



HAL
open science

Relationalism in the face of hallucinations

Roberta Locatelli

► **To cite this version:**

Roberta Locatelli. Relationalism in the face of hallucinations. Philosophy. Université Panthéon-Sorbonne - Paris I; University of Warwick (Coventry, Royaume-Uni), 2016. English. NNT: 2016PA01H213 . tel-01580461

HAL Id: tel-01580461

<https://theses.hal.science/tel-01580461>

Submitted on 1 Sep 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Université Paris 1 Panthéon Sorbonne/University of Warwick

Thèse pour l'obtention du grade de docteur de l'université Paris 1
et de l'université de Warwick

présentée et soutenue publiquement par

Roberta Locatelli

**RELATIONALISM IN THE FACE OF
HALLUCINATIONS**

**LE RELATIONNALISME FACE AUX
HALLUCINATIONS**

Directeurs de Thèse :

Maximilian Kistler, Professeur à l'UFR de philosophie,
université Paris 1 Panthéon Sorbonne, IHPST,

Matthew Soteriou, Professeur à l'université de Warwick, Department of
Philosophy
(maintenant muté à King's College London)

Naomi Eilan, Professeur à l'université de Warwick, Department of
Philosophy

Composition du jury:

Fabian Dorsch (université de Fribourg, Suisse), Professeur

Jérôme Dokic (EHESS), Directeur d'Etudes

Matthew Nudds (University of Warwick), Professeur

Mark Kalderon (University College London), Professeur

Maximilian Kistler (université Paris 1), Professeur

Date de soutenance : 15 décembre 2016

Content

Acknowledgements.....	v
Declaration	vi
Abstract	vii
Introduction	1
Chapter 1: The Relational View of Perception.....	12
1.1 Trivial and Not So Trivial Claims about Experiences	13
1.2 How Experience Strikes Us.....	21
1.3 Relational Views of Phenomenal Character.....	31
1.4 Open Questions and Terminological Distinctions.....	33
1.5 A Problematic Reasoning.....	39
1.6 The Relationalist's Commitment to Disjunctivism	43
1.7 Sense-datum Theories, Representationalism, Qualia Theories	46
1.8 Conclusions	51
Chapter 2: The Argument from Hallucination	52

2.1 The Base Case: Conflicting Appearances, Illusions and Hallucinations.....	55
2.1.1 <i>Conflicting appearances and Paradigmatic Looks</i>	55
2.1.2 <i>Optical Illusions and the Phenomenal Principle</i>	62
2.1.3 <i>Hallucinations</i>	69
2.2. The Spreading Step	70
2.3 The Indistinguishability Principle.....	75
2.4 Conclusions	83
Chapter 3: How to Be Superficial.....	86
3.1 The Purported Infallibility of Introspection	88
3.2 Superficial Properties.....	96
3.3 Observational Properties	101
3.4 The Superficiality Constraint	107
3.5 Impersonal Indiscriminability.....	112
3.6 Conclusion.....	124
Chapter 4: Hallucinating, for Real.....	127
4.1 Instinguishable from What?.....	128

4.2 Hallucination: A Fuzzy Notion	135
4.3 What Counts as Impersonally Indistinguishable from Perception	139
4.4 The Problem of Hallucination is Not a Phenomenological Problem	145
4.5 Conclusions.....	150
Chapter 5: Philosopher’s Hallucinations	152
5.1 A Metaphysical Possibility	154
5.2 Experiential Naturalism.....	162
5.3 Local Supervenience	167
5.4 Is there Empirical Support for Local Supervenience?	179
5.5 The Causal Argument from Hallucination	185
5.6 Causation and Constitution.....	196
5.7 Where Does This Leave Us?	214
5.8 Conclusions.....	222
Chapter 6: Introspecting Hallucinations	224
6.1. An Invalid Argument?.....	226
6.2 The Reversed Causal Argument	232

6.3 Attempts to Avoid the Screening Off Problem.....	240
6.4 The Negative Account of IIIHs.....	248
6.5 Objections and Clarifications	251
6.5.1 <i>The Scope of the Negative Account</i>	251
6.5.2 <i>The Alleged Need for a Positive Account</i>	254
6.5.3 <i>Very Bad Cases and Unsophisticated Hallucinators</i>	257
6.6 The Phenomenal Character as a Ghost.....	264
6.7 Introspecting	275
6.8 Response to Critics and How to Preserve Superficiality.....	280
6.9 Conclusions.....	285
Bibliography	293
Résumé de Thèse en Français.....	305

Acknowledgements

Many thanks to my supervisor Matthew Soteriou for his invaluable help and advice throughout the writing of this thesis. I can't imagine accomplishing this work without his extremely generous and insightful guidance, challenging and encouraging at once. Also thanks to my supervisor from Paris 1, Maximilian Kiestler, who made the co-tutelle between Paris 1 and Warwick possible, and to Naomi Eilan, who masterminded the delicate and laborious process of setting up the co-tutelle in the initial and final stages, and whose warm welcome at Warwick made me feel at home from day one.

