Of the Perfect and the Ordinary: Indistinguishability and Hallucination

Shivam Patel  
Florida State University  
forthcoming in *The Philosophical Quarterly*

Abstract: The claim that perfect hallucination is introspectively indistinguishable from perception has been a centrepiece of philosophical theorizing about sense experience. The most common interpretation of the indistinguishability claim is modal: that it is impossible to distinguish perfect hallucination from perception through introspection alone. I run through various models of introspection and show that none of them can accommodate the modal interpretation. Rejecting the modal interpretation opens up two alternative interpretations of the indistinguishability claim. According to the generic interpretation, hallucination is indistinguishable from perception despite the existence of possible exceptions, while according to the actuality interpretation, the indistinguishability of hallucination from perception consists in the actual failure to distinguish hallucination from perception. These alternative understandings of the indistinguishability claim have a number of significant implications for the problem of perception, including the rejection of perfect hallucination and illusion in favour of our ordinary, non-philosophical concepts of these states.

*Keywords:* disjunctivism; common kind theory; indistinguishability; modality; hallucination
1. Introduction

A prominent principle in the philosophy of perception is the *indistinguishability claim*: that perfect hallucination is introspectively indistinguishable from perception. Philosophers take the claim to be a truism of common sense, an attitude expressed in Sturgeon (2006: 119):

…every view of experience – disjunctive or otherwise – can make use of indiscriminability. The fact that experience [perfect hallucination] is indiscriminable from veridical perception is a datum of everyday life. It is a non-partisan resource in the area. Every theory can use it for whatever it sees fit.

The claim is often used against naïve realism, the view that the fundamental nature/phenomenal character of perception is constituted by features of the mind-independent world (Martin 2004; Fish 2009; Brewer 2013). Opponents of naïve realism have argued that the introspective indistinguishability of perfect hallucination from perception shows that perception has the same nature/character as perfect hallucination (e.g. Price 1932: 31-32; Farkas 2006: 222-223; Hawthorne and Kovakovich 2006: 179; Tye 2009: 560). In turn, naïve realists have responded by rejecting the inference from indistinguishability to sameness in nature/character (e.g. Martin 2006: 366-368; Fish 2009: 93-94; Brewer 2013: 98-99). Throughout the dialectic between naïve realists and their opponents, however, the indistinguishability claim itself has gone unquestioned – both sides take the claim to be inviolable. My aim in this paper is to question this attitude by arguing that the dominant interpretation of the indistinguishability claim should be rejected.

Philosophers of perception characterize perfect hallucination in terms of a *modal* reading of indistinguishability, according to which indistinguishability implies an impossibility to

---

1 Arguments in which the indistinguishability claim plays a prominent role include renditions of the argument from hallucination (Pitcher 1971: 13-14) and the causal argument from hallucination (Martin 2004: 53-54). A variant of the claim – replacing ‘perfect hallucination’ with ‘perfect illusion’ – is present in many formulations of the argument from illusion (e.g. Ayer 1940: 5-7). There are, of course, arguments against naïve realism that do not depend on the indistinguishability claim, including the time lag argument (Russell 1997) and the causal argument (Robinson 1994).
distinguish. Of course, it is possible to distinguish perfect hallucination from perception through testimony and inference. Theorists therefore posit introspection as a context that can accommodate the impossibility. However, in the first half of the paper, I argue that introspection cannot be made to serve as a context relative to which perfect hallucination is impossible to distinguish from perception. The upshot is that perfect hallucinations – in the sense intended by philosophers of perception – are metaphysically suspect.

Of course, there is something to the claim that hallucination is indistinguishable from perception. After all, psychologists, psychiatrists, and psychonauts are each in their own way interested in such states. Having rejected the modal reading of indistinguishability, in the second half of the paper I argue for two alternative interpretations of the indistinguishability claim. Based on the semantics of ability/inability ascriptions, I argue that the indistinguishability claim should be interpreted either as a generic or as having an actuality entailment. Thus, someone might be unable to distinguish hallucination from perception despite moments when she is successful in doing so (generic reading) or someone might be unable to distinguish a hallucination from a corresponding perception in the sense that she actually fails to do so (actuality reading).

Given the outsize role that perfect hallucination and indistinguishability have played in theorizing about experience, the paper has a number of ramifications for the problem of perception. First, the generic and actuality readings of the indistinguishability claim capture our notion of ordinary hallucination. After all, ordinary hallucinations – the sort that interest scientists and laymen – are sometimes told apart from corresponding perceptions despite being indistinguishable from perception. Second, inferences grounded on the modal reading of indistinguishability are no longer viable once we reject it in favour of the generic and actuality readings. Finally, it is often assumed that the general concept of experience that forms the subject matter of the problem of
perception should be characterized, at least in part, in terms of the modal reading of indistinguishability. However, if we reject the modal reading in favour of generic and actuality readings, we are forced to reassess the very subject matter of the problem of perception.

In Section 2, I explain the significance of the indistinguishability claim in the problem of perception. Section 3 characterizes the modal interpretation of the indistinguishability claim and describes two conditions on that interpretation. Section 4 shows there is no model of introspection that satisfies both conditions. Section 5 traces the failure of the modal interpretation to the assumption that indistinguishability implies an impossibility to distinguish. I then explain how ability and inability ascriptions, of which indistinguishability ascriptions are an instance, instead admit of generic and actuality interpretations. Having replaced the modal interpretation with the generic and actuality interpretations, Section 6 draws out two implications of significance for the philosophy of perception. I end with a comment on the general concept of experience in Section 7.

