Getting Acquainted with Naïve Realism: Critical notice of Perception, Hallucination, and Illusion

Heather Logue[1]

William Fish's Perception, Hallucination, and Illusion is one of the clearest and most thorough defenses of the view known as *Naïve Realism* offered to date. Naïve Realism is a claim about veridical perceptual experiences, i.e., experiences in which a subject perceives things in her environment that are as they appear to her to be.[2]

M.G.F. Martin, one of Naïve Realism's prominent proponents, characterizes the view as holding that the mind-independent objects of veridical perception "shape the contours of the subject's conscious experience" (2004: 65-8). Fish suggests that this talk of shaping should be understood in a *constitutive* sense: just as hills shape the contours of a landscape by constituting the contours of that landscape, the mind-independent objects of one's veridical perception shape the contours of one's experience by actually constituting those contours. John Campbell, another of Naïve Realism's well-known advocates, explicitly embraces this idea:

> …the phenomenal character of your experiences, as you look around the room, is constituted by the actual layout of the room itself: which particular objects are there, their intrinsic properties, such as color and shape, and how they are arranged in relation to one another and to you. (2002: 116)

One might find this idea difficult to get one's head around. What exactly does it mean to say that the phenomenal character of my current veridical experience, i.e., "what it is like" for me to have it, is *constituted* by a *banana*?

Fish offers a helpful elaboration of this core Naïve Realist claim. He suggests that the phenomenal character of veridical perception consists in the subject's being *acquainted* with mind-independent objects in one's environment and some of the properties they instantiate. Fish uses 'acquaintance' to pick out "…an irreducible mental relation that the subject can only stand in to objects that exist and features that are instantiated in the part of the environment at which the subject is looking" (p. 14).[3] By 'acquaintance', Fish doesn't mean anything particularly mysterious; as I understand him, he might as well have used the term 'perception'.[4] The basic idea is that a veridical perception has its phenomenal character in virtue of the obtaining of the acquaintance relation between the subject and the mind-independent objects of experience—for example, my veridical perception of the banana on my desk has its phenomenal

---

[1] Thanks to Alex Byrne and Susanna Siegel for helpful comments.

[2] Fish limits his task to defending a version of this claim restricted to *visual* experiences. Also, I should note that although the description of veridical experience just provided is adequate for our purposes, it isn't quite right (as it also applies to a certain kind of illusion—see fn. 5 below).

[3] All page references are to Fish's book unless otherwise noted.

[4] Presumably, he chose 'acquaintance' to circumvent distracting objections from those who hold that 'perception' picks out a relation that can be analyzed in further terms (e.g., in terms of causation). There's a substantive disagreement here, but Fish has other fish to fry in this book.

character in virtue of my being acquainted with the banana and some of its properties (e.g., yellowness, crescent-shapedness).

Fish further clarifies this claim by comparing Naïve Realism with one of its main rivals: Representationalism. According to the Representationalist, a veridical perception has its phenomenal character primarily in virtue of the subject *perceptually representing her environment as being a certain way* (e.g., as containing a yellow, crescent-shaped banana). By contrast, the Naïve Realist thinks that a veridical perception has its phenomenal character primarily in virtue of the subject being *acquainted with a certain tract of her environment* (e.g., one containing a yellow, crescent-shaped banana). In short, the dispute is about whether what it's like to have an experience is fundamentally a matter of perceptual representation or acquaintance.

It is often supposed that Representationalism enjoys a dialectical advantage over Naïve Realism, as it has a relatively unproblematic account of *non*-veridical experiences at its disposal (viz., illusions and hallucinations). A total hallucination is an experience in which the subject fails to perceive anything in her environment (e.g., an experience had by a "brain in a vat"), and an illusion is an experience in which the subject perceives something in her environment that isn't as it appears to her to be (e.g., an experience in which a yellow banana looks green).[5]

Note that while one cannot be acquainted with something that isn't there, one *can* perceptually represent a false proposition. So the Representationalist can say that the phenomenal character of a non-veridical experience consists in the subject perceptually representing her environment as being a way it is not. By contrast, Naïve Realism cannot be straightforwardly extended to non-veridical experiences. Total hallucinations don't involve acquaintance with mind-independent entities at all, and the acquaintance involved in an illusory experience doesn't obviously capture its phenomenal character (it's not immediately obvious how acquaintance with a *yellow* banana can account for the fact that it looks *green* to the subject). Thus, the Naïve Realist must embrace disjunctivism about perceptual experience, which is basically the view that hallucinations (at least) are of "a different fundamental kind" than veridical perceptions (p. 36).[6]

In chapters 1 and 2 of his book, Fish describes the current state of play with respect to Naïve Realism. In addition to clarifying what he takes to be the core Naïve Realist thesis and discussing the broad outlines of how the Naïve Realist can account for non-veridical perceptual experiences, he also canvasses the motivations that have been offered for the view. Taking off from this starting point, in chapters 3 through 6 Fish develops his own Naïve Realist theory of perceptual experience. Chapter 3 is devoted to

---

[5] This characterization of illusion isn't quite accurate, as it doesn't capture a sort of experience known as *veridical illusion*: roughly, an experience in which one perceives things in one's environment that are as they appear to be, but is still defective in some respect (e.g., the experience described in Johnston 2006: 272-3). For our purposes, however, the characterization of illusion in the main text will suffice.

