

## On the Explanatory Power of Hallucination

Dominic Alford-Duguid and Mike Arsenault

Suppose Jane glances at the mess on her desk. Her visual experience will present a distinctive scene containing the objects littering her desk (a notebook, lamp, etc.) along with some of their properties (e.g. the notebook's shape or the lampshade's colour). A theorist of perceptual experience must spell out what it is for experience to present such a scene. On the one hand, this task can seem deceptively easy. What could be more obvious to us than the nature of our experience? You might have thought that introspection simply settles the question, leaving no room for doubt that Jane's experience of the lamp involves (for example) a direct relation to the lamp and its colour. Yet on the other hand, centuries of introspection have brought no consensus, a fact which presses the question: how are we to decide between rival views of the metaphysics of experience?

This paper is about one particular version of that question—how are we to decide between *disjunctive* (particularly *naïve realist*) and *intentional* accounts of experience?—and one particular strategy for answering it. According to naïve realist disjunctivists, veridical experience consists in a certain kind of relation to objects and properties in the viewed scene (this is the 'naïve realism'), a relation that is absent when a subject undergoes hallucinations in which the relevant objects are not present. The result is that these theorists give one account of veridical experience, and another very different account of hallucination (thus 'disjunctivism'). Intentionalists, by contrast, hold that there is a common experiential kind to which at least some hallucinations and veridical experiences belong (and so deny that experience ever *consists* in a direct relation to ordinary objects and their properties).

The strategy for deciding between these views that we're concerned with, and which has received support from both sides of the aisle in recent years (e.g. Campbell 2002 in the

naïve realist camp, Pautz 2010 in the intentionalist), involves a kind of transcendental argument: we can decide which account of the metaphysics of experience is right by first deciding what is required of experience in order to explain our capacity for (perceptual demonstrative) thought about the objects and properties we perceive. This strategy offers the hope of traction in what looks like an otherwise intractable dispute. For perhaps by arguing from claims about thought to claims about perception we can avoid a deadlock of introspection-heavy arguments.

We are cautiously optimistic about this strategy (hereafter ‘THP’—Thought to Perception), and conclude the paper with suggestions for how it might be carried through. But we begin with a case study of its going awry. Pautz (2010) attempts to implement a version of THP against naïve realism. In particular, he targets the combination of naïve realism and the most minimal naïve realist account of perceptual hallucination—*negative epistemic disjunctivism*. He claims that hallucination as of a novel property—not just veridical experience of it—sometimes suffices to ground our capacity to *think* about the property. And he argues that naïve realists who accept negative epistemic disjunctivism—and so offer an account of the metaphysics of hallucination that is very different from their account of veridical experience—cannot explain how hallucination can ground this capacity in this way.<sup>1</sup>

This way of arguing against naïve realism succeeds only if Pautz can motivate his claim that hallucinations ground the capacity for demonstrative thought about a property by using considerations that are independent of disputed (and disputable) claims about phenomenology and perception. One has gotten nowhere in implementing the THP strategy if disputed claims about perception and phenomenology seep into one’s account of thought,

---

<sup>1</sup> Pautz’s is the latest in a string of objections that challenge naïve realism’s account of the explanatory power of hallucination (cf. Johnston 2004 and Matthen 2010). Where Pautz focuses on our capacity to think about sensible properties, these other objections isolate different capacities (e.g. a capacity to know higher order facts about sensible properties).

only to be projected back down into conclusions about perception.<sup>2</sup> We follow the lead of Hellie (2010, p. 104, fn. 5)—and borrow some apparatus he develops—to argue that this is just what has happened in the course of Pautz’s argument. And in consequence, we think Pautz fails to get dialectical traction against his opponents.

The paper proceeds in five parts. §1 sketches naïve realism in more detail, while §2 outlines Pautz’s objection. §3 does the heavy lifting: we show how Pautz’s case against naïve realism fails. §4 draws morals from §3, and offers suggestions as to how one might derive claims about perception from a claim about thought with a minimum of contentious intuitions or introspective data.

## §1 Naïve Realism and Negative Epistemic Disjunctivism

Naïve realism comes in several flavours. What follows is merely an outline of its general features.

*Naïve realism* is the view that in good cases—when the world and our perceptual systems behave—perceptual experience consists in a three-place relation between a subject, a perspective, and the particular objects and properties in view.<sup>3</sup> So to have a veridical experience of a rectangular notebook is just to stand in this relation to the notebook and its shape. Naïve realism thus combines two substantive metaphysical theses: (1) perceptual experience is fundamentally *relational* and (2) in veridical cases the relation holds between the perceiving subject and ordinary elements of her perceptible external environment. These theses feed into the naïve realist account of the qualitative character of a perceptual

---

<sup>2</sup> It might turn out that perceptual demonstrative thought isn’t sufficiently distinct from perception to run a version of THP in the first place. But we think this claim is unduly pessimistic at this stage, and could be justified only by further research.

<sup>3</sup> Cf. Campbell (2009) for this characterization. Recent defenses of naïve realism include Martin (2002), Campbell (2002), Brewer (2011), and Soteriou (2013).

experience (i.e. how things seem to the subject who undergoes the experience). At least in veridical cases, the objects and properties to which an experience relates a subject combine with relevant features of the subject's perceptual perspective to determine the qualitative character of the experience.<sup>4</sup>

Naïve realism is not current orthodoxy, in part due to challenges posed by hallucination and illusion (so-called 'bad' cases). Illusory perceptual experiences present an object as some way it isn't (e.g. a green object may look red). But put illusion to one side, since Pautz focuses on hallucination. In cases of hallucination, a subject has an experience as of an object (say a chair), yet no such object is available to be perceived. Naïve realists cannot characterize hallucination as they do veridical perception: there is no object (and thus, arguably, no property) to which the subject can stand in the required relation. So they appear to lack the resources to explain matching hallucinations—hallucinatory experiences a reflective subject is unable to discriminate from corresponding veridical experiences.<sup>5</sup>

Naïve realists have not ignored the challenge posed by hallucination. Because they explicitly restrict their account of the qualitative character of perceptual experience to veridical experience, many take themselves to be forced to offer *negative* accounts of the qualitative character of hallucinatory experience (cf. Martin 2002 and Brewer 2011). For instance, Martin (2002) defends a limited account of hallucination:

---

<sup>4</sup> This account of the character of perceptual experience fits nicely with a natural story about perception's role in grounding certain of our capacities to think about the properties of objects. Subjects acquire the capacity to think about colours and shapes (and other sensible properties) in part because in 'good cases' their perceptual experience confronts them directly with such properties. Perceptual experience thus provides some of the subject matter of thought. But this is just one elaboration of the old idea that veridical perceptual experience (the naïve realist's 'good case') grounds our capacities to think about the external world.

<sup>5</sup> Strictly speaking, subjects fail to discriminate what is presented, rather than the experiences themselves, but here we speak of discriminating experiences as shorthand.

*Negative Epistemic Disjunctivism.* The qualitative character of a central class of hallucinatory experiences must be fundamentally explained in terms of their reflective indiscriminability from veridical experiences.