2. The Significance of the Indistinguishability Claim

The indistinguishability claim has played a foundational role in the problem of perception. Most prominently it structures the debate between naïve realism and restrictive theories of perception. According to naïve realism, the fundamental nature/phenomenal character of perception is constituted, in part, by relations of awareness to actually existing features of the mind-independent world (Martin 2004: 39; Fish 2009: 15; Brewer 2013: 96). In contrast, restrictive theories reject the idea that perception is so constituted. Sense-datum theorists, for example, claim that the nature/character of perception is constituted by relations of awareness to sense-data, non-physical objects with sensible properties (Russell 1997; Broad 1927; Robinson
1994), while intentionalists claim that the nature/character of perception is constituted by representations that concern mind-independent objects (Harman 1990: 36; Byrne 2001: 199; Siegel 2010). Thus, where naïve realism characterizes perception in terms of actually existing features of the mind-independent world, restrictive theories characterize perception in terms that stop short of the actual existence of such features.

This difference informs two general views regarding the relationship between perception, on the one hand, and perfect hallucination and perfect illusion, on the other. According to common kind theory, perception, perfect hallucination, and perfect illusion share the same nature/character (Sturgeon 1998). Sense-datum theorists tend to be common kind theorists insofar as they argue that perception, perfect hallucination, and perfect illusion involve relations of awareness to sense-data (Robinson 1994). Intentionalists also tend to be common kind theorists insofar as they posit representations across the three categories (Harman 1990). In contrast, naïve realists tend to be disjunctivists. As I will understand it, according to disjunctivism, the fundamental nature/character of perception – constituted as it is by mind-independent features – differs from that of perfect hallucination and perfect illusion (Martin 2004).

Some restrictivists seek to undermine naïve realism by employing the indistinguishability claim. They have put forward what I will call the indistinguishability-sameness inference: if perfect hallucination is indistinguishable from perception through introspection, then perfect hallucination and perception share the same fundamental nature/character. The basic intuition behind the indistinguishability-sameness inference is that sameness in the phenomenal character of perception and perfect hallucination is the best explanation of the introspective

---

2 Some recent naïve realists reject disjunctivism (e.g. Johnston 2004; Ali 2018; and Masrour 2020). Nevertheless, these authors either explicitly adopt the modal interpretation of the indistinguishability claim (Masrour 2020: 740) or else do not obviously reject it (e.g. Ali 2018: 616).
indistinguishability of perfect hallucination from perception (for both sympathetic and critical discussion of the underlying intuition see Price (1932: 31-32); Austin (1964: 50); Martin (2008: 91-92); Smith (2002: 26); Robinson (2022: 14)). In seeking to undermine the indistinguishability-sameness inference and thereby rescue naïve realism, Martin (2006) reconceives of the relationship between introspective indistinguishability and perfect hallucination. According to Martin (2006: 369), the phenomenal character of (at least some) perfect hallucinations is exhaustively constituted by its introspective indistinguishability from perception. This manoeuvre saves disjunctivism since perfect hallucination and perception end up having distinct phenomenal characters: while the phenomenal character of perception is constituted by relations of awareness to features of the mind-independent world, the phenomenal character of perfect hallucination is exhausted by its introspective indistinguishability from perception.\(^3\)

In addition to structuring the debate between common kind theory and disjunctivism, the indistinguishability claim also plays a more foundational role by characterizing the very subject matter of the problem of perception. The subject matter of the problem of perception is *sense experience*. On at least one telling, all theories engaged in the problem of perception assume that sense experience is, at least in part, introspectively indistinguishable from perception (Martin 2006: 366). According to both common kind theorists and disjunctivists, then, perfect hallucinations, perfect illusions, and perceptions all satisfy the property of being introspectively indistinguishable from perception. However, according to the common kind theorist, introspective indistinguishability from perception is necessary but not sufficient for counting as a sense experience, while according to the disjunctivist, introspective indistinguishability from perception is both necessary and sufficient. Within common kind theory, some authors hold that a sense

\(^3\) For discussions of the tenability of Martin’s notion of perfect hallucination see Siegel (2004); Sturgeon (2006); Byrne and Logue (2008); Fish (2009); Robinson (2022).
experience, in addition to being introspectively indistinguishable from perception, must also involve sense-data (sense-datum theory) or representations (intentionalism). In contrast, since the naïve realist generally takes herself to be precluded from giving a uniform account of perception, perfect hallucination, and perfect illusion, she must treat introspective indistinguishability from perception as both necessary and sufficient for counting as a sense experience.4

The indistinguishability claim therefore plays at least two roles in the problem of perception: it serves as premise in the indistinguishability-sameness inference, and it characterizes the subject matter of the problem of perception. In Sections 3 and 4, I show that the indistinguishability claim is unworkable in the sense in which philosophers intend it.

3. Common Ground: The Modal Interpretation

It is generally agreed upon that indistinguishability implies an impossibility to distinguish. On this view, for a given particular \( x \) and kind \( F \), the indistinguishability of \( x \) from the \( F \)s is to be analysed in terms of the impossibility of distinguishing \( x \) from the \( F \)s. Timothy Williamson, M.G.F. Martin, and Katalin Farkas illustrate the agreement:

What is indiscriminability? Surface form indicates that things are indiscriminable if and only if it is not possible to discriminate between them. (Williamson 2013: 5)

…the relevant conception of what it is for one thing to be indiscriminable from another is that of not possibly knowing it to be distinct from the other. (Martin 2006: 363)

If the subject is hallucinating a teacup, what she cannot tell is that her present experience is not a veridical perception of a teacup. (Farkas 2006: 215)

---

4 Siegel (2004: 94) argues that it is not a necessary condition on being a sense experience that it be introspectively indistinguishable from perception. She presents a case in which one has a perfect hallucination of an impossible scene (e.g. an Escher staircase) concerning which one trivially cannot have a corresponding perception. Martin (2004: 80-81) responds by claiming that even if introspective indistinguishability from perception cannot be applied as a whole to perfect hallucinations of impossible scenes, parts of such a perfect hallucination are introspectively indistinguishable from corresponding parts of a perception. The debate between Siegel and Martin is orthogonal to the concerns of this paper, since the criticisms I make against the modal interpretation apply to either a ‘whole’ or ‘part’ interpretation of the indistinguishability claim.
These authors (and many others) agree that indistinguishability implies an impossibility to distinguish. However, as is routinely noted, there are means by which it is possible to know that a given perfect hallucination is not a perception (Martin 2006: 364-365; Sturgeon 2006: 125-126).