[6] *Pace* Mark Johnston, a Naïve Realist who holds that *all* sorts of experiences involve acquaintance with a "sensible profile": i.e., "a complex, partly qualitative and partly relational property, which exhausts the way the particular scene before your eyes is if your present experience is veridical" (2004: 134). On this view, subjects of non-veridical experiences are acquainted with *uninstantiated* sensible profiles; thus, this proposal is only as plausible as the claim that we can perceive uninstantiated properties and relations.

the elaboration and defense of the Naïve Realist account of veridical perception outlined in the earlier chapters. Chapters 4 and 5 deal with hallucination. In chapter 4, Fish articulates his own theory of hallucination in light of the shortcomings of previous accounts offered by Naïve Realists; chapter 5 attempts to undermine a primary source of resistance to this theory. Finally, in chapter 6, Fish develops a Naïve Realism-friendly account of illusion.

This book contains an abundance of thought-provoking material. So much, in fact, that this discussion will barely make it through two-thirds of the way through the book's title. First, I'll briefly touch on perception by discussing a motivation Fish offers for Naïve Realism, which strikes me as the most promising reason to be a Naïve Realist thus far proposed. The rest is devoted to a critique of Fish's account of hallucination. Alas, discussion of Fish's intriguing account of illusion will have to be postponed.

1. Perception

One of the aims of the first chapter is to give the reader reasons to take Naïve Realism seriously. To this end, Fish summarizes the motivations that have been offered for the view.

The first one he discusses is suggested by the view's name: the idea is that Naïve Realism is "the theory of the common [or philosophically naïve] man" (p. 18). Fish is skeptical of this motivation; for one thing, even if it's true that Naïve Realism is the folk theory of veridical perception, it's not obvious why this should count in its favor (p. 19). However, Fish is more sympathetic to a related motivation offered by Benj Hellie—viz., that Naïve Realism is supported by *expert* reflection on the phenomenology of veridical perception (pp. 19-20). Fish also briefly considers arguments for Naïve Realism from the claim that it refutes skepticism about the external world (see McDowell 1986, 2008, and Johnston 2006), and from the claim that it provides the only adequate account of our capacity to think about particular objects (see Campbell 2002).[7] Fish isn't entirely convinced that these arguments for Naïve Realism succeed, but he thinks that they are promising enough to warrant taking the view seriously.

I don't have the space to argue for this here, but in my view *none* of these arguments succeeds in motivating Naïve Realism over Representationalism. However, I think Fish has put his finger on another argument that might be able to do just that.

In Chapter 3, Fish suggests that Naïve Realism can help close what has been called "the explanatory gap". Fish sums up the problem that this gap poses with a passage by Joseph Levine:

> Let's call the physical story for seeing red 'R' and the physical story for seeing green 'G.'…When we consider the qualitative [i.e., phenomenal] character of our visual experiences when looking at ripe McIntosh apples, as opposed to looking at ripe cucumbers, the difference is not explained by appeal to G and R. For R doesn't really explain why I have the one kind of qualitative experience—the kind I have when looking at ripe McIntosh apples—and not the other. (quoted on p. 75)

[7] But see Byrne and Logue 2008: 65-8 for an argument that McDowell shouldn't be interpreted as endorsing a theory about the metaphysics of perceptual experience.

The question at issue is basically this: why is it like *this* to see ripe McIntosh apples (demonstrating the phenomenal character of my experience of a ripe McIntosh apple), rather than like *that* (demonstrating the phenomenal character of my experience of a ripe cucumber)? According to Levine, if we're hoping to find the answer by appeal to R and G, we're setting ourselves up for disappointment. After all, it's conceivable that evolution could have worked out such that R is the physical story for seeing *green* and G is the physical story for seeing *red*.

Fish suggests that "…the reason Levine cannot find the answer he is looking for is because he is searching in the wrong place" (p. 75). On Naïve Realism, a veridical experience's phenomenal character consists in the subject's acquaintance with objects in his environment and some of their properties. Thus, if Naïve Realism is true, we can say that the reason why *this* (demonstrating one aspect of the phenomenal character of my experience of a ripe McIntosh apple) is what it's like to see red things is that the relevant aspect of my experience's phenomenal character *just is* my being acquainted with something *red*. Moreover, what it's like to see *green* things consists in acquaintance with *green* things, so it doesn't really make sense to ask why it *isn't* like *that* to be acquainted with things that *aren't green* (such as ripe McIntosh apples). In short, Fish proposes that the answer to "why is it like this rather than like that" questions isn't to be found at the level of physical or functional processing, but rather in terms of acquaintance with entities in one's environment. According to the Naïve Realist, there is a constitutive connection between what it is like to perceive the world and what the world we perceive is like.[8]

By contrast, the Representational View isn't poised to provide as satisfying of an answer to the question at issue. According to the Representationalist, the phenomenal character of my veridical perception of a red McIntosh apple is a matter of my perceptually representing the proposition that something is red. On this account, we at least have a connection between what it is like to see red and redness. But what reason have we to suppose that our representation of redness is any guide to what redness is *really like*? As Johnston puts the thought, "[i]f sensory awareness [and thus phenomenal character] were representational, we would inevitably face the skeptical question of how we could know that the human style of representation is not entirely idiosyncratic relative to the intrinsic natures of things" (2006: 284-5, footnote omitted). The worry is illustrated by the intelligibility of the following questions: why is perceptual representation of redness associated with *this* phenomenal character rather than a different one? Couldn't it have turned out that perceptual representation of redness goes with the phenomenal character that actually goes with perceptual representation of greenness? In short, the Representational View appears to *stipulate* a connection between redness and what it is like to see it, rather than *explaining* that connection.[9]