The ‘central class’ are those hallucinations that share a proximal cause—a certain neural state, for instance—with a corresponding indiscriminable veridical experience. Purely negative accounts of hallucination such as Martin’s are often motivated by ‘screening-off’ worries: if we provide a *positive* account of such hallucinations, the features mentioned therein supposedly screen off, and so render explanatorily redundant, the explanatorily relevant features of the indiscriminable veridical experiences (cf. Martin 2004).

Martin’s view remains controversial, but we can sidestep that debate. Nothing much in our paper hangs on how one develops negative epistemic disjunctivism. We focus on this austere account of hallucination both because it is one of Pautz’s explicit targets, and because we think that if austere versions of naïve realism can answer Pautz, then so can less austere versions.

## §2 Pautz’s Argument from a Distance

Pautz’s argument depends on the *Grounding Intuition*: if a subject undergoes a hallucination as of a property F, she can thereby acquire the additional capacity for beliefs that are true iff something is F (a so-called ‘F-involving belief’).<sup>6</sup> He thinks this intuition applies to (at least) both shape and colour properties. Paradigm F-involving beliefs are those that represent F by means of demonstrative concepts (such as one might express with ‘this’ or ‘thus’). We shall thus treat demonstrative beliefs as one species of F-involving beliefs, sometimes using ‘perceptual demonstrative thought’ in place of ‘F-involving belief’.

---

<sup>6</sup> For a statement of the *Grounding Intuition*, see p. 266 of Pautz (2010). He discusses paradigm cases of F-involving beliefs in Pautz (ms).

It's best for our purposes to present Pautz's argument in the context of a character such as Jackson's (1982) Mary.<sup>7</sup> Mary knows everything physical about the brain and the world in purely 'physical' terms, but because she lives in a black and white room she has never seen red (or any other colour), and so does not know 'what it's like' to see red (*cf.* Nagel 1974 for the 'what it's like' terminology). Suppose she then undergoes a hallucination as of a red square (and so becomes 'Hallucinating Mary'). The *Grounding Intuition* entails that this hallucination could ground Hallucinating Mary's capacity to think a new kind of thought about the property red, just as if she had really seen it—a kind of thought she could not have had before 'experiencing' red. If *Grounding* is correct, then it doesn't matter whether the grounding experience is veridical or hallucinatory.

With this case in mind, we can state Pautz's argument as follows. Supposing that pre-hallucination Mary lacks the capacity to think thoughts involving redness, then:

- 1) [*Grounding*] A hallucinatory experience as of a red patch, *H*, is such that Mary's "having it *explains* [her] new capacity to have beliefs involving redness..."
- 2) [*Incompatibility*] "This could not be true of the mere property of failing to satisfy a certain epistemic condition." (i.e. of failing to be discriminable from the corresponding veridical experience).

So,

- 3) "...contrary to epistemic negative disjunctivism, *H* cannot be identical with the property of failing to satisfy a certain epistemic condition."<sup>8</sup>

*Grounding* is simply a statement of the *Grounding Intuition*, whereas *Incompatibility* is a theoretical claim about the explanatory powers of negative epistemic disjunctivism. 1 is supposed to entail that Hallucinating Mary can form F-involving beliefs on the basis of *H*

---

<sup>7</sup> This is not Pautz's case, but we use it here because it represents the hardest case for the argument we present below. If our argument succeeds in showing that *Grounding* fails in Mary's case, then it should fail in less extreme cases as well. These include real life cases such as the hallucinations involved in Charles Bonnet syndrome, for which see Fernandez et al. (1997). Thanks to an anonymous referee from this journal for bringing the relevance of Charles Bonnet syndrome to our attention.

<sup>8</sup> For this argument, see pp. 277-279 of Pautz (2010). It is important to emphasize that Pautz's argument is meant to work in the case of shape as well. We focus on the case of colour only for ease of exposition.

*because* of its phenomenal character. In conjunction with 2, it follows that naïve realists cannot explain Mary’s acquisition of the capacity to have an F-involving belief about this shade of red on the basis of her hallucination of it. Pautz then uses this conclusion (3) as part of an argument to the best explanation against naïve realism, and in favour of a view on which there is an explanatorily relevant common factor between veridical and hallucinatory experiences. Naïve realism is false according to this abductive argument because it isn’t part of the best explanation of *Grounding*.<sup>9</sup>

### §3 Why Pautz’s Argument Against Naïve Realism Fails

In §3.1 we develop our primary response to Pautz’s argument, which is that *Grounding* (i.e. 1) has no dialectical traction against the naïve realist. Though this argument stands on its own, in §3.2 we present a secondary argument against *Incompatibility* (i.e. 2) for those readers who, despite our best efforts, remain wedded to *Grounding*.

#### §3.1 Against *Grounding*

Our argument against *Grounding* proceeds in three steps. First, we identify the basic problem with Pautz’s implementation of the THP strategy. Second, we put pressure on *Grounding* by pointing to an unwanted consequence of it which, third, justifies the invocation of an error theory that explains away any remaining intuitive plausibility associated with *Grounding*.

##### §3.1.1 Dialectical Traction

---

<sup>9</sup> There are various ways of understanding the relationship between transcendental and abductive arguments, depending especially on how ‘transcendental argument’ is understood. Nothing of substance in our paper turns on how one understands the relationship between them (e.g. we could allow that transcendental arguments are a subspecies of abductive arguments).

In questioning whether Pautz's argument gets dialectical traction against the naïve realist we are taking up the baton from an argument of Mike Martin's, to which Pautz responds, and according to which because naïve realism and negative epistemic disjunctivism together imply that *Grounding* is false, Pautz cannot appeal to *Grounding* in arguing against naïve realism.<sup>10</sup>

Pautz's reply to Martin's objection is perfunctory: *Grounding* begs the question only if every argument that begins with intuitions inconsistent with the target theory thereby begs the question against that theory. After all, intuitions like these are supposed to be part of the data an account of perception is supposed to explain.

Pautz is certainly correct that we cannot condemn every argument that draws on intuitions inconsistent with its target. But this is not the sort of charge we (or, presumably, Martin) wish to make. Abductive arguments persuade because they demonstrate that opponents' explanations fail as the best explanations of data which all parties at least find compelling. But the dispute between naïve realists and their opponents is in part a dispute over which putative data are, or should be thought, compelling. And yet it seems that *Grounding* takes us no great distance from claims that were already in dispute. The original dialectic between intentionalists and naïve realists had the intentionalist insisting that phenomenal indiscriminability means that perceptions and hallucinations share an important experiential common factor. The naïve realist resists this move. In appealing to *Grounding* Pautz is simply reduplicating this original dispute one level up: now the intentionalist insists that since we can achieve the same *perception based* F-involving thought *via* veridical perception or hallucination, veridical and hallucinatory perception must share an important common factor. But once a naïve realist has denied that the experiences on which these

---

<sup>10</sup> For Pautz's attribution of the argument to Mike Martin, together with his reply, see Pautz (2010, p. 278, fn. 19).



thoughts depend are of the same kind, why should she accept that the thoughts that are based on them are of the same kind? To make headway in arguing for *Grounding* the intentionalist needs to make it plausible that hallucinations could serve just as well as veridical experiences in grounding predicative thought by appeal to independent considerations *about the nature of predicative thought*. So if (as we'll argue in the next section) *Grounding* is, or depends upon, a disputable claim about perception and phenomenology, its presence at the foundation of Pautz's argument immediately ruins any hope of gaining traction in a debate about the metaphysics of experience by appeal to (relatively) independent claims about thought.

