, presumably we will still face the question which prices are worth paying, which theories are on balance credible, which are the unacceptably counterintuitive consequences and which are the acceptably counterintuitive ones. On this question we may still differ. And if all is indeed said and done, there will be no hope of discovering still further arguments to settle our differences. (Lewis 1983, x–xi)

Has the debate on relationalism reached the point where, as Lewis says, all has been said and done and there’s no hope to settle our differences? I will argue that this is not
the case. There is still a lot to do to fully assess the costs and benefits of relationalism. The deep-seated scepticism with which disjunctivism is often met suggests that some hidden assumption is at play in the debate. In the next section I will briefly outline what this implicit assumption is and how it can be used to mount a further argument against relationalism, which promises to highlight the true cost of embracing it. The assumption in question—implicit in the very notion of phenomenal character—is the idea that the phenomenal character of experience is superficial with respect to introspection. For this reason, I dub the ensuing argument ‘the argument from superficiality’.

Like the argument from hallucination, the argument from superficiality focuses on the possibility of hallucinations and takes the form of a *reductio ad absurdum*. If Lewis is right, we should not expect any argument (including *reductiones*) to conclusively refute a theory (despite what some proponents of the argument from hallucination purport). What *reductiones* show is that a theory is incompatible with certain allegedly compelling claims. However, *reductiones* can be more or less persuasive—their persuasive force being conditional to finding the premises of the argument more convincing than the claim to be rejected, in this case relationalism.

As mentioned at the beginning of this paper, relationalists contend that the premises used in the various versions of the argument from hallucination beg the question against naive realism, and are less compelling than the intuition relationalism seeks to accommodate—if not downright false. I will argue that the argument from superficiality relies on premises that relationalists cannot easily dismiss as question begging and that are prima facie at least as compelling as the relationalist intuitions.

With this argument in place, we can then examine whether relationalists can accommodate its premises and at what costs, and what would they lose if they fail or refuse to accommodate them. In other words, by considering what relationalists can say
in response to the argument, we can better understand the true costs of accepting relationalism. Depending on the results, this may either reveal that the cost to be paid is higher than relationalists had so far realized, or finally appease some of the worries of the opponents.

Because this paper focuses on the scope and methodology of the philosophy of perception, I won’t attempt any detailed evaluation of the argument or discussion of what relationalists could say in response to it. The aim of this paper is not to establish whether the merits of relationalism exceed its costs. Rather, it is to show that there is scope for progress in the debate, and in what direction we should move in order to achieve it.

4. Phenomenal Character and Introspection

The recent debate has focused on versions of the argument from hallucination that make use of the ‘same cause, same effect’ principle or the principle of local supervenience and pivot around a thought experiment presented above, featuring hallucinations that are brought about by the same brain activity involved in perception.

However, I think we can learn something important and advance the debate if we temporarily leave aside the later instantiations of the argument from hallucination and we focus on understanding what lies at the root of the original argument from hallucination and the Indistinguishability Principle (the idea that if two experiences are introspectively indistinguishable through introspection, they must have the same phenomenal character).

Proponents of the Indistinguishability Principle are motivated by certain assumptions about introspection and its link to the phenomenal character of perception.
Many think that, while introspection is generally fallible, it is infallible when it comes to the access it offers to the phenomenal character of experience (see Farkas 2006; Gertler 2012; Giustina and Kriegel 2017).

This claim of restricted infallibility is highly contentious and vehemently rejected by relationalists. While I share these reservations towards the Indistinguishability Principle and the idea of local introspective infallibility, I think that the thought that there is a special link between phenomenal character and introspection which disjunctivism makes difficult to accommodate deserves further examination. The problem is that this thought has so far been mischaracterized by the opponents of relationalism, steering the debate in an unfruitful direction. In what follows I will seek to offer a more accurate characterization of this link and examine what constraints it puts on any account of phenomenal character and how this may put pressure on disjunctivism.

‘Phenomenal character’ is a term of art introduced to focus on what it is like to have an experience, among all the other aspects of perception we may be interested in (its functional role, its object, its causal history, its representational content, if any). In turn, what it is like to have an experience is often specified in terms of what is manifest to the experiencer through introspection.