This motivation for Naïve Realism is still in the formative stages; Fish devotes only a couple of pages to it, and more needs to be said to clarify the details and to

---

[8] Of course, Naïve Realism solves no puzzles regarding the phenomenal character of *non-veridical* experiences. According to Fish, there aren't any such puzzles to solve in the case of hallucination because, on his view, hallucinations lack phenomenal character altogether (more on this claim below). As for illusions, this motivation for Naïve Realism is effective only insofar as the Naïve Realist can account for them in a way that answers questions of this sort.

[9] See p. 78, fn. 12 for Fish's own reservations about Representationalist attempts to co-opt this motivation.

determine whether there is a version of Representationalism that could accommodate it. Nevertheless, Fish has succeeded in giving expression to a profound idea that seems to have been lurking in the background, piquing peoples' interest in Naïve Realism despite their inability to articulate exactly what's drawing them to it.

2. Hallucination

Let us now turn to Fish's account of hallucination. He holds that endorsing Naïve Realism requires endorsing disjunctivism about perceptual experience, and a specific form of it at that—one on which the mental nature of hallucination is exhausted by the property of being indiscriminable from a particular kind of veridical perception. Moreover, Fish claims that Naïve Realism is committed to denying that hallucinations have any phenomenal character whatsoever. The road from Naïve Realism to the first claim is considerably longer than the road to the second; we'll take the longer journey first.

To briefly review the road from Naïve Realism to disjunctivism: on Fish's characterization of Naïve Realism, the phenomenal character of veridical perception consists in the subject's being acquainted with things in his environment and some of their properties. Plausibly, in specifying what the phenomenal character of an experience consists in, we've specified the experience's fundamental kind or nature. But of course, *total hallucination* does not involve the subject's being acquainted with anything in her environment. Thus, the Naïve Realist must insist that whatever the nature of hallucination is, it's fundamentally different from that of veridical perception—which is to endorse a version of disjunctivism about perceptual experience. So if Naïve Realism is to even get off the ground as a theory of perceptual experience, it has to embrace disjunctivism in some form (p. 37).[10]

However, Fish thinks that embracing disjunctivism isn't sufficient to get the Naïve Realist out of hallucination-related hot water. For, as M.G.F. Martin argues, what we say about the nature of hallucinations has implications for the nature of veridical perceptions. Martin's argument for this claim is roughly as follows: suppose we characterize the nature of hallucination as involving representation of a proposition about one's environment. Plausibly, a hallucination as of a yellow banana has a proximate neural cause; this cause would be sufficient for perceptual representation of the relevant proposition (e.g., that there is a yellow banana before one). It's also plausible that a proximate neural cause of this type could generate a *veridical perception* of a yellow banana. If this type of proximate neural cause is sufficient for representation of the relevant proposition in the case of hallucination, it would be *ad hoc* to deny that it's sufficient for representation of the relevant proposition in the case of *veridical perception* as well (Martin 2004: 52-8). Call this Martin's *causal argument*.

This result, while not strictly speaking incompatible with Naïve Realism as stated, nevertheless seems to take the wind out of its sails. For it's plausible that if veridical perception and hallucination have basic psychological commonalities, then these constitute the best explanation of anything else they have in common (e.g., generation of the same kinds of beliefs and behaviors). Since the obtaining of the acquaintance relation is unique to veridical perception, it drops out as explanatorily idle with respect to any features shared by hallucination and veridical perception. The

[10] Again, *pace* Johnston (cf. fn. 6).

upshot is that the disjunctivist's characterization of hallucination must be such that it doesn't result in the "screening off" of the obtaining of the acquaintance relation with regard to explaining key aspects of veridical perception (Martin 2004: 58-63). Call this Martin's *screening-off* argument.

Martin's response to the causal and screening-off arguments is to claim that the mental nature of hallucination is exhausted by the property of being indiscriminable from a veridical perception of a certain kind. For example, the mental nature of a hallucination as of a yellow, crescent-shaped (etc.) banana is exhausted by the property of being indiscriminable from the kind: veridical perception of a yellow, crescent-shaped (etc.) banana. Of course, a veridical perception of a yellow, crescent-shaped (etc.) banana trivially has the property of being indiscriminable from the kind: veridical perception of a yellow, crescent-shaped (etc.) banana—nothing is discriminable from a kind of which it is an instance. But this property is not one that could screen off the obtaining the acquaintance relation from explaining anything. The idea is that the explanatory power of the indiscriminability property is *parasitic* on the explanatory power of the property of being a veridical perception of a certain kind. To borrow an example from Martin, the fact that James is having an experience that is *indiscriminable* from a veridical perception of a giant spider can explain why he shrieked only if James's having a *veridical perception* of a giant spider could explain his shrieking. Thus, even though the relevant indiscriminability property is instantiated by a veridical perception, it doesn't screen off the obtaining of the acquaintance relation from explaining anything (Martin 2004: 68-70).