### §3.1.2 Symmetry

We've observed that if the intuitive support for *Grounding* comes from disputable claims about perception, then Pautz's argument is in trouble. To see whether it is founded on such intuitions, we can compare Pautz's application of *Grounding* to cases of perceptual demonstrative thought that are not in dispute: to begin with, few would suggest that the analogue for *Grounding* goes through for thought about objects (as opposed to thought about properties). The putative objects of hallucinations fail to exist. And we cannot have object-involving beliefs—beliefs that are true iff the object has some property—about objects that fail to exist.

The fact that the *objects* of hallucinations don't exist, whereas the properties apparently represented in hallucinations—such as redness—do exist (even if they are not actually instantiated in the subject's environment), might seem to justify this asymmetry in our intuitions about the truth of *Grounding* with respect to objects and properties.<sup>11</sup> But this

---

<sup>11</sup> Indeed, Pautz (2010, p. 266-7) tries to push just this sort of response.

rather obvious difference hides an important *symmetry* between our demonstrative thought about properties and objects, a symmetry revealed more clearly by cases of so-called ‘veridical hallucination’.

Hallucinations are “veridical” when there just happens to be a corresponding real object where the hallucinatory object appears to be. Suppose Bill is having a veridical hallucination, on the basis of which he then tries to think about the very basketball he (merely) seems to see in the doorway. Bill’s attempt at a demonstrative thought fails in this case, as in other such cases of hallucination. Even though there is in fact a basketball in the doorway, because he’s not seeing the real ball in the doorway he cannot think an object-involving demonstrative thought about it. *Prima facie*, what’s missing from Bill’s veridical hallucination isn’t just that the basketball he seems to see in the doorway does not exist, but that he has no *access* to the ball that really is in the doorway. In Bill’s case, the missing access seems to be (at least) missing causal access. Because his experience is caused by an evil neuroscientist, say, and not by the real ball’s affecting his sensory apparatus in the right sort of way, he’s lacking (at minimum) the direct causal link perception ordinarily provides with objects in his environment.<sup>12</sup> This suggests that perception subserves perceptual demonstrative thought about objects partly through establishing a direct causal link with the object to be thought about (we leave open whether the access demanded by perceptual demonstrative thought must include something beyond a mere causal link—some have made the stronger claim that perceptual demonstrative thought about an object demands *rational* access to the object).<sup>13</sup>

---

<sup>12</sup> Our appeal to causation here is meant to be intuitive. We intend nothing more full-blooded or theoretical in appealing to causation than what is already implicated in our pre-theoretical judgements about cases, such as this case of veridical hallucination.

<sup>13</sup> For the claim that perceptual demonstrative thought about an object demands that perception provide rational access to the object, see Smithies (2011).

But if perception subserves demonstrative thought about objects by establishing a direct causal link with those very objects, why think that we can have perceptual demonstrative thoughts about properties without being (or ever having been) similarly linked to instances of those properties? Whatever causal story applies to demonstrative thought about objects, *Grounding* is incompatible with the analogous story for thought about properties, because in undergoing a hallucination as of a red object one is not causally related to anything red.<sup>14</sup> So accepting *Grounding* requires revising this straightforward story about how experience makes possible our thought about properties.

The trouble for Pautz’s argument against the naïve realist is that he hasn’t mustered any independent considerations about *thought* to justify treating objects and properties differently here. Intuitions pumped by our trying to imagine ourselves (or Mary) hallucinating may tell us something about our attitudes to colour experience, but they do nothing to explain why *thought* about properties should make different demands of perception than *thought* about objects. The attitudes about perception and colour pumped by these cases, such as that (some of) the objects of awareness are quasi-mental, that colours are wholly revealed in our perception of them, etc. provide no independent vantage from which to refute naïve realism using the THP strategy.<sup>15</sup>

### §3.1.3 Error Theory

Pautz holds that hallucinations ground F-involving thoughts, such as the thought purportedly expressed by Mary’s saying “I believe something is *that way*” (where “that way” is

---

<sup>14</sup> It is important to resist the feeling that Mary *has* come in causal contact with an instance of redness after all, in undergoing her hallucination, since her hallucination is itself red. Her hallucination is not red – at best it is indiscriminable from an experience *of* something red. For more on whether *Grounding* is compatible with a (partly) causal story about perception’s role in our account of thought, see our response to Pautz’s *Argument from Access* against the compatibility of naïve realism and *Grounding* in §3.2.

<sup>15</sup> One naïve realist who has worked against these attitudes to secure the objectivity of perceptible properties (especially colour) is John Campbell (1993, 2002, 2005, 2006). See also Hellie (2010) and McDowell (1986).

meant to pick out the relevant property F). We want to offer an error theory of such statements. According to our error theory, this statement of Mary's is not true (we'll remain neutral about whether it's false or neither true nor false). We think it's not true because "that way" is empty—it fails to pick out a property.

That's our error theory in a nutshell. It is backed by the intuitive idea that perception grounds thought about objects *and* properties by confronting us with the subject matter of thought, that is, by the considerations of symmetry discussed above, in the last section.

But we want to add to the plausibility of our error theory by explaining why it might seem that Mary is speaking the truth in saying "I believe something is *that way*", even though, if we're right, she's not. That is, we want to explain away the *Grounding* intuition, not merely show that it's false.

We think that *Grounding* seems true in the case of the properties Pautz considers (in contrast with objects) because of the way colour and shape properties are organized along dimensions of variation, such as hue and saturation in the case of colour (the case of shape is obviously more complex).

We propose to bring out the relevance of this consideration to the difficult case of Mary by working up to that case in steps, starting with the more benign case of Hume's missing shade of blue thought experiment. Hume (*Enquiry* 2.1.6) suggests—seemingly contrary to his own theory—that a subject could think imaginatively about a shade of blue she has never seen before by exploiting her prior familiarity with other colours (and especially, presumably, with other shades of blue). His idea was not simply that one can exploit one's background capacities to think *descriptively* about the relevant colour, but that one can recover something of the *phenomenology* of experiencing it—what it's like—by way of this sort of exercise of the imagination.

If it is plausible that we can recover the missing shade of blue in imagination, it's only because colours occupy determinate locations in colour space. It is our familiarity with the dimensions of variability of colour, together with knowledge of the colours in the vicinity of the missing shade, that explain Hume's confidence in our ability to "lock-on" to the missing shade of blue in imagination-based judgements so as to be able to think an F-involving thought about it.

By contrast, we could have no such confidence in judgements formed about particulars with which we're unacquainted. No matter how many humans you have met, it would be absurd to try to form the equivalent<sup>16</sup> of a perceptual demonstrative thought about Sam by exploiting your acquaintance with Bill and Sally and Jennifer. For Sam is not defined by his position alongside Sally and Jennifer and Bill in anything like the way colours are defined by their positions in colour space (or shapes in some geometric space).