Thus, the role the term ‘phenomenal character’ is designed to fulfil in the current philosophical jargon implies that it is inextricably intertwined with the notion of

---

6 See Schwitzgebel 2008, 2012; Churchland 1984; Bayne and Spener 2010; Shoemaker 1996 for various arguments (coming from very different perspectives, none of which presupposes relationalism) against this restricted version of introspective infallibility.

phenomenal character. This is not to say that the phenomenal character of experience
metaphysically depends on introspection: that would be a controversial commitment
about its nature that most philosophers would reject.\footnote{Some see higher-order theories of consciousness (Lycan 1987, 2001; Rosenthal 2000, 2005a, 2005b; Carruthers 2004) as committed to this metaphysical dependence. However, most accounts construe introspection as an activity that involves attention and the deployment of higher-order concepts, they maintain that phenomenal consciousness does not require these higher-order capacities (as creatures devoid of them, or humans that are inattentive or cognitively impaired, possess phenomenally conscious experiences).} Yet ‘direct introspective
ostension’ (Kriegel 2015, 47) seems to be the only way to fix the reference for
‘phenomenal character’.\footnote{As Block (1978) famously notes, when asked to define phenomenal consciousness, a natural response is to answer as Louis Armstrong did when asked to define jazz: ‘if you have to ask, ain’t never gonna know.’}

This explains why it is so tempting for many to accept (more or less explicitly) the
Indistinguishability Principle and dismiss disjunctivism as utterly unbelievable or
confused. If an appeal to introspection is part of the descriptor that fixes the reference
for ‘phenomenal character’, then introspection is the ultimate authority for knowing
what the phenomenal character of an experience is: whatever the phenomenal character
an experience seems to have through introspection is the phenomenal character the
experience has, because there is nothing else to the phenomenal character of an
experience than what appears through introspection. Or so the reasoning goes.

But this is a non sequitur. If we want to assess whether disjunctivism is a viable
option, we need to tease apart which constraints are simply part of the function we have
assigned to the notion of ‘phenomenal character’ from further controversial

---

\footnote{Some see higher-order theories of consciousness (Lycan 1987, 2001; Rosenthal 2000, 2005a, 2005b; Carruthers 2004) as committed to this metaphysical dependence. However, most accounts construe introspection as an activity that involves attention and the deployment of higher-order concepts, they maintain that phenomenal consciousness does not require these higher-order capacities (as creatures devoid of them, or humans that are inattentive or cognitively impaired, possess phenomenally conscious experiences).}

\footnote{As Block (1978) famously notes, when asked to define phenomenal consciousness, a natural response is to answer as Louis Armstrong did when asked to define jazz: ‘if you have to ask, ain’t never gonna know.’}
commitments about consciousness, perception, introspection, and so on, which might be surreptitiously built into it. Offering an informative yet neutral characterization of the link between introspection and phenomenal character is a delicate exercise of conceptual negotiation that I can’t hope to settle once and for all here. But I will propose an initial approximation and open up the negotiations.

I find that the best way to begin is to focus on the function the notion of phenomenal character is designed to fulfil (that of singling out certain aspects of experience from others, on the basis of their privileged relation to introspection) and to compare it with another notion, which plays a similar role—that of contrasting a definite class of properties from others on the basis of their privileged relation to a source of information. The notion I have in mind is that of ‘observational properties’, used to refer to those properties (such as colour, taste, odours, sound, shape, size) that are thought to bear a special relation to perception. This is in contrast to properties which can also be seen but do not seem to be linked to perception in the same special way, such as natural kinds (being a lemon), artefactual kinds (being a chair), or causal-historical properties (being fresh).\(^\text{10}\)

---

\(^{10}\) There isn’t consensus about what properties count as observational (for instance, some would claim that colours count but shapes don’t), or even whether there is any property which is truly observational. This is irrelevant for my present purposes. All is needed for the present purposes is to accept that there is a meaningful distinction to be drawn between observational and non-observational properties and if one were to maintain that either colour, smell, shape, or size were observational, the best way to account for observationality is in terms of their being superficial with respect to perception.
Similarly to what happens with the notion of phenomenal character, often philosophers either fail to specify the nature of observational properties’ relation to perception or give clearly inadequate characterizations, yet assume as obvious the

I also don’t need to take a stance on whether we (directly) perceive natural kinds, artefactual kinds, and so on, often referred to as higher-level properties. As I understand it, the question as to what makes certain properties observational is more fundamental than the question as to whether higher-level properties can be directly perceived. We would not be able to ask that question if we didn’t have the intuition that some properties (what I refer to here as observational properties and what are referred to as ‘low-level properties’ in the context of the debate on what can be directly perceived) are more intimately linked to perception than others—as testified by the fact that we take for granted that they are directly perceived while we consider it an open question whether higher-level properties are directly perceived.