For example, imagine I am placed by researchers in a machine that produces perfect visual hallucinations, but am told when I am undergoing a hallucination. Or again, I may come to infer when it is that I am undergoing a perfect visual hallucination on the basis of some accompanying occurrence, e.g. an itch. In both cases I would know that the experience I am then having is a perfect hallucination and not a perception. Given these possible scenarios, theorists must specify some particular context relative to which it is impossible to distinguish perfect hallucination from perception. Theorists thus tend to adopt what I will call the modal interpretation of the indistinguishability claim.

**The Modal Interpretation:** $x$ is indistinguishable from a perception of an $F$ if and only if relative to context $C$, it is not possible to know that $x$ is not a perception of an $F$.

The context that ends up being selected is introspection: relative to the context of introspection, it is impossible to distinguish perfect hallucination from perception (Martin 2006: 364-365; Sturgeon 2006: 122-124; Brewer 2013). Introspection serves the role of bracketing possible exceptions in which one distinguishes perfect hallucination from perception.

It is true that theorists sometimes slide between talk of ‘introspective’ indistinguishability and ‘qualitative’ indistinguishability (e.g. Robinson 1994: 160-160). Qualitative indistinguishability suggests that perfect hallucination is impossible to distinguish from perception with respect to qualitative properties shared between perfect hallucination and perception. The problem with this interpretation, and the reason I do not consider it at length in this paper, is that the indistinguishability claim is supposed to be assumable by both common kind theorists and disjunctivists without prejudging the debate between them (see Sturgeon 2006: 119). However, if
acceptance of the indistinguishability claim amounted to acceptance of the qualitative identity of perfect hallucination and perception, then the common kind theorist and disjunctivist could not find common ground in the indistinguishability claim. It is crucial to the possibility of debate that the chosen context not presuppose a resolution of the debate. The context of introspection appears to be such a context. I suggest, then, that authors who slide between talk of ‘introspective’ and ‘qualitative’ indistinguishability should be read with charity as invoking the context of introspection. The neutrality of introspection makes it apt to serve as the contextual restrictor in the modal interpretation.

However, there is also a deeper reason why introspection is chosen: there are two conditions that must be satisfied by the contextual restrictor in the modal interpretation, and introspection seems uniquely placed to satisfy both. On the one hand, the contextual restrictor must explain the impossibility to distinguish perfect hallucination from perception. Call this the explanatory condition. The explanatory condition is needed because shifts in context can change whether items are indistinguishable. For example, two letters may be indistinguishable relative to uncorrected vision but distinguishable relative to corrected vision. Or again, relative to the context of testimony it is possible to distinguish perfect hallucination from perception, while relative to the context of introspection it is impossible. Shifts in context can change whether items are distinguishable or indistinguishable because contexts are explanatory of whether items are distinguishable or indistinguishable. There is something about uncorrected vision – e.g. interactions between light and cornea shape – that explains why the two letters are indistinguishable and something about corrected vision – e.g. interactions between light, glasses, and cornea shape – that explains why the letters become distinguishable. So too there is something about testimony that explains why perfect hallucination is distinguishable from perception when
testimony is the operative context and there is something about introspection that explains why perfect hallucination is impossible to distinguish from perception when introspection is the operative context. Thus, the context relative to which it is impossible to distinguish perfect hallucination from perception must explain why it is impossible to do so.

On the other hand, the contextual restrictor must also be transparent: features of perfect hallucination and/or perception must fully explain the impossibility of distinguishing perfect hallucination from perception. Call this the transparency condition. If a proposed context is not transparent, then the indistinguishability claim could not play the two roles noted in Section 2. Suppose there is an elixir that has some magical property that prevents one from distinguishing perfect hallucination from perception, and thereby explains the impossibility of distinguishing perfect hallucination from perception. If the elixir served as the contextual restrictor, sameness in phenomenal character would not even be in the running as a possible explanation of the impossibility of distinguishing perfect hallucination from perception. Rather, the most obvious explanation would appeal to the interaction between the magical property and features of perfect hallucination and perception. Thus, adopting a context like the elixir would prevent the indistinguishability-sameness inference from structuring debate between common kind theory and disjunctivism. Moreover, if the elixir and introspection served equally well as impossibility-accommodating contexts, then it would be an arbitrary choice to characterize experience in terms of one or the other. However, philosophers would not characterize experience in terms of indistinguishability relative to the elixir. What is wrong with the context of the elixir is that it makes an explanatory contribution to the impossibility of distinguishing perfect hallucination from perception over and above the contributions of perfect hallucination and/or perception. The issue
is resolved, however, if we select a context – introspection – that allows features of perfect hallucination and perception to alone determine the impossibility to distinguish the states.

Introspection is thus chosen as the impossibility-accommodating context because it is pictured as a kind of see-through medium. Because it is a medium through which it is impossible to distinguish perfect hallucination from perception, introspection explains why it is impossible to distinguish perfect hallucination from perception. But since introspection is see-through, it allows perfect hallucination and perception alone to determine the impossibility. On the one hand, introspection is similar to the contexts of testimony, inference, and the elixir insofar as all help to explain either the possibility or impossibility of distinguishing perfect hallucination from perception. But, unlike testimony, inference, and the elixir, introspection is transparent, permitting features of perfect hallucination and perception to be solely responsible for the impossibility of distinguishing the one from the other. Introspection – pictured as a kind of see-through medium – seems uniquely placed to satisfy both the explanatory and transparency conditions.