I benefitted immensely from supervisions with Mike Martin and Paul Snowden during my research stay at the University of London (Institute of Philosophy) before I enrolled at Warwick. Mike continued reading some of my work afterwards, always providing sharp and penetrating comments.

I am grateful to Hemdat Lerman, who, with her astonishing eye for detail, took the time to read and thoroughly discuss earlier drafts of some chapters; and to Ivan Ivanov, for reading and commenting on some chapter drafts, and for countless philosophical discussions. Joe Cunningham, Anna Savva Grinevich, Nicole Osborne, Jack Shardlow, and Barnaby Walker read and commented on chapters or sections of this thesis, and provided useful stylistic advice, which helped to make this thesis a more pleasant read. The responsibility for any inaccuracies or stylistic errors remaining herein is my own.

Thanks also to Bill Brewer, Wesley Chai, Tony Cheng, Tom Crowther, Joe Cunningham, Jérôme Dokic, Juan Camilo Espejo-Serna, Craig French, Christoph

Hoerl, Guy Longworth, Sofia Miguens, Matthew Nudds, Michele Ombrato, Sebastian Sanhueza, Susanna Schellenberg, Irina Schumski, Maarten Steenhagen, Charles Travis, Daniel Vanello, Alberto Voltolini, Barnaby Walker, Sander Werkhoven, Keith Wilson, and Hong Yu Wong for their helpful discussions, suggestions, and support along the way.

I would also like to express my gratitude to Paolo Spinicci and Clotilde Calabi, from the University of Milan, who initiated me to the delights and pains of the philosophy of perception, and to Jocelyn Benoist, who supported me in many ways when he was my supervisor at Paris 1, before I crossed the English Channel. Finally, thanks to Matteo Fiorillo for his loving patience, and to my parents, for their abiding support.

Declaration

I hereby declare that this thesis is the result of original research carried out by me while in candidature for a doctoral degree at the University of Warwick and at the University of Paris 1 in the context of a cotutelle agreement; that I am the sole author and researcher and that the work presented herein is entirely my own; that no portion has been submitted for a degree or qualification at this or any another university; and that all source materials have been clearly and adequately acknowledged and cited.

Abstract

Relationalism claims that the phenomenal character of perception is constituted by the obtaining of a non-representational psychological relation to mind-independent objects. Although relationalism provides what seems to be the most straightforward and intuitive account of how experience strikes us introspectively, it is very often believed that the argument from hallucination shows that the view is untenable. The aim of this thesis is to defend relationalism against the argument from hallucination. The argument claims that the phenomenal character of hallucination and perception deserves the same account, and that relationalism cannot be true for hallucinations, therefore relationalism must be rejected. This argument relies on the Indistinguishability Principle (IND), the claim that two experiences that are introspectively indistinguishable from each other have the same phenomenal character. Before assessing the plausibility of this principle, I first consider and dismiss versions of the argument which wouldn't depend on IND.

Although widely accepted, no satisfactory support for IND has been presented yet. In this thesis I argue that defending IND requires that we understand the notion of 'indiscriminability' employed in IND in an impersonal sense. I then identify what underwrites IND: the intuition that, in virtue of its superficiality, the nature of a phenomenal character must be accessible through introspection, together with the claim that it is not possible to deny IND without denying the superficiality of phenomenal characters too.

I argue that the relationalist can deny IND while preserving the superficiality of phenomenal characters. This can be done by adopting a negative view of hallucination and an account of introspection whereby the phenomenal character doesn't exist independently of one's introspective awareness of it and where having introspective access to our experience depends on our perceptual access to the world.

Resumé

Le relationnalisme affirme que le caractère phénoménal de la perception est en partie constitué par les objets que l'on voit. Malgré son attrait intuitif, il est souvent rejeté en vertu de l'argument de l'hallucination. L'objectif de cette thèse est de défendre le relationnalisme. L'argument fait valoir que, puisque le relationnalisme ne peut être vrai pour les hallucinations, il doit forcément être faux pour les perceptions aussi. En cela, l'argument repose sur le principe de l'indiscernabilité (IND), affirmant que deux expériences qui sont introspectivement indiscernables les unes des autres ont le même caractère phénoménal. Je considère et rejette les autres versions de l'argument qui ne reposent pas sur l'IND.

Bien que largement accepté, aucun support satisfaisant pour l'IND n'a encore été présenté. Dans cette thèse, je soutiens que la défense de l'IND requiert que l'on entende la notion d'indiscernabilité employée dans l'IND en un sens impersonnel. Ensuite, j'identifie ce qui motive l'IND : l'intuition que, en vertu de sa superficialité, la nature d'un caractère phénoménal doit être accessible par l'introspection, de concert avec l'idée qu'il est impossible de nier l'IND sans nier par là même la superficialité des caractères phénoménaux.