We can understand those who deny that we can perceive higher-level properties (e.g. Byrne 2009; Price 2009; Brogaard 2013) as offering a simple explanation of the intuitive distinction between observational and higher-level properties: for them what distinguishes observational properties is that they can feature in perception. However, even proponents of the claim that we can perceive higher-level properties (e.g. Bayne 2009; Siegel 2006, 2009; Nanay 2011a, 2011b) can maintain that there is a distinction: what they will need is an alternative explanation for this distinction which is not simply in terms of what can and cannot be directly perceived. What I offer here can be seen as an attempt to clarify what is special about observational properties which remains neutral with respect to the debate on what type of properties feature in perception. Thanks to Heather Logue for pointing out the need to clarify this.
existence of a special connection of some sort with perception. However, the notion of ‘observational properties’ is less elusive than that of ‘phenomenal character’, and we can rely on the fact that the relation between perception and observational properties has been more broadly examined. Thus, we can hope to learn some lessons about the phenomenal character of perception by first looking at the relatively more manageable literature on observational properties.

Peacocke (1983) notices that observational and non-observational concepts (such as, for instance, concepts of natural kind properties) have a specific pattern of epistemic possibilities. It is possible for something to look like a tomato from all the different angles it may be seen, by a subject whose perceptual mechanism works properly and has the relevant concept of being a tomato, and yet it not be a tomato, but, say, a fake plastic tomato or even a synthetic replica of tomato produced by Kraft laboratories. On the other hand, something cannot seem to a subject with good sight and no relevant cognitive deficiency to be consistently yellow across time, under changing lighting conditions.

---

11 Here is an example of inadequate characterization: ‘Properties are observational in so far as they are presented or represented in perceptual experience. Properties are represented in experience by means of appearances. Thus we can say that observational properties are those properties that objects can appear to have when we perceive them’ (Langsam 2000, 69). This characterization is inadequate because it is too inclusive: objects can appear to have all sorts of properties, not just observational properties such as colours and tastes. An object can appear to be a lemon, stale, or lonely. We can see his characterization of observationality as assuming that we can’t perceive higher-level properties. Or else we lose the ability to distinguish between observational and non-observational properties, because any property that we can perceive would count as observational.
conditions which include optimal ones, without actually being yellow (see Peacocke 1983, 99).12

If we fail to see that an object is a fake tomato rather than a real tomato, it might be our perception’s fault (maybe our vision is blurry and we can’t tell that the plastic tomato out of a child’s grocery shop set is a fake tomato). But it’s also possible that the fault doesn’t lie with perception: in the case of the synthetic fake tomato, for instance, there is no perceivable difference between the real and the fake tomato. The difference is hidden to perception: it lies in their respective chemical composition. It can only be discovered by non-perceptual means of investigation. In the case of redness, if something appears red but is not red, this can only be because something is wrong with either the conditions of observation or the perceptual mechanism, not because the difference between real red and ‘fake red’ is hidden to perception. An effective metaphor often used to characterize this distinction is in terms of superficiality. A property like redness is superficial in the sense that we can know that it’s instantiated by something on the basis of perception alone.

5. The Superficiality Constraint

This metaphor of superficiality and the related notion of epistemic possibilities are useful for understanding what constraint applies to an account of the relation between phenomenal character and introspection. Remember that the task at hand is to characterize the intuitive link between introspection and phenomenal character in a way which is informative, yet neutral with respect to controversial commitments.

12 Peacocke focuses on observational concepts, but it is plausible to assume that observational concepts refer to classes of properties to which the epistemic constraint he identifies applies.
Philosophers have tried to characterize it with various notions, such as infallibility, certainty, incorrigibility, or luminosity. But these are all controversial.

The idea that the phenomenal character is superficial with respect to introspection in the same way observational properties are superficial with respect to perception is weaker than any of these alternatives, yet it is informative, inasmuch as it grasps the intuitive distinction between, on one hand, phenomenal character and, on the other hand, other aspects of our mental life that may be introspected, but do not bear the same deep link to introspection.

We often know through introspection alone what mental kinds we instantiate: we can reflect on our experience and judge that we are undergoing a perceptual experience, or an episode of recollection, or that we feel certain emotions. However, in some cases, whether we are in one kind of mental state as opposed to another is something that goes beyond the scope of introspection. For example, introspection doesn’t offer any clue as to whether the experience is a veridical perception or a hallucination that is subjectively indistinguishable from it. The only way to adjudicate that is to assess how the experience is brought about.