However, worries emerge when we try to specify exactly what the indiscriminability property is supposed to be. Martin proposes that an experience is indiscriminable from a veridical perception of a certain kind just in case the subject of the experience is *not in a position to know by reflection alone* that her experience is not a veridical perception of that kind (Martin 2006: 364). As Fish explains, this construal of the indiscriminability property faces some serious *prima facie* problems. Perhaps the most extensively discussed problem is that Martin's proposal is not well-equipped to characterize the hallucinations of "cognitively unsophisticated" creatures—creatures that cannot form beliefs, and hence aren't in a position to know anything at all. For example, a dog undergoing a hallucination as of a *yellow banana* presumably isn't in a position to know by reflection alone that its experience isn't a veridical perception of a *giant brown bear*. Although the mental nature of the dog's hallucination surely isn't exhausted by *this* indiscriminability property, Martin's account lacks the resources to explain why this is so (see Siegel 2004, 2008 for extended discussions of this sort of objection).

Fish endorses Martin's general strategy for dealing with hallucinations, but implements it in a way that avoids the worries raised for Martin's account. Fish agrees that the mental nature of hallucination is exhausted by an indiscriminability property, but he gives a different gloss on the notion of indiscriminability involved. Essentially, he suggests that we should characterize it not in terms of relatively advanced cognitive *capacities*, but rather in terms of cognitive *effects*. In particular, an experience is indiscriminable from a certain kind of veridical perception iff it produces the same cognitive effects as that kind of veridical perception would have produced, holding the subject's other mental states fixed (pp. 93-5). Importantly, the cognitive effects include not just beliefs but also behaviors. This is how Fish's proposal sidesteps the "dog problem"— the dog would have behaved in the same way had he been veridically perceiving a yellow banana (holding fixed his other mental states, e.g., his desires). But

the dog would have behaved differently had he been veridically perceiving a giant brown bear, so the dog's experience counts as discriminable from such a veridical perception on Fish's characterization of indiscriminability (pp. 101-2).

According to Fish, it's not enough for the Naïve Realist to account for hallucination in terms of indiscriminability. He must also deny that hallucinations have phenomenal character at all—he must deny that there is *anything it is like* to undergo a hallucination. Fish's argument for this counterintuitive claim is surprisingly simple (see pp. 41-2):

> 1. Phenomenal character consists in acquaintance with mind-independent objects and properties instantiated in one's environment.
>
> 2. Hallucination does not involve acquaintance with mind-independent objects and properties instantiated in one's environment.
>
> C. Hence, hallucinations lack phenomenal character.

Since the Naïve Realist maintains that phenomenal character is a matter of being acquainted with mind-independent objects and some of their properties, and hallucination by definition doesn't involve this, it seems to follow from Naïve Realism that hallucinations can't have phenomenal character.

We have arrived at Fish's account of hallucination, which he sums up as follows:

> for a mental event, *e*, to be a pure hallucination, it must lack phenomenal character yet produce the same cognitive effects that a veridical experience, *V*, would have produced in a rational subject with the same background beliefs, desires, and other mental states (i.e., in the same 'doxastic setting' as the hallucinator). (pp. 94-5)

In other words, a mental event is a pure hallucination iff it lacks phenomenal character and is indiscriminable from a veridical perception (in the sense Fish has specified). A mental event is a hallucination as of, say, a yellow banana iff it lacks phenomenal character and is indiscriminable from a veridical perception of a yellow banana (again, in the relevant sense of 'indiscriminable').

I have several reservations about this account of hallucination, foremost among them having to do with the eyebrow-raising suggestion that there's nothing it's like to hallucinate. I'll begin, however, with objections concerning the indiscriminability component of Fish's account.

First, if hallucinations really lack phenomenal character, then Naïve Realism (as stated by Fish, at least) doesn't require an account of hallucination in terms of indiscriminability. To see this, suppose (*pace* Fish) that the fundamental nature of hallucination is a matter of perceptually representing propositions. Applying Martin's causal argument, a veridical perception with the same proximal cause as a hallucination also consists in perceptual representation of a proposition. Recall that the basic thrust of Martin's screening-off argument is that if a hallucination and a veridical perception have fundamental psychological commonalities, then anything else they have in common is best explained in terms of such commonalities (subject to the proviso that they aren't explanatorily parasitic on some other property had by only one of the experiences). Now, given that hallucinations *lack* phenomenal character, this

representational property doesn't screen off the obtaining of the acquaintance relation from explaining the veridical perception's phenomenal character. For the hallucination and the veridical perception don't have a common phenomenal character; this wouldn't be a case of a fundamental psychological commonality being best suited to explain *something else the experiences have in common*. Thus, a Naïve Realist can accept that veridical perception involves perceptual representation without being forced to accept that its *phenomenal character* is explained in terms of this property.

Of course, the common representational property will screen off the obtaining of the acquaintance relation from explaining all sorts of things (e.g., the beliefs generated by both the hallucination and the veridical perception). But this is no problem for Fish's Naïve Realism, which is a thesis about *phenomenal character*. In short, *if* we accept Fish's claim that hallucinations lack phenomenal character, then there's no need for the Naïve Realist to go in for an indiscriminability account of hallucination. In what follows, though, I'll consider Fish's account of hallucination in its entirety (setting aside the worry that the conjunction of claims it consists in isn't well-motivated).