Since it seems that Hume can rely upon imagination rather than perception to think an F-involving thought about the missing shade, the foregoing suggests that non-hallucinatory perception may not be *absolutely* indispensable for our F-involving thought about colours (despite its indispensability to perceptual demonstrative thought about objects). For maybe Hume is right that imagination can make us aware of novel shades provided one has seen enough of the spectrum of colours. But accepting that a subject could recover the missing shade of blue in imagination *does not* force us to accept that non-hallucinatory perception is (as *Grounding* might suggest) everywhere dispensable where F-involving thoughts are concerned. The imaginative capacities that enable subjects to recover

---

<sup>16</sup> This use of 'equivalent' is meant only to make allowance for the fact that we might not want to call demonstrations of imagined objects and properties "perceptual" demonstrations, even when the species of imagination in question is visual, or auditory, etc. It also bears remarking that we are of course not denying that you can think descriptively about people you have never met. We are concerned throughout with the capacity for perceptual/ 'imaginative' demonstrative thought in particular.

the missing shade of blue ultimately depend on previous encounters with, and so successful perceptual demonstrative thought about, other colours. It thus turns out that the missing shade of blue case is the exception that proves the rule: in general, *non-hallucinatory* perception plays the same sort of indispensable role in our demonstrative thought about properties as it does in our thought about objects. For even in the missing shade case, the capacity for F-involving thought about the shade is ultimately explained by a background of perceptual interaction with other colours and a grasp of their dimensions of variation.<sup>17</sup>

Nonetheless, the missing shade of blue case identifies a genuine difference between objects, on the one hand, and colours and shapes, on the other, which is relevant to our thought about them. The relevance of this difference to Hallucinating Mary and *Grounding* emerges when we consider an intermediate case: Hallucinating Hume.

Hallucinating Hume is just like Hume in the original missing shade of blue scenario, except that in his case the missing shade comes unbidden in the form of a hallucination, rather than through the intentional exercise of imaginative capacities. The question now is whether, in exclaiming “That’s new!” upon hallucinating the missing shade of blue, Hallucinating Hume succeeds in expressing an F-involving thought about the missing shade. According to our error theory, he *does not* succeed: he tries to exploit visual contact with a shade of blue when there is no such contact to be had.<sup>18</sup>

---

<sup>17</sup> One might wonder: why doesn’t a history of colour hallucinations suffice when it comes to recovering the missing shade of blue? We do not need to show that this is definitely impossible, but only that it could hardly be agreed to be obvious that it is possible. It is not as though one can test its plausibility by trying to imagine oneself conjuring up a shade of blue one has never seen before while thinking “Also, all my experiences of colour thus far have been hallucinations.” How could one tell, in the course of such a thought experiment, whether what one was imagining was a successful conjuring-up of a shade of blue one had never seen before, rather than a failed attempt at thinking an F-involving thought about a shade of blue? So it looks as though nothing can be done with the missing shade of blue thought experiment to show that one can think imaginatively about a colour without ever having seen *any* colours.

<sup>18</sup> *Mutatis mutandis*, we take our response to the case of Hallucinating Hume to cover actual cases such as those mentioned in fn. 7, as well the illusory perceptions of novel colours studied in Billock and Tsou (2010), and discussed in Pautz (2013).

If we're right that Hallucinating Hume fails to essay the relevant F-involving thought, we have to explain why one might nonetheless think that *Grounding* is true in the case of Hallucinating Hume. Our answer is that Hallucinating Hume's failed attempt to think an F-involving thought is ripe for sense-making, or the practice of making adjustments in one's understanding of other subjects so as to make sense of them when they might otherwise be unintelligible (think of how effortlessly one sometimes massages a struggling student's ideas into something coherent in the vicinity of what they did "say"—which might have been strictly speaking incoherent). In our application of sense-making to perceptual phenomena, we follow Hellie (2010). His account of the perspective-dependence that runs through sense-making provides the apparatus that we'll use to undermine the remaining intuitive force of *Grounding*.

There exists considerable room for debate over how sense-making works, as well as over when precisely it is appropriate. But we can remain safely neutral on these questions, since sense-making clearly applies to cases like Hallucinating Hume's. For there should be little doubt that we do (and ought to) try to understand others as having expressed contents available to their perspective, which licence their utterances and other aspects of their behaviour, if the alternative is peevish insistence that they had not thought or said anything. Accordingly, although no such F-involving thought was available from Hallucinating Hume's perspective—since he was not *seeing*, and so in causal contact with, anything of the appropriate shade—we engage in sense-making and so treat him as though he had succeeded in thinking about the shade in order to make sense of his perspective.<sup>19</sup>

One reason it might seem natural to treat Hallucinating Hume as though he really had succeeded in thinking an F-involving thought is that he can't tell he's not seeing the

---

<sup>19</sup> We owe the point about sense-making to Hellie (2010, p. 104, fn. 5).

relevant shade, and so not thinking such a thought. But indiscriminability can't be doing all the work, because it is present in the object case too, and yet we don't find *Grounding* plausible for thought about objects. Here again the fact that colours are organized in colour space does important work: because the structure of colour space is mirrored in the structure of Hume's understanding of the colours, we can be sure there were many nearby—i.e. highly determinate—ways of picking out the relevant shade of blue (*via* description, or imagination, etc.) available from his perspective, even though he didn't avail himself of them in fact. Accordingly, we can explain our tendency to treat Hallucinating Hume as though he had succeeded in thinking an F-involving thought as reflecting our knowledge of the availability of these highly determinate *descriptive* contents. In ideal conditions, we might be able to make explicit the various ways of picking out the shade from Hume's perspective: e.g. "Ah, there is a shade of blue between my wallpaper and my pillow case after all!" But in ordinary conditions one is unlikely to know *precisely which* highly determinate descriptive contents to supply unless one is the evil neuroscientist who knows which colour Hume seems to see. But this lack of full information only adds to the naturalness of proceeding as though Hume had succeeded in *demonstrating* the shade, which is why *Grounding* seems plausible in the first place. For until one knows what colour Hume seems to see, one's best bet is to proceed as though he had demonstrated some colour or other, as a sort of place holder, or peg on which to hang subsequent descriptions ("Ah, he thinks he saw a new colour—we'll see which one as we go along").

The fact that the sense-making content is highly determinate is important, because the more determinate and faithful to Hume's perspective it is, the more of Hume's behaviour we can explain. After all, Hallucinating Hume will no doubt behave in ways characteristic of someone who had succeeded in thinking an F-involving thought about a



highly determinate shade of blue: he might later think he was re-identifying the same shade (though of course he couldn't actually be re-identifying the colour, as he never succeeded in identifying it in the first place), and he might "succeed" in replicating it with paint, etc. This too adds to the tendency to gloss over his error. After all, Hallucinating Hume certainly looks as though he can do everything someone who had succeeded in thinking an F-involving thought can do. Nonetheless, we do not need to go in for *Grounding* to explain his performances: these are amply explained in terms of indiscriminability together with his pre-existing understanding of colours and colour space, and the descriptive contents such understanding makes available for describing the character of his hallucination.

In appealing to sense-making, we have begun to adjudicate the question of whether Hallucinating Hume has succeeded in thinking a thought from the third-person perspective. Of course, there's no point in trying to adjudicate the question from Hallucinating Hume's own perspective, since he will think he's succeeded because he thinks he's seeing (rather than hallucinating). But to this one might respond: Couldn't you imagine discovering you were hallucinating, and yet continuing to think you had succeeded in thinking an F-involving thought? After all, this case might seem more relevant to attempts to think our way into Hallucinating Hume or Mary's perspective. If that's right, then insofar as these cases are 'purely first-personal', it might seem that we cannot re-describe them in terms of sense-making, since these re-descriptions seem to require a third person perspective. The trouble for this response is that even this sort of diachronic case isn't 'purely' first-personal in the relevant sense. For in taking a stance on the veracity of your earlier experience, you are effectively adopting a third person perspective on that earlier self, so that issues of sense-making and (admittedly privileged) interpretation come in all the same. In short, you engage in the project of making sense of your past (and present) self, given what you now know.