On the other hand, it doesn’t seem possible for a difference in phenomenal character to go beyond the scope of what is accessible through introspection and only be discoverable through non-introspective enquiries. To put it in terms similar to those used by Peacocke with respect to observational properties, while introspective mistakes are possible (for instance because one doesn’t pay enough attention or because one’s introspective capacities don’t work properly), the pattern of possible mistakes is constrained by the following rule, which I call ‘the Superficiality Constraint’:

(SC): It is not possible that an experience seems through introspection to have a certain type of phenomenal character, while it doesn’t actually have that
phenomenal character, and there is no disabling or interfering condition in place that prevents one from introspectively realizing that the experience doesn’t, in fact, have that type of phenomenal character.

The Superficiality Constraint is a very minimal claim. It is an epistemic constraint that merely restricts the pattern of possible mistakes about the phenomenal character of experience and remains neutral with respect to various commitments about the nature of introspection, phenomenal character, perception, and consciousness. As such, it is prima facie compatible with disjunctivism and cannot be said to beg the question against it.

To appreciate how the Superficiality Constraint is indeed very modest and neutral, it is useful to contrast it with two other claims in its vicinity: (a) revelation and (b) luminosity.

(a) Revelation was introduced by Johnston as an alleged feature of our commonsense understanding of colour:

(R): ‘the essential nature of (for example) canary yellow is fully revealed by visual experiences as of canary yellow things.’ (Johnston 1992, 138)

It might be tempting to think that the Superficiality Constraint is akin to Revelation for the phenomenal character of experience. But this would be a mistake. While revelation is about the essential nature of a colour, that which makes a colour the property it is and not something else, the Superficiality Constraint is not about essential properties: it remains silent with respect to the nature of phenomenal characters. It is not about what makes the phenomenal character what it is rather than something else: it more modestly dictates the conditions under which an experience may seem to have a phenomenal character it does not have (whatever makes the phenomenal character what it is).
(b) Luminosity is a thesis about mental occurrences that Williamson (2000) defines as follows:

‘(L) For every case α, if in α [the mental occurrence] C obtains, then in α one is in a position to know that C obtains’ (Williamson 2000: 95).

While mental occurrences such as pain and bodily sensations are often thought to be luminous, Williamson famously argues that no mental state is luminous. By allowing that there might be interfering conditions that impair one’s introspective capacities, the Superficiality Constraint is weaker than luminosity. If disabling or interfering conditions are in place, one may not be in a position to know that an experience with a certain phenomenal character obtains, although it remains true that the phenomenal character is superficial, because the fact that a certain phenomenal character is present could in principle be detected (where the interfering or disabling conditions are not in place) through introspection alone.

Once we have distinguished the Superficiality Constraint from Revelation and Luminosity, we can more easily see in what respects the Superficiality Constraint is minimally committal and, as such, ‘relationalist-friendly’:

(a) Unlike Luminosity—and unlike the Indistinguishability Principle, the principle of Same Cause, Same Effect and Local Supervenience used in the arguments from hallucination—the Superficiality Constraint is compatible with the disjunctivist commitment: due to the possibility of interfering and disabling conditions, it allows for hallucinations to seem to have a (relationalist) phenomenal character that they in fact lack.

(b) Unlike Revelation, the Superficiality Constraint doesn’t commit to the idea that introspection fully reveals the nature of a phenomenal character. This fits very well with what relationalists are likely to say about the phenomenal character
of both perception and hallucination. They clearly want to avoid the idea that the nature of the phenomenal character of hallucination is fully revealed by introspection. And while they are likely to think that there is some introspective support for their claims about the phenomenal character of veridical perception, they are likely to think that additional extra-introspective considerations are required to support their view.\footnote{An example of argument which heavily relies on extra-introspective considerations is the argument from transparency in Martin (2002).}

As we have seen in the previous section, the idea of a special link between introspection and phenomenal character is part and parcel of the notion of phenomenal character. Denying such a link would be tantamount to refusing to use ‘phenomenal character’ in the way it is arguably designed to be used when introduced—i.e. to refer to that aspect of experience that bears a special connection to introspection. If one were to deny the existence of such a connection, one would not simply embrace a controversial explanation of what the phenomenal character of perception is, one would use the locution ‘phenomenal character’ in a way that changes the explanandum, and fails to refer to what everybody else in the debate means by it.