My second worry about Fish's account of hallucination is that it doesn't have the resources to distinguish experiences that are intuitively very different. Compare a dog's hallucination as of a black cat against a green wall in broad daylight with a dog's hallucination as of a white cat against a green wall in broad daylight. Plausibly, these are quite different experiences, but it seems that Fish's account lacks the resources to distinguish them. We're going to be hard pressed to find any differences in cognitive effects across these experiences (assuming for the sake of argument that the dog isn't cognitively sophisticated enough to form beliefs). The color of the cat is unlikely to make any difference to the way the dog is disposed to behave towards it, given the optimal viewing conditions. (For example, the example is set up such that the dog can discern the "hallucinatory cats" from the "hallucinatory backgrounds" equally well in both cases, and so there's no reason to expect that the dog would be slightly more halting in its efforts to chase the hallucinatory white cat as opposed to the black one.) Thus, both experiences count as indiscriminable (in Fish's sense of the term) from a veridical perception of a black cat against a green wall in broad daylight and from a veridical perception of a white cat against a green wall in broad daylight. (And of course, we can't distinguish the experiences by appeal to their phenomenal character, as neither of the experiences has any.) So it seems that Fish's account incorrectly lumps together very different hallucinations.

Fish might reply by questioning the demand for a sorting of hallucinatory experiences into nice and neat kinds, as he does in response to a similar objection. Take, for example, "Martin's hasty and inattentive subject John, who treats samples of scarlet and vermillion in the same way" (p. 111). Given John's disposition to make hasty judgments regarding color, a hallucination as of a scarlet patch has exactly the same cognitive effects for John as a hallucination as of a vermillion patch, a veridical perception of a vermillion patch, *and* a veridical perception of a scarlet patch (holding his other mental states fixed). In this case, it seems that Fish's account predicts that intuitively (albeit subtly) different kinds of hallucination have exactly the same mental nature (in this case, something like: being indiscriminable from a veridical perception of a scarlet patch *and* being indiscriminable from a veridical perception of a vermillion patch). One might suggest that if an account of the mental nature of hallucination doesn't have the resources to distinguish different hallucinatory experiences, so much the worse for that account.

Fish's reply to worries of this sort is that there's no obvious reason why we should expect to be able to sort such hallucinations into nice, neat kinds (p. 112). Perhaps the metaphysical structure of hallucination consists in a messy conjunction of indiscriminability properties, or maybe it's indeterminate. (We might say that John's hallucination as of a scarlet patch consists in an indiscriminability property, but it's just not determinate which one.) On the face of it, it's not clear what's so bad about these proposals.

I'm not convinced that this attempt to shift of the burden of proof is enough to quell the worry just outlined. But even if it is, it certainly isn't enough to quell the worry with which we began. Saying that subtly different hallucinations have the same mental nature is one thing. Saying that *very* different hallucinations (e.g., a dog's hallucination as of a black cat and a dog's hallucination as of a white cat) have the same mental nature is quite another. Maybe we shouldn't expect to be able to distinguish hallucinations as of scarlet and vermillion patches. But surely we can expect our account of hallucination to distinguish hallucinations as of *black* and *white* cats.

At this juncture in the dialectic, I expect that Fish would insist that we examine the assumptions behind the claim that a dog's hallucination as of a white cat and a dog's hallucination of a black cat are fundamentally different kinds of experience. Plausibly, what motivates the claim that these experiences are very different is the background assumption that they are *phenomenally* very different (as he suggests on p. 112, fn. 30). But this begs the question against Fish's account of hallucination—on his view, hallucinations lack phenomenal character altogether. Fish would likely insist that, once we give up the idea that hallucinations have phenomenal character, there just aren't any other grounds for distinguishing hallucinations that have exactly the same cognitive effects. For example: if it *really is* the case that a dog's veridical perception of a white cat and a dog's veridical perception of a black cat would have exactly the same cognitive effects, then, given Fish's denial that hallucinations have phenomenal character, we can't even make *sense* of the idea of a dog's hallucination as of a *black* cat as opposed to a dog's hallucination as of a *white* one. In short, it seems that Fish can avoid the worry by insisting that it's generated by an assumption that he has argued against.

This response to the objection is only as plausible as the claim that hallucinations lack phenomenal character—a rather unorthodox claim to which we will turn our attention shortly. But first, I'd like to raise another concern about the indiscriminability component of Fish's account of hallucination. Since Fish elucidates indiscriminability in terms of cognitive effects, a natural type of objection to his account appeals to hallucinations that don't have any cognitive effects at all. For example, Susanna Siegel objects to Fish's account on the grounds that it "…should be possible to have a hallucination that is indiscriminable from a veridical perception of a butterfly, even if the standard effects…don't actually come about" (2008: 217, quoted on p. 104). Such hallucinations don't count as indiscriminable from any kind of veridical perception on his view, and thus his theory of hallucination doesn't seem to have the resources to account for them.

Fish's response to such objections is that they've failed "…to grasp the core of [his] approach" (p. 104, see also p. 100, p. 114). In short, he claims that on his view, a state counts as a hallucination only insofar as it has cognitive effects. If it doesn't have any such effects, it's simply not a hallucination. So it "misrepresents the claim on the table" (p. 100) to suggest that one could undergo a hallucination without it having any cognitive effects.