4. Attempts to Thread the Needle: Two Models of Introspection

I will now argue that there is no model of introspection that can satisfy the explanatory and transparency conditions. I consider observational and non-observational models of introspection, showing that neither can successfully satisfy both conditions. The acquaintance and inner sense models are both observational approaches to introspection (e.g. Russell 1997; Armstrong 1993). The constitutive model is a non-observational approach to introspection (e.g. Shoemaker 1988; Martin 2004).
4.1. Introspection as Observation

According to the acquaintance model, in introspection we are acquainted with phenomenal properties of experience. Acquaintance is a non-causal, non-inferential, and unmediated form of awareness (Russell, 1997) or experiencing (Conee 1994: 140-141). A more theoretically loaded way of unpacking the relation of acquaintance is in terms of constitution or embedding (Chalmers 2003: 235; Gertler 2001: 307-309). The basic idea of this embedding view is that the content of introspection refers to a phenomenal property in virtue of that content embedding or being partly constituted by that phenomenal property. Gertler (2001) claims that embedding is made possible by demonstrative attention to phenomenal properties (see also Giustina 2022: 128). From the point of view of acquaintance models, then, acquaintance is the context relative to which it is impossible to distinguish perfect hallucination from perception. In contrast to acquaintance models, inner sense models construe introspection in terms of a causal relation. For example, Armstrong (1968: 324) holds that introspection is akin to a kind of scanning process that acquires information about underlying sense experience. According to inner sense models, then, ‘scanning’ is the context relative to which it is impossible to distinguish perfect hallucination from perception.

Though these characterizations are brief, what matters for our purposes is just that both models of introspection are observational in the following sense: the content of introspection – how things seem from the perspective of introspection – is determined by a mental state whose existence and nature is independent of that content. This is most clearly illustrated in the embedding view, according to which phenomenal properties embed the content of introspection, but the nature and existence of the phenomenal properties do not depend on being so embedded (e.g. Gertler 2001: 308). Observational theorists thus picture introspection as set upon a reality whose existence and nature are independent of introspection.
Though many find this observational aspect of acquaintance and inner sense models attractive, it renders the explanatory and transparency conditions flatly incompatible. It cannot be that direct awareness (Russell) or scanning (Armstrong) of perfect hallucination explains the impossibility (explanatory condition), and also that features of perfect hallucination alone explain the impossibility (transparency condition). If the relatum of the relation of introspection alone explains the impossibility, then there is no room for the relation itself to do explanatory work. Indeed, any model that holds that introspection and perfect hallucination are distinct cannot possibly satisfy both the explanatory and transparency conditions. Perfect hallucination cannot alone explain the impossibility while also making room for something else to explain the impossibility.

The only way to reconcile the two conditions, then, is to deny that introspection and perfect hallucination are distinct. A proponent of the embedding view might point out that this is just what that view does. Recall that on the embedding view the content of introspection is in part constituted by phenomenal properties. Since phenomenal properties of perfect hallucination are constituents of the content of introspection, introspection itself – the introspective state with its content – explains the impossibility. And at the same time, the embedding view can allow that phenomenal properties of perfect hallucination alone explain the impossibility. Introspection explains the impossibility (explanatory condition) in virtue of one of its constituents explaining the impossibility (transparency condition).

Despite appearances, however, the incompatibility remains. The embedding view conceives of introspection as a relation between an introspective state and an introspected state (e.g. Gertler 2001: 308). However, as mentioned earlier, if the relatum – the introspected state – alone explains the impossibility of distinguishing perfect hallucination from perception, then no
explanatory work could be done by the relation itself. For example, the following relational features could not explain the impossibility: the fact that perfect hallucination is embedded in the content of introspection, the fact that the content of introspection refers to perfect hallucination, the fact that introspection involves demonstrative attention to perfect hallucination, and so on. But these are just the sort relational features that need to do explanatory work if the explanatory condition is to be satisfied. I conclude that the explanatory and transparency conditions are mutually incompatible on an observational model of introspection.

4.2. Introspection as Non-Observational

We have seen that in order to satisfy both the explanatory and transparency conditions we need to deny that introspection and perfect hallucination are distinct. The problem with the embedding view stemmed from its claim that the content of introspection is constituted by phenomenal properties of perfect hallucination. However, a solution presents itself if we flip things around: instead of perfect hallucination constituting introspection, we should instead view introspection as constituting perfect hallucination. Indeed, Martin (2004, 2006) offers a constitutive model of introspection on which the perspective of introspection exhaustively constitutes the nature of (at least some) perfect hallucinations.\(^5\) Consider Martin (2006: 392):

However things seem from the subject’s perspective with respect to her phenomenal consciousness is how phenomenal consciousness must be….The subject’s perspective on her own sense experience constitutes sense experience being that way for her…

In contrast to observational models, on Martin’s constitutivist view, the nature of perfect hallucination is not fixed independent of one’s introspective take on the experience. Introspection

\(^5\) The idea that the perspective of introspection exhaustively constitutes the nature of perfect hallucination stems from Martin’s response to the causal argument from hallucination. In response to that argument, Martin (2006: 369) claims that ‘…there is no more to the phenomenal character of [at least some perfect hallucinations] than that of being indiscriminable from corresponding visual perceptions…’. In order for there to be ‘no more’ to the nature/character of perfect hallucination than its indistinguishability from perception, it must be that the introspective perspective exhaustively constitutes perfect hallucination in the sense that I argue for in this Section.
is not pictured as set upon a reality whose existence and nature are independent of the perspective of introspection. Instead, according to Martin, one’s introspective take on the experience exhaustively constitutes the nature of perfect hallucination.

This reversal allows Martin’s view to satisfy both the explanatory and transparency conditions. Martin’s view satisfies the explanatory condition because his view can allow that the impossibility of distinguishing perfect hallucination from perception is explained in terms of the perspective of introspection. But the view also satisfies the transparency condition. Since the introspective perspective exhaustively constitutes perfect hallucination, the impossibility of distinguishing perfect hallucination from perception is explained in terms of the nature of perfect hallucination alone. Perfect hallucination alone explains the impossibility (transparency condition) in virtue of the perspective of introspection explaining the impossibility (explanatory condition). By having the perspective of introspection exhaustively constitute the nature of perfect hallucination, Martin’s model seems to satisfy both the explanatory and transparency conditions.