The idea that some special link exists between introspection and the phenomenal character of perception is what has more or less implicitly motivated the widespread resistance towards relationalism. However, opponents so far have failed to identify the specific nature of this link, which hindered any progress in assessing whether or not there is some tension between how we think of phenomenal character and disjunctivism. With the notion of the Superficiality Constraint on hand, we can now hope to make some progress in this direction.
6. Advancing the Debate with the Argument from Superficiality

The Superficiality Constraint is something that relationalists can prima facie accept, because, contrary to the Indistinguishability Principle, it allows for introspective mistakes about phenomenal character.

Disjunctivists can explain indistinguishable hallucinations without violating the Superficiality Constraint. They can say that something interferes with or disables the subject’s introspective capacities. For instance, the subject might not be attentive enough, or their brain might not work properly. However, it becomes problematic when it comes to accommodating the type of hallucinations introduced in the thought experiment presented in section 3—hallucinations that are indistinguishable from a perception and brought about by the same neural activity. In this thought experiment, ex hypothesi, there is no interfering or disabling condition to which disjunctivists can appeal to explain why hallucination seems to have a relationalist phenomenal character (one that relationally puts us in contact with objects and is constituted by them). It is

---

14 What counts as an interfering or disabling condition for introspection depends in part on how we conceive of introspection. The more pluralist one’s account of introspection is (where multiple processes and capacities are involved, rather than a single mechanism), the more opportunity for disabling and interfering conditions there are. Without committing to any specific account of introspection, we can say that, since attention is certainly a crucial component of introspection, any factor that may interfere with or disable one’s introspective attention counts as a potential disabling or interfering condition. This might be the presence of competing stimuli that divert attention (such as overwhelming emotions), altered brain states due to illness or drug use, or impaired cognitive capacities.
part of the set-up of the thought experiment that the brain activity is identical in the two situations, that the subject is employing exactly the same cognitive capacities, is exerting the same degree of attention, and no other occurrence in their stream of consciousness interferes with one’s introspective capacities. The only difference between a perception and the corresponding perfect hallucination is that the object is present in the former case and absent in the latter. But, in order for a condition to interfere with introspection, it must be something that is potentially relevant to introspection. It is not clear how the mere absence of an object can be relevant to one’s introspective capacities. In this respect, introspection is different from perception, and the analogy between phenomenal character and observational properties breaks down. Perception is sensitive to objects in the world. For this reason, facts about the world (e.g. the presence of an occluding object, or of certain lighting conditions) count as disabling or interfering conditions for seeing that an object has a certain property. But introspection tracks mental occurrences, so it is natural to think that only facts about one’s psychological state and the neurological mechanisms that underlie them can count as disabling and interfering conditions.

Thus, disjunctivism seems committed to denying that the phenomenal character of these types of hallucination is subject to the Superficiality Constraint. This makes the position untenable, as the way we use the notion of phenomenal character in the philosophical jargon requires us to accept the Superficiality Constraint. Or so the argument goes.

This argument, which I call ‘argument from superficiality’, only relies on unpacking implicit assumptions that are part and parcel of the way the notion of ‘phenomenal character’ is employed in the philosophical debate. Therefore, it is more
compelling than the various versions of argument from hallucination, whose key assumptions are deemed question-begging by relationalists.

Of course there are a number of things disjunctivists can say in response to this argument. I don’t claim to have offered a conclusive argument against disjunctivism: as we have seen, theories survive their refutations. On the contrary, having identified and spelt out the source of the (as of yet implicit) resistance towards the view can give disjunctivists the tools to address these worries and win over some opponents, especially those who might appreciate the benefits of relationalism but are worried by its implications.

The first option open to relationalists is of course to reject the Superficiality Constraint. They might grant that the Superficiality Constraint is constitutive of the way most people use the notion of phenomenal character, but they use it differently. What they mean by ‘phenomenal character’ is the property of relating the subject to certain objects in the environment. If this commits them to violating the Superficiality Constraint, then tant pis for superficiality: for them the phenomenal character of an experience goes beyond what is accessible through introspection.

This, however, is not an easy solution. I have argued that the Superficiality Constraint doesn’t follow from a particular view of phenomenal character, but rather simply from the way we assign the reference of ‘phenomenal character’. There can’t be a disagreement about the nature of phenomenal character if we don’t agree on what we talk about when we talk about phenomenal character. And if my analysis in the previous section was correct, they couldn’t engage in any debate with their opponents about the nature of phenomenal character if they were to deny the Superficiality Constraint: they would just use the same word to refer to something else entirely. If relationalists seriously want to pursue this strategy, they cannot simply reject the superficiality
constraint, but they have to offer an alternative understanding of the relation between phenomenal character and introspection, one that accommodates our intuitions equally well as the Superficiality Constraint.