There are at least two causes for concern about this line of response. First, it seems dialectically inappropriate. The objector intends to reject Fish's account of hallucination based on a certain intuition, viz., that one could undergo a hallucination without it having any cognitive effects. It won't do to reply that on the conception of hallucination under discussion, such a situation simply isn't possible. For the whole problem *in the first place* is that the account predicts that such a situation isn't possible, contrary to our pre-theoretical grasp of what hallucination involves. Now, Fish can respond to such objections by attempting to explain away the intuitions his account is in tension with (see, e.g., p. 101). Whether or not such attempts are successful, the point here is simply that the objectors haven't failed to grasp the core of Fish's approach. Rather, they've compared Fish's approach with a conflicting intuition, and concluded that the intuition is worth keeping.

Second, and more importantly, Fish's account fails to discharge what is arguably an important explanatory burden. According to Fish, "…we can explain everything we need to explain (in other words, everything a hallucinating subject thinks, says, and does) by appeal to the beliefs that the hallucinating subject forms" (p. 98). However, the beliefs the hallucinator forms can't explain one crucial thing: *why the hallucinator has those beliefs in the first place*. Plausibly, part of the point of giving an account of the nature of hallucination, and of perceptual experience more generally, is to help *explain why* a given experience has the cognitive effects that it does. Fish's account of hallucination cannot perform this function, as it maintains that hallucination has no mental nature other than the production of certain cognitive effects. There's simply nothing in the account to do the explanatory work that one would reasonably expect of an account of hallucination.

Let me elaborate on this alleged explanatory role by focusing on a particular explanandum: *perceptual belief*. We might ask: why do I believe that there's a yellow banana before me? As a first pass, a plausible answer is that I'm having a perceptual experience as of a yellow banana. But there's a further question we can raise: what makes my experience the sort of thing that naturally generates the belief that there's a yellow banana before me in the first place? What must my experience be like, what fundamental structure must it have, in order for my experience to play this role? In other words, there seems to be a deep connection between my perceptual experience and the belief that there is a yellow banana before me. What exactly does this connection consist in?

If the question at issue isn't entirely perspicuous, perhaps the possible answers one might give will help to clarify it. One possible answer is that my perceptual experience is fundamentally a matter of perceptually representing the proposition that there's a yellow banana before me, and the connection between my experience and my belief is that I go on to believe the representational content of my experience (other things being equal). When it comes to veridical perception, the Naïve Realist can say that my perceptual experience is fundamentally a matter of being acquainted with the fact of the banana's being yellow, and there's obviously a tight relationship between the fact of the banana's being yellow and in front of me and the proposition that I go on to believe (that there's a yellow banana before me).

But what can Fish say in response to this demand for explanation in the case of hallucination? Since, on Fish's account, a hallucination as of a yellow banana is fundamentally a matter of being a state that has certain cognitive effects (among them the belief that the banana is yellow), the fundamental nature of this hallucination doesn't *explain why* it generates this belief. Rather, the fact that this hallucination

generates this belief is simply a matter of *definition*. Thus, given that a primary goal of a philosophical account of perceptual experience is to elucidate the deep connections between experiences and the other mental states they cause, it would seem that Fish's account of hallucination fails to discharge an important explanatory burden.

Fish is sensitive to this worry. As he notes on p. 114, "I imagine that it might be alleged that this account will be incomplete until some account of *why* hallucinations come to have the same kinds of effects as certain kinds of veridical perception has been provided." However, he responds to this worry with the questionable dialectical maneuver discussed above:

> …given the definition of hallucination presented here, such a demand is misguided…Because a mental event qualifies as a hallucination only inasmuch as it has the same kinds of effects as a certain kind of veridical perception, asking why *hallucinations* have these kinds of effects would be akin to asking what bachelors have in common in virtue of which none of them are married. Hallucinations *just are* those events that have the same kinds of effects as, and are therefore indiscriminable from, veridical perceptions of a certain kind (p. 114).

The objector rejects Fish's account on the grounds that it cannot discharge an important explanatory burden. It won't do for Fish to insist that, if his account of hallucination is correct, the correct account of hallucination cannot discharge this explanatory burden. For the whole problem in the first place is that his account cannot discharge this explanatory burden, contrary to our antecedently held ideas about the point of giving an account of perceptual experience.

Of course, Fish might respond by arguing that these antecedently held ideas are misguided—that in fact providing explanations of experiences' cognitive effects isn't (or shouldn't be) a constraint on accounts of hallucination or of perceptual experience in general. In any case, it's worth emphasizing that whether or not Fish can offer such an argument at the end of the day, the objection under discussion doesn't constitute a failure to grasp the core of Fish's approach. Rather, we have noted that Fish's approach is in tension with a plausible constraint on theories of perceptual experience, and we have no reason as yet to believe that we should give up this constraint.

Now let us turn to the aspect of Fish's account of hallucination that many will find most difficult to accept: the claim that hallucinations lack phenomenal character.[11] As a result of the argument outlined earlier, Fish thinks that the Naïve Realist is committed to this claim. Many would be inclined to characterize this argument as a reductio of Naïve Realism, but Fish devotes a whole chapter to arguing that this would be too quick—for our reasons for thinking that hallucinations have phenomenal character are not as good as is commonly supposed.