There is thus no standpoint from which someone could pump pro-*Grounding* intuitions while avoiding our re-descriptions in terms of sense-making.

Our claim so far has been that a subject's facility with colour concepts and attendant grasp of colour space grease the wheels of sense-making so as to generate the specious plausibility of *Grounding*. But Hallucinating Mary's grasp of colour space is much more deficient than Hallucinating Hume's: he is missing only one shade, while she is missing every colour. Consequently, there will be few (if any) highly determinate descriptive contents available to her perspective for picking out the shade she seems to see. Does this not show that our error theory fails to apply to subjects as deficient as Mary?

It does not, but this objection brings out why Hallucinating Mary's case presents the greatest challenge for our view. Our error theory explains the specious plausibility of *Grounding* as it applies to Mary because the attempts to think our way into Mary's perspective that pump the *Grounding* intuition are relatively insensitive to differences between Hallucinating Hume and Mary. To make this plausible, two points need to be observed.

First of all, it is easy and natural in trying to assume Mary's perspective to overestimate the determinacy of the descriptive contents available from that perspective. After all, not only is she familiar with all the colours 'described physically', she is familiar *perceptually* with the black, white, and shades of grey that cover the surfaces in her room. It is easy to mistake this familiarity with the monochromatic shades for an (admittedly) weaker analogue of the pre-existing framework of concepts that Hallucinating Hume has. Accordingly, when trying to imagine her perspective from the inside it is hard to keep firmly in view that these monochromatic shades are not arrayed in colour space alongside the other

colours, so it is often tempting to treat her monochromatic vision as closer to our colour vision than it really is.<sup>20</sup>

There are other differences it's hard to keep track of. While Mary would no doubt claim to be able to re-identify the shade on subsequent occasions, just as Hallucinating Hume (or we) would, it's not obvious that she would know how to produce it from other colours, or that she could reproduce it, or re-identify objects of that colour under different lighting conditions. The trouble is that using these differences to characterize her perspective by subtraction from our own leaves little to fall back on when characterizing her hallucinatory perspective<sup>21</sup>, let alone in deciding whether her hallucination affords her a capacity to think about red. In consequence, it is natural to play up similarities (spurious or not) and downplay differences in the course of trying to understand her perspective. And the more we do so, the more we treat her like Hallucinating Hume, and so find sense-making similarly natural.

Moreover, since in non-ideal (i.e. actual) cases we cannot settle upon determinate descriptive content until the interaction progresses, the expedient of treating Hallucinating Mary as though she had *demonstrated* (or actually had an F-involving thought about) the shade seems as natural in her case as it does in Hallucinating Hume's. Consequently, the fact that there might not have been a highly determinate descriptive content available from her

---

<sup>20</sup> Assimilating black and white to the colours is very natural. It is no accident that many early theories of colour and colour vision, including Aristotle's (see *De sensu*. 3), explain the other colours in terms of black and white. For a fascinating account of these early theories, informed by contemporary philosophy and vision science, see Kalderon (2015), esp. chs. 5-6.

<sup>21</sup> In a more positive vein, it is tempting to employ heuristics of dubious value. For example, it is tempting, but it won't do in fact, to imagine her perspective as being similar to our own when, in the course of watching a hitherto black and white movie, an (actually) red rose suddenly appears. But even this strategy inadvertently assimilates Mary to us, for our experiences of the shades of grey on the screen retain important associations with colours (bananas ordinarily look the way yellow things look, and so are represented by a lighter shade of grey than red roses, etc.) that are missing for her.

perspective to licence the sense-making is easily lost in the fray, so that it is perfectly natural that her case should pump the same misleading intuitions as Hallucinating Hume's.

It will help in rounding-out our discussion to come back to the issue of re-identification. Suppose Mary later *sees* the shade of red she merely seemed to see before. Is it not plausible that she could identify that shade as the shade she saw/seemed to see before? And if she could so identify the shade, would that not show that she had grasped some particular shade all along, contrary to our hypothesis? We think not. A comparison with the object case brings this out. Suppose John had had a hallucination as of some (novel) man whom he then claims to *see* again the next day. Nobody should be inclined to say that the presence of this disposition to “re-identify” a look-alike shows that he had actually succeeded in demonstratively identifying that actual man (or any other) on the basis of his previous hallucination, since his hallucination didn't put him in touch with anyone in the first place (see our point about access in §3.1.2). But the same is true of Mary, whether or not there is a colour property corresponding to her hallucination: there's no reason to treat the mere presence of a putative disposition to “re-identify” the shade as evidence that she originally succeeded in thinking an F-involving thought.

A summary is in order. First, we argued that Pautz's appeal to *Grounding* has no dialectical traction unless he can show that it *really is* a generally appealing, or compelling, datum. Then we challenged its appeal by pointing out that it is inconsistent with an important *prima facie* symmetry between thought about properties and thought about objects. Finally, we explained how someone could find *Grounding* appealing even though its false by showing how our tendency towards sense-making leads us to treat hallucinating subjects' mistaken reports of their own thoughts as though they were in order.

### §3.2 Against *Incompatibility*

In the previous section we challenged the status of *Grounding* as an intuitively compelling introspective claim. This suffices to block Pautz' argument without our saying anything about *Incompatibility*, since if *Grounding* is not intuitively compelling then naïve realism has not been shown incompatible with anything a theory of perceptual experience must explain. Nonetheless, in this section we lay out Pautz's argument for the incompatibility of *Grounding* with naïve realism. Pautz offers three *positive* arguments for *Incompatibility* (cf. Pautz 2010, p. 278), the first two of which we'll treat together.

The first is what we'll call the *Argument from Mystery*, which begins by emphasizing that a commitment to negative epistemic disjunctivism forces naïve realists to explain our capacity to form F-involving beliefs on the basis of hallucinations in terms of a subject's failure to stand in a particular epistemic condition (i.e. the condition in which her experience is not reflectively indiscriminable from a veridical experience). The argument from mystery then simply contends that it is not clear what such an explanation would look like.

The *Argument from Access* conjectures that our capacity to form F-involving beliefs requires that a subject stand in some physical relation to the property F. But, Pautz insists, failure to stand in a particular epistemic condition cannot explain the fact that a subject stands in such a relation to F.

We are not impressed by the arguments from *Mystery* and *Access*. Nobody has a satisfactory account of what it is to have an F-involving belief, so the explanatory gap highlighted by *Mystery* applies to naïve realists and their opponents equally. This objection simply reinforces one of the morals of this paper, which is that in deciding questions about perception by deploying the THP strategy we have to first make independent decisions about thought.

One might object, as did an anonymous reviewer from this journal, that at the very least the intentionalist promises a relatively non-mysterious account of hallucination-based perceptual demonstrative thought: Perceptual hallucinations can ground corresponding beliefs because both perceptions and beliefs are intentional states with contents.