Another option is to accept the Superficiality Constraint but avoid the problem it poses for hallucinations by arguing that hallucinations don’t have any phenomenal character. This is a strategy already adopted by some relationalists (Fish 2009; Dokic and Martin 2012; Logue 2012) in response to the traditional argument from hallucination. Eliminativism about the phenomenal character of hallucination faces several objections, the most compelling of which, in my opinion, is the following. In this view, the apparent sensory conscious phenomenology is due to cognitive phenomenology (see Fish 2009, 98–9, n. 19). But it is not clear that one can accurately account for the (apparent, according to eliminativists) sensory phenomenology of hallucinations in terms of cognitive phenomenology (see Vega-Encabo 2010, 190).\(^\text{15}\) Additionally, this strategy won’t work, because as long as we concede that it falsely seems to a hallucinating subject that they are in a state with phenomenal character, one would still violate the superficiality constraint.\(^\text{16}\)

What seems to me the most promising option for relationalists is to show that their view is compatible with the Superficiality Constraint. They could do so by insisting that in the hallucinatory case there is something that interferes with one’s introspective

\(^{15}\) As Pautz (2013) notices, the viability of this proposal depends on how satisfactorily one can counter traditional arguments against Rosenthal’s ‘higher-order thought theory’ of consciousness, on which Fish’s idea that consciousness of hallucination is a by-product of cognitive attitudes heavily relies.

\(^{16}\) Thanks to Heather Logue for pointing this out to me.
capacities. The relationalist could insist that, since the object constitutes the phenomenal character of one’s veridical perception, and this is something one can introspect, the lack of the object that seems to constitute the phenomenal character of hallucination is relevant to introspection.

Interestingly, some relationalists have proposed theories of introspection that could support such a claim. The most developed theory of introspection of this kind I am aware of is proposed by Soteriou (2013, ch. 8). He argues that introspection works by focusing one’s attention on the object of perception by focusing on how they subjectively strike us: introspection is then structurally parasitical on perception. So the mere fact that in introspection one fails to be in the perceptual relation to the world one takes herself to be is enough to impair one’s ability to introspect properly and access the real phenomenal character of perfect hallucination.  

There are certainly more things relationalists could say to try and accommodate the Superficiality Constraint or to replace it with a less problematic specification of the relation between phenomenal character and introspection. Exploring their options will in any case lead to a better understanding of their commitments and will offer new elements to help us decide whether the costs of accepting disjunctivism exceed its benefits. What’s noteworthy is that trying to deny the Superficiality Constraint and trying to accommodate it seem to both depend on a clarification of what relationalists take introspection of the phenomenal character to be. This suggests that the main disagreement between relationalists and their opponents may lie in how they

---

17 It seems to me that a view of introspection along these lines plays a central role in much of the arguments for disjunctivism in Martin (2006). A similar theory has been proposed by Logue (2012), who pairs it with a commitment to eliminativism about the phenomenal character of hallucination.
respectively understand introspection and that the focus of the debate should shift in that direction.

Fish identifies a lack of progress in the debate, and attributes it to a structural (and, as such, insurmountable) problem: the incommensurability of the paradigms adopted by proponents of alternative views. Instead, I have argued that the current stall is due to specific shortcomings—in particular, failing to acknowledge, from both sides of the debate, the deep-seated intuition that underlies the rejection of disjunctivism (i.e. the Superficiality Constraint). This diagnosis of the debate is preferable because it makes space for the possibility of adjudicating the debate and offers some suggestion as to what could tip the scale one way or another.\(^{18}\)

References


\(^{18}\) I would like to thank Hemdat Lerman, Matthew Soteriou, Maarten Steenhagen, Daniel Vanello, and the editors Heather Logue and Louise Richardson for their insightful and helpful comments on drafts of this paper. I also benefited greatly from feedback by many people, when I presented some of the material discussed here in various occasions. Among those, I would like to especially thank Bill Brewer, Chiara Brozzo, Dan Cavedon-Taylor, Mark Kalderon, Keith Wilson, and Hong Yu Wong.


Fish, William C. this volume. ‘Perceptual Paradigms’. In Heather Logue & Louise Richardson (eds.), *Purpose and Procedure in Philosophy of Perception*. Oxford: Oxford University Press.