According to Fish, the primary motivation for the claim that hallucinations must have phenomenal character is the *local supervenience principle*: "if the proximal neural conditions that occurred when one perceived an object were created in the absence of that object, the subject would have an experience with the same phenomenal character

[11] Fish points out that this claim does not entail that there is *nothing* it's like for a hallucinating subject: "…if there is something it is like to believe [as some argue], then there will be something it is like for the subject, so the hallucinator will not be in the same position as a philosophical zombie" (p. 99, fn. 19)

as the original perceptual experience" (p. 41). The idea is that any two experiences with the same proximal neural causes have the same phenomenal character; the phenomenal character of an experience locally supervenes on its proximal neural causes.

It's a small step from the local supervenience principle to the conclusion that hallucinations have phenomenal character. If phenomenal character locally supervenes on proximal neural causes, then a hallucination with the same proximal neural causes as a given kind of veridical perception has the same phenomenal character as veridical perceptions of this kind (and *a fortiori*, it has phenomenal character).

While the route from the local supervenience principle to the conclusion that causally matching hallucinations have phenomenal character is unproblematic, Fish argues that we have no compelling reason to accept the starting point. In chapter 5, Fish considers three broad types of argument for the local supervenience principle and finds them all wanting: arguments from the "very occurrence of hallucinations" (i.e., from the fact that subjects actually mistake their hallucinations for veridical perceptions), arguments by way of thought experiment (e.g., from the conceivability of a brain in a vat that has an experience phenomenally identical to the one I'm having right now), and the idea that the principle is "…an empirical hypothesis borne out by the results of brain stimulation experiments" (p. 123).

For the sake of argument, let's just grant that there's no good case for the local supervenience principle, and thus that it is inadmissible as support for the claim that hallucinations have phenomenal character. However, there is another way to argue for the latter claim that Fish hasn't considered. Suppose that you're having a hallucination as of a yellow, crescent-shaped banana. Suppose further that you introspect your mental situation, and upon doing so you come to the conclusion that you're having an experience with phenomenal character—in particular, you conclude that you're having an experience as of a yellow, crescent shaped banana, and that there is something that it is like for you to be in this experiential state. Fish can grant all of this, of course, but he's committed to saying that your faculty of introspection makes a *massive error*. It has generated the belief that you're having an experience with phenomenal character, but, if Fish's account of hallucination is correct, you're not. In general, introspection is totally unreliable when it comes to the phenomenal character of hallucinations (or lack thereof).

Now, you don't have to think that introspection is an *infallible* source of information about your mental states in order to find this result unacceptable. You might think that one's introspective access to what it's like for one to be in a given mental state is fallible; moreover, it's plausible that in at least some circumstances introspection can deliver the wrong result about something as coarse-grained as *which mental states one is in*. For example, introspection may generate the belief that you desire to make amends with your estranged friend, but close examination of your behavior (e.g., your avoidance of her calls, your spreading of malicious rumors about her) would give the lie to this notion.

So introspection might be unreliable when it comes to certain fine-grained mental matters, and even about coarse-grained mental matters when it is employed in unfavorable contexts (e.g., in the presence of motives conducive to self-deception). But surely introspection isn't *so* bad as to systematically generate the belief that one has an experience with phenomenal character when in fact *the state lacks it entirely*. It is plausible that introspection is extremely reliable when it comes to the very coarse-grained question of whether one is in a state with phenomenal character, at least when the context is favorable for its operation (e.g., one isn't suffering the delirium typical of

dreams).[12] Even Eric Schwitzgebel, who is as much of a skeptic about the reliability of introspection as you'll find, would assent to this claim (see his 2008: 253).

In response to this argument, one might suggest that any situation in which one is subject to a total hallucination is an unfavorable context for the operation of one's faculty of introspection. However, this suggestion is rather *ad hoc*. Surely it's in principle possible for a hallucinating subject to be exceedingly attentive to her experience, to suffer no distractions from such a careful introspective endeavor, to have no motives conducive to self-deception regarding the phenomenal aspects of her mental situation, etc.

Another potential response is that while an introspecting hallucinator might correctly judge that she's in a state with phenomenal character, she'd be wrong if she went on to judge that her *perceptual experience* is what has that phenomenal character. For example, if there is something it's like to *believe*, hallucinating subjects might systematically mistake cognitive phenomenology for perceptual phenomenology. This suggestion isn't very promising, however, as the phenomenology associated with perceptual experience is rather distinctive. Even if there is something it's like to believe, what it's like to *believe* that there is a yellow banana before me is completely different from what it's like to have an *experience* as of a yellow banana before me. (Whatever it's like for me to believe that there's a banana before me, it doesn't involve a banana seeming to be *right there*, "staring me in the face.") The difference between perceptual and cognitive phenomenology (if there is any) seems to be not a matter not of degree but of *kind*. Thus, the suggestion that hallucinating subjects systematically mistake the phenomenal character of beliefs for the vivid, "immersing" phenomenology typically had by veridical perceptions seems implausible and *ad hoc*.

In short, even if we grant that the local supervenience principle isn't adequately supported, there is another way to substantiate the intuition that hallucinations have phenomenal character (the argument from introspection) that remains unaddressed.

Finally, it's not clear that Fish even has to go to the lengths that he does to undermine the claim that hallucinations have phenomenal character. For Fish's argument from Naïve Realism to the claim that hallucinations lack phenomenal character is too quick. The argument went as follows:

1. Phenomenal character consists in acquaintance with mind-independent objects and properties instantiated in one's environment.