It is easy to overstate the ease of the intentionalist's story. First, it is not at all clear how this view should be spelled out so as to satisfy the *Grounding* intuition, according to which *perceptual* content would have to be more basic than, and so explanatory of, the corresponding thought content.<sup>22</sup> Moreover, this strategy runs the risk of solving one mystery by generating another, bigger, mystery. For the more we assimilate perception and thought, the harder it is to explain the apparent difference between them. Indeed, Pautz ends up treating the distinction between believing and “sensorily entertaining” a content as brute (2007, 519). *Prima facie*, though, the apparent difference between perception and thought is a feature of perception that we might have expected a theory of perception to explain. So it seems there is enough mystery to go around, and the naïve realist need not be particularly concerned with the argument from *Mystery*.

Similarly, the *Argument from Access* does not pose a special challenge to naïve realism. For the argument to work, it must isolate a physical relation that both explains *Grounding* and fails to beg the question against naïve realism. But it is not at all clear how advocates of *Grounding* could appeal to a physical relation in explaining a hallucination's capacity to put one in touch with redness. It is tempting to think that an evolutionary story could be told, according to which certain brain states evolved to represent redness, so that a causal/counterfactual relationship holds between those brain states and instantiations of

---

<sup>22</sup> Both Campbell (2002, p. 122) and Dickie (2015, §4.6) argue that for perception to ground thought, the intentionality (or ‘content’) characteristic of perception must be different in kind from the intentionality (or ‘content’) characteristic of thought.

redness. But if we allow that hallucinations as of redness are multiply realizable by different brain-states, some of which have *not* evolved to subserve seeing red things in the environment, then the argument falls apart. Since *Grounding* is presumably supposed to be phenomenologically plausible no matter what brain state underlies the relevant hallucination, advocates of *Grounding* can't restrict its application to subset of indiscriminable hallucinations without sacrificing the purported phenomenological support for *Grounding*.

But what other feature of one's mental state could causally co-vary with redness, if not the brain state itself? If it is suggested that it is the hallucinatory experience that somehow co-varies with F-ness in the world, then if the co-variation claim is not simply false—what reason is there to think that hallucinations of redness co-vary with instances of red?—it is true only if veridical and hallucinatory experiences share a common experiential kind. But surely that simply begs the question against the disjunctivist naïve realist, since this common kind claim is precisely what disjunctivists deny. So Pautz's chances of telling a non-question begging positive story about our capacity to think F-involving beliefs in terms of a physical relation look dim. What looked like a problem for naïve realism looks more like a problem for *Grounding*.

Pautz's final (and strongest) argument requires more care. Call it the *Argument From Explanation*. It begins with a plausible thesis about explanation, which Pautz attributes to David Chalmers (Pautz 2010, p. 248): if an object's possession of a property  $\phi$  explains its possession of some other property  $\Psi$ , then if the object had not possessed  $\phi$  it would not have possessed  $\Psi$ . But Pautz contends that a subject would not lack the capacity to form F-involving beliefs on the basis of hallucination if her experience were reflectively discriminable from veridical experience of an object that is F. So a naïve realist cannot explain the capacity of subjects to form F-involving beliefs in terms of reflective

indiscriminability. Hence naïve realists cannot explain our capacity to form F-involving beliefs on the basis of hallucination.

In order to make our case against this objection, let's get it more clearly in view. We need to know which cases of the negative epistemic property's failing to apply are relevant to the argument. Here's one case: Suppose we're trying to explain why my having an experience of *blue* instead of red explains why I am unable to think a redness-involving thought on the basis of my *blue* experience. Here the negative epistemic disjunctivist's account seems right: it's because this experience of blue is discriminable from an experience of red – it's just plain different with respect to the colour it is an experience of – that it fails to generate a capacity to think about redness. So Pautz's argument fails when the discriminable experience is an otherwise ordinary non-F veridical experience. The remaining option is that Pautz's argument applies when the discriminable experience is reflectively discriminable from any *possible* veridical experience.

Here the relevant case is that of afterimages, such as those we experience when we shut our eyes after staring at the sun. If a red afterimage is discriminable from any veridical experience of redness, and yet it generates a capacity to think about redness, the argument from explanation succeeds. But as Phillips (2013) has recently argued, in a large range of cases it is simply not clear that afterimages actually *are* discriminable from any corresponding veridical experience. For there could be veridical experiences in which we encounter a certain kind of light phenomenon projected from our own perspective that turn out to be indiscriminable from afterimages. If Phillips is right, any afterimages that fail the negative epistemic condition will be especially peculiar. And the more peculiar they are, the less likely it is that they could put a subject in touch with a genuine feature of the world, as would be required for their use in Pautz's *Grounding* argument.



#### §4 Deriving Claims About Perception From A Claim About Thought

Our limited defence of both naïve realism and Pautz’s proximate target—the negative epistemic account of hallucination adopted by many naïve realists—stands on its own. But the aims of this paper are not exclusively negative. We believe some version of the strategy we called ‘THP’—a strategy that derives claims about perception from claims about thought—holds promise as a methodology for resolving disputes between naïve realists and their opponents. In particular, the strategy offers an attractive alternative to excessive reliance upon introspection. But our study of Pautz’s argument suggests that it is all too easy in implementing THP to slide into trading on the very claims about experience to which this strategy promised an alternative. We will therefore use this opportunity to make some suggestions as to how THP might be implemented without contamination from controversial assumptions about sense perception, and so vindicate our faith in the strategy’s promise.

Pautz’s appeal to *Grounding* failed to gain dialectical traction against negative epistemic disjunctivists. One might worry that this problem generalizes to undermine all but the most toothless versions of THP. So we must ask: when does appeal to a claim about thought gain traction against a rival account of perception?

THP tends to employ one of two kinds of claim about thought. The first concerns *when* a given sort of thought is available. Pautz’s *Grounding* is a claim of this kind, since it says that F-involving thoughts are available even when a subject undergoes an hallucination as of something F. Claims of the second kind concern *what makes it the case* that a given sort of thought is available. The parallel claim for Pautz would concern what enables a subject to

form a genuinely perception based F-involving thought (as opposed to a thought that fails to represent F).

Now there's nothing dialectically ineffective *per se* about versions of THP that begin from a claim about *when* a particular sort of F-involving thought is available. Some claims of this kind are uncontroversial. Virtually all parties agree that veridical perceptual experience of an object can enable a subject to think perceptual demonstrative thoughts about it (similarly for veridical experience of properties and thought about them).<sup>23</sup> And they'll agree that there is some point at which the perceptual link with an object becomes too tenuous to support a perceptual demonstrative thought (yet surely we cannot settle upon where to draw this boundary in advance of some account of perceptual demonstrative thought).

However our complaint about Pautz's deployment of THP is *not* that he starts from non-veridical experience. There's no reason that a proponent of THP need confine herself to the veridical case; she might venture into non-veridical cases when she has the backing of independent claims about thought. For example, it is a mark of perceptual demonstrative thought that a subject can think such thoughts even about an object (some of) whose properties she gets wrong. This feature of demonstrative thought helps to distinguish it from descriptive thought, since in paradigm cases of descriptive thought a subject must get (at least) the property mentioned in the description correct.<sup>24</sup> It also ensures that illusory (and so non-veridical) experiences sometimes enable subjects to think perceptual demonstrative thoughts about perceived objects.