However, as I shall now argue, the model does so only by sacrificing a central feature of the modal interpretation. A given perfect hallucination can at the same time possess two properties: the property of being introspectively indistinguishable from perception and the property of being testimonially distinguishable from perception. There is supposed to be no clash because each property carves out a different context – introspection and testimony, respectively – relative to which perfect hallucination is either indistinguishable or distinguishable from perception. However, if we assume Martin’s constitutive model of introspection, then there is a clash between these properties.

On the current model, the perspective of introspection constitutes perfect hallucination. But what could it mean for a perspective – the perspective of introspection – to constitute perfect
hallucination? Martin (2006: 392) provides the answer: ‘*however things seem* from the subject’s perspective with respect to her phenomenal consciousness is how phenomenal consciousness must be…’ (italics added). In effect, according to Martin, the content of the introspective perspective – ‘*however things seem*’ – fixes the nature of perfect hallucination. In other words, the perspective of introspection constitutes perfect hallucination in the sense that if, from the perspective of introspection, some perfect hallucination is $F$, then the perfect hallucination is $F$ *tout court*. Whatever introspection says also goes. Now, from the perspective of introspection, *perfect hallucination is impossible to distinguish from perception* – this content is our relevant substitution instance for $F$. But, given what it means for the introspective perspective to constitute perfect hallucination, this just means that perfect hallucination is impossible to distinguish from perception *tout court*. However, this results in a clash with testimonial distinguishability: it cannot both be that perfect hallucination is impossible to distinguish from perception *tout court* and possible to distinguish from perception through testimony.

The constitutivist may try to avoid this conclusion by revising the content of introspection. They might argue that the content of introspection is not that perfect hallucination is impossible to distinguish from perception. Rather, the content of introspection is that perfect hallucination is impossible to distinguish from perception *relative to introspection*. The rejoinder works by having the perspective of introspection show up *twice over*: once as the perspective constitutive of perfect hallucination and again as an element in the content of the perspective. Thus, if we substitute $F$ with the revised content – that perfect hallucination is impossible to distinguish from perception relative to introspection – it turns out that perfect hallucination is impossible to distinguish from perception relative to introspection. The result is that the clash with testimonial distinguishability
is avoided: it can both be the case that perfect hallucination is impossible to distinguish from perception through introspection and possible to distinguish from perception through testimony.

The problem with this response is that introspection, as it figures in the revised content, is either itself a constitutive perspective or it is not. If it is a constitutive perspective, then its content – that perfect hallucination is impossible to distinguish from perception – fixes the nature of perfect hallucination. The result is the previously observed clash between introspective indistinguishability and testimonial distinguishability. However, if introspection, as it figures in the revised content, is not a constitutive perspective, then the response turns on an ambiguity in the role designated to the introspective perspective: on the one hand, the perspective constitutes perfect hallucination, and yet on the other hand, the perspective would not constitute perfect hallucination, serving instead as a mere context relative to which perfect hallucination is impossible to distinguish from perception. Introspection must play one or another of these roles, but cannot play both.

4.3. The Dilemma

The particular problems with the observational and non-observational models bring into relief a dilemma facing any attempt to develop a model of introspection that satisfies both the explanatory and transparency conditions. On the modal interpretation, contextual restrictors – testimony, inference, introspection – are all pictured as external vantages on perfect hallucination, not fixing the nature of perfect hallucination, but determining whether it is possible or impossible to distinguish perfect hallucination from perception. Once this picture is in place, however, it is hard to see how the explanatory and transparency conditions could be compatible: it cannot be that the external vantage explains the impossibility while perfect hallucination and/or perception alone
explain the impossibility. The solution just considered rejects the idea that introspection is an external vantage on perfect hallucination, instead picturing introspection as moulding the nature of perfect hallucination. However, as I have just argued, this solution eliminates the mark of relativity implicit in the notion of a contextual restrictor. Once introspection is pictured as moulding the nature of perfect hallucination, how things are ‘relative’ to introspection ends up just being how things are with perfect hallucination *tout court*.

5. **Towards Alternatives: Generic and Actuality Interpretations**

The debate between the common kind theorist and disjunctivist implicates indistinguishability in two ways. First, there is disagreement over whether the indistinguishability of perfect hallucination from perception is best explained via sameness in nature/character. Second, both sides share a concept of experience characterized, at least in part, in terms of its indistinguishability from perception. The lesson from the previous Section is that the debate between the common kind theorist and disjunctivist imposes incompatible constraints – the explanatory and transparency conditions – on the context relative to which perfect hallucination is impossible to distinguish from perception. I suggest that the version of the indistinguishability claim that is supposed to underlie the debate is unworkable.

I will now argue that a proper understanding of the indistinguishability claim involves rejecting the modal interpretation: indistinguishability does not imply an impossibility to distinguish. Alternatives to the modal interpretation open up once we see that claims of indistinguishability are claims of *inability*. The indistinguishability claim concerns an inability to do something, namely, an inability to distinguish hallucination from perception. I will argue that the inability to distinguish does not imply an impossibility to distinguish by appeal to the semantics
of ability ascriptions put forward by Bhatt (1999), who views a subset of such ascriptions in non-modal terms. The ability ascriptions that concern us exhibit what Bhatt calls ‘the phenomenon of disappearing modality’.

On what Bhatt calls the *actuality entailment reading*, ‘S is able to F’ entails that S manages to F. This reading comes out in the following example from Bhatt (1999: 174):

(1) Last night, a masked assailant attacked me on my way home. I was able to wrestle him to the ground. #But I didn’t do anything since I am a pacifist.

The use of ‘able’ entails that the speaker manages, in fact, to wrestle the assailant to the ground. Bhatt restricts the actuality entailment reading to ability attributions that use the past perfective as in (1). However, Bhatt’s restriction is not well motivated, since ability attributions in the present tense can also generate the actuality entailment. Take the following example:

(2) I am able to see the bird.