2. Hallucination does not involve acquaintance with mind-independent objects and properties instantiated in one's environment.

C. Hence, hallucinations lack phenomenal character.

Rather than accept the conclusion, the Naïve Realist could insist that there is *more than one way* for a state to have phenomenal character. For example, a Naïve Realist could say that while a veridical perception has phenomenal character in virtue of consisting in

[12] It would be disingenuous to suggest that introspection is reliable when it comes to the coarse-grained question of whether or not a state has phenomenal character because, as a matter of fact, total hallucinations are rare. For what's at issue is whether introspection might systematically produce the wrong result *in cases of total hallucination*, and the argument here is that an affirmative answer is implausible.

acquaintance with mind-independent objects and some of their properties, a hallucination has phenomenal character in virtue of, say, being indiscriminable from a certain kind of veridical perception. (This wouldn't commit the Naïve Realist to saying that veridical perception has phenomenal character in virtue of being indiscriminable from a certain kind of veridical perception; the veridical perception has that indiscriminability property in virtue of *being* a veridical perception, which in turn consists in acquaintance with mind-independent objects and some of their properties.) In short, there is an at least *prima facie* plausible way of interpreting the premises such that they don't entail the conclusion.

This Naïve Realist strategy is not without its problems. For example, one might worry that it doesn't afford an adequate explanation of the indiscriminability of hallucinations from veridical perceptions. On the proposal under consideration, we cannot say that a hallucination is indiscriminable from a certain kind of veridical perception in virtue of the fact that it shares a phenomenal character with veridical perceptions of that kind. For the idea is that the order of explanation goes in the other direction: a hallucination has the phenomenal character it does in virtue of having a certain indiscriminability property. Moreover, suppose (plausibly) that a characterization of hallucination's mental nature is supposed to explain why a given type of hallucination tends to generate the types of beliefs and behaviors that it does (as I suggested above). It follows that we cannot explain why a hallucination is indiscriminable from a given kind of veridical perception by appealing to the fact that it generates beliefs and behaviors of the same type that are generated by a veridical perception of that kind. For, given the supposition, the fact that the hallucination fundamentally consists in such an indiscriminability property is supposed to explain why it generates such beliefs and behaviors.

In short, on the proposal under consideration, the indiscriminability of a hallucination from a certain kind of veridical perception may just be a brute fact, inexplicable in further psychological terms.[13] While some might find this claim implausible, I submit that it is considerably more plausible than the claim that hallucinations lack phenomenal character altogether. Thus, it seems that the Naïve Realist would do better to pursue the idea that there are different ways for a state to have phenomenal character, rather than to pin her hopes on the denial that there is anything it is like for a subject to hallucinate.

3. Conclusion

In Perception, Hallucination, and Illusion, Fish does an admirable job of summarizing the current state of the debate about Naïve Realism, as well as advancing the dialectic beyond that state. Most importantly, he identifies a promising yet hitherto overlooked motivation for Naïve Realism, one which should bring even Naive Realism's most trenchant critics to admit that the view is worth taking seriously. In my view, Fish's account falters when it comes to hallucinations, but its fault lines intersect with fundamental issues that require further discussion and debate (e.g., the purpose of giving a philosophical theory of perceptual experience, and the extent of our introspective access to experience). In short, anyone on either side of the debate over

---

[13] However, it's not obvious why an explanation in purely neural terms (e.g., it terms of sameness of proximal neural causes) wouldn't be sufficient.

Naïve Realism, and those wanting to learn what all the fuss is about, would do well to study Fish's book closely.

References

Byrne, A., and H. Logue. 2008. Either/Or. *Disjunctivism: Perception, Action, Knowledge*, ed. A. Haddock and F. Macpherson. Oxford: Oxford University Press.
Campbell, J. 2002. *Reference and Consciousness*. Oxford: Oxford University Press.
Fish, W. 2009. *Perception, Hallucination, and Illusion*. Oxford: Oxford University Press.
Johnston, M. 2004. The obscure object of hallucination. *Philosophical Studies* 120: 113-83.
———. 2006. Better than mere knowledge? The function of sensory awareness. *Perceptual Experience*, ed. T. Gendler and J. Hawthorne. Oxford: Oxford University Press.
Martin, M. G. F. 2004. The limits of self-awareness. *Philosophical Studies* 120: 37-89.
———. 2006. On being alienated. *Perceptual Experience*, ed. T. Gendler and J. Hawthorne. Oxford: Oxford University Press.
McDowell, J. 1986. Singular thought and the extent of inner space. *Subject, Thought, and Context*, ed. J. McDowell and P. Pettit. Oxford: Oxford University Press.
———. 2008. The disjunctive conception of experience as material for a transcendental argument. *Disjunctivism: Perception, Action, Knowledge*, ed. A. Haddock and F. Macpherson. Oxford: Oxford University Press.
Schwitzgebel, E. 2008. The unreliability of naive introspection. *Philosophical Review* 117.
Siegel, S. 2004. Indiscriminability and the phenomenal. *Philosophical Studies* 120: 91-112.
———. 2008. The epistemic conception of hallucination. *Disjunctivism: Perception, Action, Knowledge*, ed. A. Haddock and F. Macpherson. Oxford: Oxford University Press.