---

<sup>23</sup> Cf. Evans (1982); Campbell (2002); Smithies (2011); Recanati (2012); Dickie (2015). For a dissenting voice, see Armstrong (1968). Since he treats perceptions as beliefs, perceptions can't *ground* beliefs.

<sup>24</sup> Some version of this claim is presupposed by all of those theorists mentioned in the previous footnote. There are very nice questions about whether a parallel claim applies to thought about properties (as opposed to perceptual demonstrative thought about objects).

By contrast, we think Pautz's *Grounding* is controversial (and ultimately dialectically ineffective) because it *both* concerns non-veridical experience, *and* yet has no *independent* basis in a claim about thought (it therefore neither enjoys the intuitive support of the veridical case, nor the theoretical support potentially afforded by independent considerations drawn from our theory of thought). The lesson for a proponent of THP is clear: to successfully implement the strategy one must be careful to stray beyond the 'obvious' veridical cases only if armed with independent considerations drawn from our theory of thought.

Pautz's argument against the negative epistemic disjunctivist account of hallucination began with a claim about *when* F-involving thought is available. Yet a largely unexplored (and potentially fruitful) alternative implementation of THP to thought about properties would instead begin with a claim about *what makes it the case* that perception enables F-involving thought (see Campbell 2002 for this alternative implementation of S to the parallel case of perceptual demonstrative thought about objects).

We believe the question of how perception enables F-involving thought remains wide open. But we want to illustrate how one might motivate a THP strategy using various accounts of thought. Our entry point is the question of whether hallucination can provide *knowledge* of what it would be for something to possess a sensible property (e.g. a colour or shape property). To see where *knowledge* enters the picture we must introduce a plausible and widely held thesis about what it takes to represent a property in thought: our capacity to think about a property requires *knowledge* of what it would be for something to have the property. A quick (and controversial) argument for this thesis runs as follows:

1. Subjects can only think those thoughts they understand.

2. To understand a thought of the form  $\langle \alpha \text{ is } \phi \rangle$  a subject must know what it would be for the thought to be true.<sup>25</sup>

Yet

3. A subject possesses this knowledge only if she grasps  $\langle \alpha \rangle$  and  $\langle \phi \rangle$ .

What does it take to grasp  $\langle \phi \rangle$ ? Assuming a thought of the form  $\langle \alpha \text{ is } \phi \rangle$  is true iff the bearer of  $\langle \alpha \rangle$  instantiates  $\phi$  (i.e. the property represented by  $\langle \phi \rangle$ ), many philosophers have found the following answer compelling:

4. To grasp  $\langle \phi \rangle$  requires knowing what it would be for something to be  $\phi$ .<sup>26</sup>

But 1-4 entail 5:

5. To think about  $\phi$  using a thought of the form  $\langle \alpha \text{ is } \phi \rangle$  requires that a subject know what it would be for something to be  $\phi$ . [from 1-4]

To resist the move to 5 would require that we reject one of 1-4. But none of these premises are individually in tension with either the commitments of naïve realism or those of its opponents. 5 suggests that hallucination enables us to think about a sensible property only if it provides knowledge of what it is for something to have that property.<sup>27</sup> Thus a question ripe for genuine and fruitful disagreement arises: what does it take to know what it would be for something to have a sensible property? Only after first answering this sort of question

---

<sup>25</sup> We use ‘<’ and ‘>’ to represent thoughts (and their components), and ‘ $\langle \square \rangle$ ’ and ‘ $\langle \phi \rangle$ ’ to pick out representationally relevant components deployable in thought.

<sup>26</sup> This claim—that grasp of a predicative concept requires knowledge of what it would be for the referent of the concept to be instantiated—has found many adherents. For instance, we find versions of it in Strawson (1959), Dummett (1981), Evans (1982), Peacocke (1992), and Campbell (2002). Also, see Reimer (2002).

<sup>27</sup> One way for someone like Pautz to block this entailment would be to deny that the *Grounding Intuition* concerns predicative representation of sensible properties. Hallucination might instead enable us to think about sensible properties only via name-like nominal representations (e.g. the thought expressed by ‘redness is not similar to greenness’). The apparent neutrality of Pautz’s notion of F-involving belief (i.e. beliefs whose truth requires something to be F) seems to support this dodge. But this neutrality should be taken with a grain of salt. Predicative concepts are our standard means for thinking about properties, since properties are fundamentally entities we *ascribe* to objects.

should we then ask whether a given account of hallucination satisfies the conditions required for perception based F-involving thought.<sup>28</sup>

Which way someone goes at this point will thus depend on her view about what it takes to know what it would be for something to have a given property. There exist (at least in principle) a broad spectrum of views. As in the case of object-directed thought, a right account of F-involving thought might invoke different kinds of knowledge for different species of F-involving thought. Of course, a proponent of THP will be most concerned with the sort of knowledge required by perception based F-involving thought. We'll canvas two views of this sort of knowledge in order to give the reader a flavour of what someone might say about perception based F-involving thought. We take no stand on the truth of the views themselves.

One relatively non-demanding view would be to treat the relevant knowledge as grasp of a kind of primitive mechanics for a property—of the principles that capture the interaction of the putative property with other properties, objects, or events. For example, knowledge of what it would be for something to be round might involve, *inter alia*, knowing that round objects roll when on a smooth incline. Knowledge of such principles is knowledge of what *difference* possession of a property makes to an object.<sup>29</sup>

A more demanding view would deny that grasp of a primitive mechanics for a property suffices for the requisite knowledge of what it would be for something to have that

---

<sup>28</sup> There's no doubt that we're going to have to appeal to intuitions about perception and phenomenology at some point. The trick in implementing THP is to delay those appeals for as long as possible, so as to stick to prior and independent considerations about thought as much as possible. This will ensure a maximum of common ground between the two camps.

<sup>29</sup> For the notion of a primitive mechanics, and a discussion of its role in thought about properties, see Evans (1980). Shoemaker (1980) defends a view on which the identity of a property is exhausted by its contribution to the causal role of objects that instantiate it—by the difference it makes to an object.

property.<sup>30</sup> Such a view might insist that we also require knowledge of a state of affairs in which the property is instantiated. Veridical experience provides such knowledge, since it presents objects and their actual properties; hallucinations do not.

Again, we take no stand on which (if either) of these two views is correct. Our interest is in illustrating how independent claims about thought might be brought to bear in choosing between accounts of perception, and so in what sorts of considerations might figure in a dialectically effective version of THP. The two views we've canvassed impose different constraints on a right account of perception. The first view entails that perceptual experience enables F-involving thought only if it provides knowledge of a primitive mechanics for F. The second view denies that such knowledge suffices for the requisite knowledge of what it would be for something to have F. It is therefore possible that the same account of hallucination could enable F-involving thought when combined with a view of the first sort, but preclude such thought if combined with a view of the second sort. As a result, someone who endorses *Grounding* prematurely runs a real risk. For it may turn out that, given a right account of how perception explains F-involving thought, the theorist's chosen account of hallucination is incompatible with *Grounding*. So it may turn out that *Grounding* poses as serious a challenge for its proponents as for negative epistemic disjunctivists—we must wait for a right account of F-involving thought to find out (this point echoes similar remarks we made in §3.2).