In a standard context, if a subject responds to the question ‘Are you able to see the bird?’ with (2), then her affirmative answer is false if she does not actually see the bird, but it is nevertheless possible for her to see the bird. Despite not being in the past perfective, (2) has an actuality entailment: it entails that I actually see the bird. The actuality entailment reading is thus one way in which the phenomenon of disappearing modality can be realized.\(^6\)

According to Bhatt, there is also is a *generic reading* of ‘to be able’. On this reading, ‘S is able to F’ entails that S generally manages to F. Consider the following example:

(3) I am able to eat five apples in an hour.

According to Bhaat, (3) means that I generally manage to eat five apples in an hour. The hallmark feature of a generic statement is that it is tolerant of exceptions. Thus (3) is true even if there are

\(^6\) One might object that the actuality implication is not an entailment of ‘is able to F’, but an implicature. However, Bhatt presents evidence from French and Hindi showing that it is an entailment (see also Hacquard (2020) for further discussion).
times when I am not able to eat five apples in an hour despite trying. Although Bhaat focuses attention on ability ascriptions that employ the past imperfective, the generic reading is also available when it comes to ability attributions that employ the present tense. Thus, the generic reading of ‘to be able’ represents a second way of realizing the phenomenon of disappearing modality.

Although Bhatt’s discussion is devoted to ability attributions, there is nothing to bar application of Bhatt’s two readings to inability attributions. According to the actuality entailment reading, ‘S is unable to F’ entails that S does not manage to F, while on the generic reading, ‘S is unable to F’ implies that S does not generally manage to F. Suppose that we are atop a mountain overlooking a valley, but some trees are in our way. You spot a bird and in response I utter:

(4) I am unable to see the bird.

(4) does not assert that it is impossible for me to see the bird. (4) is true even if it is possible for me to see the bird, but I do not now see it. This suggests that (4) has an actuality entailment reading: that, in fact, I do not now see the bird.

Bhaat’s generic reading also applies to inability attributions. Though some inability attributions use the word ‘unable’, a number of inability attributions take the form of adjectives with an ‘-able’ or ‘-ible’ suffix. Such attributions are particularly apt for our purposes, since attributions of indistinguishability often take this form. Consider the following inability attribution modified by a degree word:

(5) The undergrowth is also thicker and more impenetrable, with a well-developed moss and lichen layer, interspersed with low shrubs, large lianas and monocarpic bamboos.7

---

7 Retrieved via the iWeb corpus on 21 July 2023 and sourced from https://www.wildmadagascar.org/overview/forest_types.html
(5) does not assert that the undergrowth is more impossible to penetrate than other forms of undergrowth. Indeed, it is not clear what it would mean for one state of affairs to be more impossible than another. This can be seen if we delete the comparative aspects of (5):

(6) The undergrowth is also thick and impenetrable, with a well-developed moss and lichen layer, interspersed with low shrubs, large lianas and monocarpic bamboos.

The undergrowth can be impenetrable while it is nevertheless routinely penetrated by, say, snakes, spiders, and hell-bent humans. I suggest that (6) asserts that the undergrowth generally fails to be penetrated by creatures, but not that it is impossible for creatures or any subset of creatures to penetrate the undergrowth. I suggest that a generic reading is often most natural when it comes to inability attributions with an ‘-able’ or ‘-ible’ suffix, e.g. ‘inflexible’, ‘inscrutable’, and so on.

The actuality and generic readings of inability claims pave the way for us to interpret the indistinguishability claim non-modally.

5.1. The Generic Interpretation

The hallmark of generics is that they are true despite the existence of exceptions. Indistinguishability claims seem to have this property. Consider:

(7) This piece of soap is indistinguishable from a real lemon.

(7) is true despite the fact that pieces of soap shaped and coloured like a lemon are often easy to tell apart from a lemon. For example, I can take a knife to it and know, by the lack of juice, that it is not a lemon. One may respond that the proper interpretation of (7) mentions the context relative to which it is impossible to distinguish the piece of soap from a lemon. The context in question would serve to bracket the way one came to distinguish the piece of soap. One might furnish such a context as follows:

(8) This piece of soap is indistinguishable from a real lemon through sight and without impingement.
But (8) too is true despite the existence of exceptions. For example, I can move the piece of soap under a lamp, and without impinging it tell by sight that it is not a lemon – its look under bright light is just too waxy. I think the proper reaction to sentences (7) and (8) is that claims of indistinguishability admit of a generic interpretation. On this view, \( x \) is indistinguishable from the \( F \)s despite the existence of exceptions in which one does distinguish \( x \) from the \( F \)s. Just as it is true that tigers are striped despite the existence of exceptions (some tigers have no stripes), so too claims of indistinguishability can be true despite the existence of exceptions. And there is nothing that bars sometimes giving the indistinguishability claim a generic reading:

**The Generic Interpretation:** \( x \) is indistinguishable from a perception of an \( F \) despite the possible existence of cases in which \( x \) is distinguished from a perception of an \( F \).

Notice that in providing the generic interpretation I have not appealed to the semantic structure of generic statements. This is because there are a number of vexed issues about how to properly analyse generics and I wish to remain neutral concerning them in this paper (e.g. Leslie and Lerner 2022). Instead, I have characterized the generic interpretation in terms of a property, being tolerant of exceptions, that all sides agree characterize generic statements.

### 5.2. The Actuality Interpretation

Above we considered the claim that a piece of soap is indistinguishable from a lemon (sentence (7)). We then discovered exceptions. We blocked these exceptions by the use of add-on conditions (sentence (8)). But we found exceptions to the resulting claim as well. In response, the supporter of the modal interpretation might dig in her heels by further specifying the context relative to which the piece of soap is indistinguishable from a lemon. Here is an option:

(9) This piece of soap is indistinguishable from a real lemon through sight and without impingement and **without moving the relevant object from where it is and without moving any other object into the scene**.
However, exceptions to (9) are also discoverable. It might be that the piece of soap ends up falling off its support, rolling across the floor (without impingement), and we see that its roll is a relatively slow, waxy one. Through sight alone we thus tell that it is not a lemon. The reader sees where this is going: for every add-on condition offered, there will be an exception where one tells apart the soap from the lemon.