More generally, a given account of perceptual experience could be ruled out if, when combined with one of the two views just sketched (or some other view in the wide spectrum of possible views), it entailed that we *never* have F-involving thoughts. This result would run

---

<sup>30</sup> Someone who pursues this option, though not quite in the way we sketch, is Campbell (2002, 2006). He emphasizes the importance of our knowledge of the *categorical* property that grounds dispositions captured by the primitive mechanics.

counter to even the most uncontroversial claims about when F-involving thoughts are available. THP thus remains viable even if we begin with a claim about *what* makes it the case that F-involving thoughts are available.

The potential dialectical power of THP comes from the fact that it begins with claims about thought that we can accept or reject without committing ourselves to controversial theses about perception. Part of what we've sought to do in this section has been to highlight some of the wide range of claims about thought that do not fall prey to the problems that dog *Grounding*, and which might thus serve as the basis of a successful implementation of THP. These include certain claims about which perception based thoughts subjects can form in response to non-veridical perceptual experience.

Obviously what we've said in this final section does not constitute a decisive defence of THP—even the most promising strategy must earn its place through success. But we hope to have said enough to ward off the worry that the problems that afflict Pautz's implementation of THP render the strategy a non-starter.

## §5 Conclusion

This paper pursued two goals. The first was primarily negative. We showed how Adam Pautz's objection to naïve realism fails, not just as a matter of detail, but from the very start. Rather than beginning with an independent claim about thought, Pautz packs disputed claims about perception into his claim about thought (i.e. *Grounding*), and then tries to walk a path back from this claim about thought to a disputed claim about perception.

Our second goal was forward-looking. Pautz's objection constitutes a failed attempt to derive a claim about perception from a claim about thought. But we're optimists about the prospects for a strategy that begins with a claim about thought—for instance a claim

about *when* subjects can form F-involving thoughts—and derives a claim about perception. We called this strategy ‘THP’. We believe some version of THP promises progress in the debate between naïve realism and its opponents. To vindicate our optimism we sketched some examples for those who wish to pursue the strategy in a fashion that avoids the pitfalls that undermine Pautz.

### Works Cited

- Armstrong, D. M. (1968). *A Materialist Theory of the Mind*. London: Routledge.
- Barnes, J. (Ed.). (2014). *Complete Works of Aristotle, Volume 1: The Revised Oxford Translation: The Revised Oxford Translation* (Vol. 1). Princeton University Press.
- Billock, V., & Tsou, B. (2010). Seeing Forbidden Colors. *Scientific American*, 302: 72–77.
- Brewer, B. (2011). *Perception and its Objects*. Oxford: Oxford University Press.
- Campbell, J. (1993). A Simple View of Colour. In John J. Haldane & C. Wright (eds.) [\*Reality: Representation and Projection\*](#). Oxford: Oxford University Press.
- Campbell, J. (2002). *Reference and Consciousness*. Oxford: Clarendon Press.
- Campbell, J. (2005). Transparency vs. Revelation in Color Perception. *Philosophical Topics* 33: 105-115.
- Campbell, J. (2006). Manipulating Color: Pounding an Almond. In Gendler and Hawthorne (eds.) *Perceptual Experience*. Oxford: Oxford University Press.
- Campbell, J. (2009). Consciousness and Reference. In Brian McLaughlin, Ansgar Beckermann and Sven Walter (eds.) *Oxford Handbook of Philosophy of Mind*. Oxford: Oxford University Press: 648-662.
- Chalmers, D. (2006). Perception and the Fall from Eden. In Gendler and Hawthorne (eds.) *Perceptual Experience*. Oxford: Oxford University Press.
- Dickie, I. (2015). *Fixing Reference*. Oxford: Oxford University Press.
- Dummett, M. (1981). *Frege: Philosophy of Language* (Second Edition). London: Duckworth.



- Evans, G. (1980). Things Without The Mind. In *Philosophical Subjects*. Oxford: Oxford University Press.
- Evans, G. (1982). *The Varieties of Reference*. J. H. McDowell (Ed.). Oxford: Clarendon Press.
- Fernandez, A., Lichtshein, G., & Vieweg, W. V. (1997). The Charles Bonnet Syndrome: a Review. *The Journal of nervous and mental disease*, 185(3), 195-200.
- Genone, J. (forthcoming). Appearance and Illusion. *Mind*.
- Hellie, B. (2010). An Externalist's Guide to Inner Experience. In Bence Nanay (ed.) *Perceiving the World*. Oxford: Oxford University Press.
- Hellie, B. (2011). There It Is. *Philosophical Issues* 21: 110-164.
- Hellie, B. (2015). Love in the Time of Cholera. In Berit Brogaard (ed.) *Does Perception Have Content?* Oxford: Oxford University Press.
- Hume, D. (1748). *An Enquiry Concerning Human Understanding*. Ed. L. A. Selby-Bigge.
- Kalderon, M. E. (2015). *Form Without Matter: Empedocles and Aristotle on Color Perception*. Oxford University Press.
- McDowell, J. (1986). Singular Thought and the Extent of 'Inner Space'. In John McDowell & Philip Pettit (eds.), *Subject, Thought, and Context*. Clarendon Press.
- Jackson, F. (1986). What Mary Didn't Know. *Journal of Philosophy* (83): 291–295.
- Johnston, M. (2004). The Obscure Object of Hallucination. *Philosophical Studies* 120(1), 113-183.
- Martin, M. G. (2002). The Transparency of Experience. *Mind & Language* 17(4): 376-425.
- Martin, M. G. (2004). The Limits of Self-Awareness. *Philosophical Studies* 120: 37-89.
- Matthen, M. (2010). Color Experience: a Semantic Account. In Mohan Matthen and Jonathan Cohen (eds.) *Color Ontology and Color Science*. Bradford Books.
- Nagel, T. (1974). What is it like to be a bat? *The Philosophical Review* 83(4): 435-450.
- Pautz, A. (2007). Intentionalism and Perceptual Presence. *Philosophical Perspectives* 21 (1): 495-541.
- Pautz, A. (2008). How visual consciousness reaches to the world. Manuscript.
- Pautz, A. (2010). Why explain visual experience in terms of content? In Bence Nanay (ed.) *Perceiving the World*. Oxford: Oxford University Press.

- Pautz, A. (2013). Do the Benefits of Naïve Realism Outweigh the Costs? Comments on Fish, *Perception, Hallucination and Illusion*. *Philosophical Studies*, 163(1): 25-36
- Peacocke, C. (1992). *A Study of Concepts*. The MIT Press.
- Phillips, I. (2013). Afterimages and Sensation. *Philosophy and Phenomenological Research* 87(2): 417-453.
- Recanati, F. (2012). *Mental Files*. Oxford University Press.
- Reimer, M. (2002). Do Adjectives Conform to Compositionality? *Nous* 36 (s16): 183-198.
- Shoemaker, S. (1980). Causality and Properties. In Peter van Inwagen (ed.), *Time and Cause*. D. Reidel: 109-135.
- Smithies, D. (2010). ‘What is the Role of Consciousness in Demonstrative Thought?’ *Journal of Philosophy* 108 (1): pp. 5-34.
- Soteriou, M. (2013). *The Mind’s Construction*. Oxford University Press.
- Strawson, P. F. (1959). *Individuals*. London: Methuen.