What underlies the futility of adding on conditions? It is notable that the add-on conditions are negative conditions. In particular, the conditions specify ways in which the context is not to be altered relative to the context at which one actually fails to distinguish the piece of soap from a lemon. This suggests that if the full cascade of add-on conditions is specified it will just amount to the purely negative condition given in (10):

(10) This piece of soap is indistinguishable from a real lemon through sight if there is no change in the context at which one actually fails to distinguish them.

But what could explain the impossibility of distinguishing the piece of soap from a lemon as it arises in (10)? It cannot be features of the context at which one actually fails to distinguish the piece of soap from a lemon. The lesson from (7)-(9) is that whatever those features happen to be there will always be an exception in which one succeeds in distinguishing the piece of soap from a lemon at a context characterized in terms of them. The only other candidate to explain the impossibility of distinguishing the piece of soap from a lemon is the actual failure to distinguish the piece of soap from the lemon. But the actual failure to distinguish cannot explain the impossibility of doing so. What, then, accounts for the truth of sentences like (10)?

I suggest (10) is true because it has an actuality entailment. Notice that there are clear instances where claims of indistinguishability have an actuality entailment. For example, suppose two black luggage bags – mine and another’s – are coming down the airport carousel. I am going
back and forth in my head trying to assess which one is mine. The following claim is true in the context:

(11) I am unable to tell the two bags apart.

(11) is true despite the fact that it is common knowledge that it is possible to distinguish the two bags at that very context. This makes sense only if (11) has an actuality entailment. (11) entails that I do not manage to distinguish the two bags. I suggest, then, that some claims of indistinguishability – including (10) – have an actuality entailment. There is nothing that bars reading the indistinguishability claim as sometimes having an actuality entailment:

**The Actuality Interpretation:** $x$ is indistinguishable from a perception of an $F$ if and only if there is an actual failure to distinguish $x$ from a perception of an $F$.

On the actuality interpretation, $x$ being indistinguishable from a perception of an $F$ is consistent with it nevertheless being possible to distinguish $x$ from a perception of an $F$. This is because the actuality interpretation entails only that one actually fails to distinguish $x$ from a perception of an $F$.

6. Implications

I have presented a two-pronged argument against the version of the indistinguishability claim that is supposed to underlie the debate between the common kind theorist and disjunctivist. Section 4 argued the explanatory and transparency conditions are not jointly satisfiable. Section 5 argued that we cannot interpret claims of indistinguishability in a manner that brackets all exceptions. Instead, we found that claims of indistinguishability are amenable to generic and actuality readings. The two-pronged argument has a number of significant consequences for the issues discussed in Section 2, three of which I explain across Sections 6 and 7.
6.1. Undermining the Indistinguishability-Sameness Inference

Recall the indistinguishability-sameness inference from Section 2: if perfect hallucination is indistinguishable from perception through introspection, then perfect hallucination and perception share the same fundamental nature/character. The common kind theorist and disjunctivist disagree about whether we are justified in inferring the conclusion from the premise. According to both camps, the question is whether an epistemic feature of the relationship between perfect hallucination and perception – impossibility to distinguish – justifies us in inferring an ontological relationship between the two – sameness in nature/character. The common kind theorist answers in the affirmative, since, according to them, the best explanation of the impossibility is that the two kinds of experience in fact share the same nature/character (e.g. Price 1932: 31-32; Tye 2009: 560). Disjunctivists resist the inference by appeal to cases in which it is impossible to distinguish objects or states that nevertheless differ in nature/character (e.g. Austin 1964: 50; Martin 2008: 91-92). The modal interpretation of the indistinguishability claim thus sustains an important debate between common kind theorists and disjunctivists.

In contrast, the actuality and generic interpretations cannot sustain debate over the indistinguishability-sameness inference. If we assume the actuality interpretation, we get the following inference: if one actually fails to distinguish a hallucination from perception, then the experiences share the same nature/character. The problem is that, from the point of view of the actuality interpretation, it is possible that in the next moment the subject actually succeeds in distinguishing the experiences. But now we are in an awkward position: when the subject actually fails, the experiences possess the same nature/character, but when the subject actually succeeds, the same experiences do not possess the same nature/character. Not even the common kind theorist would put forward such an unstable inference.
Now consider the generic interpretation. If we assume the generic interpretation, we get the following inference: if one generally fails to distinguish a hallucination from perception, then the experiences share the same nature/character. As explained in Section 2, the common kind theorist claims that sameness in nature/character best explains the indistinguishability of perfect hallucination from perception. However, the sameness in nature/character could not explain the generic interpretation of the indistinguishability claim. In particular, the fact that perception and perfect hallucination share the same nature/character cannot explain how it is that one sometimes succeeds but sometimes fails at distinguishing perfect hallucination from perception. The sameness in nature/character can only explain why it is impossible to distinguish perfect hallucination from perception. Not even the common kind theorist would put forward the indistinguishability-sameness inference if indistinguishability were interpreted as a generic. Thus, if we reject the modal interpretation in favour of the actuality and generic interpretations, the indistinguishability claim can no longer ground debate between common kind theorists and disjunctivists.

6.2. Undermining Philosophical Notions of Hallucination and Illusion

Philosophers often draw a distinction between ordinary or psychologists’ hallucinations and perfect or philosophers’ hallucinations (Robinson 2022; Byrne and Manzotti 2022: 332). Although a large part of the literature assumes that perfect hallucinations are possible, there is a small minority who question its empirical possibility. These discussions question existing psychological evidence that such hallucinations occur or else attempt to explain away putative cases of philosophers’ hallucinations (Fish 2009; Masrour 2020; see also Byrne and Manzotti 2022). Unlike these other discussions, however, this paper suggests a route toward arguing against the metaphysical possibility of perfect hallucination. Philosophers of perception characterize
perfect hallucination, at least in part, in terms of the modal interpretation of the indistinguishability claim. However, we found that we were unable to find a contextual restrictor that could satisfy the explanatory and transparency conditions. This suggests that perfect hallucination, in the sense intended by philosophers of perception, is not metaphysically – indeed, not even conceptually – possible.

But this is not to say there is no viable notion of hallucination, since, as I shall now argue, the generic and actuality interpretations put into relief our notion of ordinary hallucination. Subjects who undergo hallucinations sometimes have insight into their hallucinatory status. Insight is characterized, in part, as the recognition that one’s behaviours and experiences are caused by one’s disorder (David 1990). In the literature on insight and hallucination, it is observed that those with more persistent auditory hallucinations have less frequent insight into their unreality (Lera et al. 2011). Nevertheless, there are many cases in which those with persistent external hallucinations do have insight into their unreality. Aubrey Lewis (1934: 346) recounts such a case:

...his appearance changed greatly, he appeared frightened and dejected and said, “I had a terrible feeling, I thought my mother was dead. I heard a voice say in my ear ‘Mother’”. Asked if there was anything in it really, he answered, “No, I just got that impression. I do get these funny sort of feelings, sometimes.”

Cases of insight into the unreality of hallucination are relatively frequent (Sass and Pienkos 2013). Despite such insight, it remains true that hallucinations are indistinguishable from perception. This makes sense only if the indistinguishingability claim is interpreted as a generic: it is true that hallucinations are indistinguishable from perception, even though it is also the case that hallucinations are frequently distinguished from perception. The generic interpretation captures both of these aspects of ordinary hallucination.

The shift from the modal to the generic and actuality interpretations has a bearing on illusion as well. Just as we distinguished between perfect and ordinary hallucination, we can also distinguish between perfect and ordinary illusion. Perfect illusion is characterized, at least in part,
in terms of the modal interpretation: perfect illusion is impossible to distinguish from perception through introspection. The problems with the modal interpretation are equally applicable to the case of perfect illusion. But this does not mean that there is no viable notion of illusion, since the generic and actuality interpretations capture our notion of ordinary illusion: ordinary illusions are indistinguishable from perceptions despite often being told apart from perceptions. Consider Austin’s (1964: 49) observation:

…it is simply not true to say that seeing a bright green after-image against a while wall is exactly like seeing a bright green patch actually on the wall; or that seeing a white wall through blue spectacles is exactly like seeing a blue wall… or (once again) that seeing a stick refracted in water is exactly like seeing a bent stick. In all these cases we may say the same things (‘It looks blue’, ‘It looks bent’, &c.), but this is no reason at all for denying the obvious fact that the ‘experiences’ are different.

One way of interpreting Austin here is that he is agreeing that illusions are indistinguishable from perception – we say things like ‘It looks blue’, ‘It looks bent’, and so on – but that it is nevertheless palpable to the subject that they differ from corresponding perceptions. This would make sense only if the generic interpretation captured our notion of ordinary illusion: illusions are indistinguishable from perception even though they are routinely told apart from perceptions. In general, I suggest, there are two aspects of our notions of ordinary hallucination and illusion: on the one hand, hallucination and illusion are indistinguishable from perception, while on the other hand, hallucination and illusion are frequently told apart from perception. The generic and actuality interpretations capture these two aspects and thus limn the contours of these ordinary notions.

7. Getting at the Rootlets of ‘Experience’

J.M. Hinton noted that philosophers who deny the existence of sense-data often also claim that perception and perfect hallucination/illusion share the same phenomenal character. Hinton took these philosophers as his target:
It is certainly strange that people should readily agree that one must not introduce a private Polly-Ann who is seen whenever Polly or Ann is seen, and yet go on introducing – or at least not extraditing – a private Polly-Ann-sighting that is done whenever Polly or Ann is sighted: a quasi-hallucination, for which one must no more supply an actual object than for a genuine hallucination, but which occurs whenever one has a veridical experience. (1967b: 227)

Or again:

The view that in all cases of seeing (and visual hallucination) we actually see a mental picture or image is now widely rejected, and I for one would also want to reject the marginally different view that a sort of as-it-were-picture-seeing occurs as a common constituent of illusion and true perception. (1967a: 10)

According to Hinton, the supposition that perception and perfect hallucination share the same phenomenal character – that in both cases there is ‘a private Polly-Ann-sighting’ or an ‘as-it-were picture-seeing’ – is a hangover of the long-since-rejected idea that one is aware of sense-data in experience (Byrne and Logue 2008). If a philosopher rejects sense-data, according to Hinton, she should also reject the idea that perception and perfect hallucination share the same phenomenal character. In making good on this directive, Hinton claimed that perception and perfect hallucination do not possess the same phenomenal core: there is no ‘private Polly-Ann-sighting’ or ‘as-it-were-picture-seeing’ common to both disjuncts. In offering what came to be known as ‘disjunctivism’, Hinton took himself to be undermining a background notion of experience animating discussions of perception at the time.

I hope that I am getting at some of the rootlets that keep the shattered trunk of ‘visual experiences’ standing, though deeply undermined and pruned of its topmost sense-datum twigs. (1967b: 227)

By rejecting the idea that perception and perfect hallucination share the same phenomenal character, Hinton took himself to be ‘get[ting] at the rootlets’ of an objectionable concept of experience.

From the perspective of this paper, however, Hinton failed to see that the root system that sustains an objectionable concept of experience goes even deeper. As explained in Section 2, parties to the problem of perception assume that experience is defined, at least in part, by the property of being impossible to distinguish from perception through introspection. The common
kind theorist takes the property to be necessary but not sufficient for counting as an experience, while the disjunctivist takes the property to be both necessary and sufficient. However, if this property is not workable, as the paper has argued, we must reject the background notion of experience animating both common kind theory and disjunctivism. It then remains an open question what would count as a viable notion of experience.\(^8\)

Shivam Patel

Florida State University, USA

\(^8\) Many thanks to Zina Ward, John Schwenkler, Nat Stein, Umrao Sethi, Casey Doyle, and students in my graduate seminar on the problem of perception at Florida State University for comments on previous drafts.
References