Je soutiens que le relationnaliste peut nier l'IND tout en préservant la superficialité des caractères phénoménaux en adoptant une thèse négative de l'hallucination et en reconsidérant la nature de la relation entre le caractère phénoménal d'une expérience et l'accès introspectif qu'il peut y avoir.

Keywords:

Perception, Hallucination, Illusion, Disjunctivism, Naïve Realism, Relationalism, Phenomenal Character, Introspection, Sense-data, Representationalism, Philosophy of Mind, Psychopathology of Hallucinations, Causal Argument from Hallucination, Local Supervenience, Superficial Properties, Observational properties, Indistinguishability.

Mots Clés :

Perception, hallucination, illusion, disjonctivisme, réalisme naïf, relationnalisme, caractère phénoménal, introspection, sense-data, philosophie de l'esprit, indiscernabilité, représentationnalisme, psychopathologies des hallucinations, argument casuel de l'hallucination, survenance locale, superficialité, propriétés observationnelles.

Ecole Doctorale de philosophie (ED 280) – IHPST (Institut d'histoire et de philosophie des sciences et des techniques) UMR85590

Department of Philosophy, University of Warwick

You have led me, it seems, step by step, till I am got I know not where. But I shall try to get out again, if not by the way I came, yet by some other of my own finding (Berkeley, *Alciphron* sec.10).

Introduction

Relationalism claims that the phenomenal character of perception is constituted by the obtaining of a non-representational psychological relation to mind-independent objects. It provides what seems to be the most straightforward and intuitive account of how experience strikes us introspectively.

Some relationalist commitments have long been a feature of the philosophy of perception, most often as part of so-called ‘naïve realism’, a view attributed to the unsophisticated ordinary person and used by philosophers as their target against which they frame their positions. However, only in recent years have we seen attempts to systematically present the commitments, merits and implications of relationalism, chiefly thanks to the work of Martin (1997; 2002; 2004; 2006), Snowdon (1980; 1990), Campbell (2002), Brewer (2011), Fish (2009), and Soteriou (2013).

The recent debate over relationalism has primarily focused on one of its implications: disjunctivism. Disjunctivism claims that the phenomenal character of hallucination shouldn’t be accounted for in the same way as we account for the phenomenal character of perception. This is a corollary of the relationalist view that the phenomenal character of perception is partly constituted by the mind-independent objects to which perception seems to

relate us, together with the observation that in hallucination the object to which one seems to be related doesn't exist. The term 'disjunctivism' derives from Hinton's discussion of a particular range of propositions that he calls 'perception/illusion disjunctions' (Hinton 1967a; 1967b; 1973). Hinton argues that there is no reason to believe that the general category of experience under which philosophers lump together perception and hallucination really captures an event that happens whenever one either perceives or hallucinates. Hinton claims that expressions like 'I have an experience as of a flash of light' or 'I seem to see a flash of light' is simply 'a more compact way of saying' that a disjunctive claim is true: 'Either I see a flash of light, or I have an illusion of a flash of light' (Hinton 1967b: 217).

Disjunctivism can take different forms, depending on the aspect with respect to which perception and hallucination are claimed to be distinct. There is a disjunctive position about the fundamental nature of perception and hallucination (which leaves open what determines their respective fundamental nature), that Byrne and Logue (2008) call 'metaphysical disjunctivism'. This version of disjunctivism is attributed to Snowdon (1980; 1990), who presents it in the context of a criticism of the causal theory of perception. An epistemic disjunctive claim, stating that perception and hallucination put the subject in different epistemological positions, was famously introduced by McDowell (1986, 1982, 1995). Finally, a disjunctive claim about the phenomenal character of experiences—sometimes called 'phenomenal disjunctivism'—was introduced by Martin (1997, 2002b, 2004, 2006).

I will focus on the latter, because this is the disjunctive claim that is entailed by relationalism. Accordingly, in what follows I will use ‘disjunctivism’ to refer exclusively to disjunctivism about phenomenal characters, and I will bracket all other versions of disjunctivism. There are interesting questions regarding how disjunctivism about phenomenal characters relates to the other forms.¹ However, I will leave these questions aside, as they are not relevant to the aim of this thesis, which is to defend relationalism from the argument from hallucination.

Beyond being an unavoidable corollary of relationalism, disjunctivism plays a crucial role in the strategy rationalists adopt to counter the main argument against it, the argument from hallucination. The argument seeks to establish that, since hallucinations that are indistinguishable from perceptual experience are possible, and since hallucination cannot be constituted by the obtaining of a relation to the mind-independent object one seems to be experiencing, perception cannot be a relation to the mind-independent object it seems to relate us to either.

A large part of the recent debate over relationalism has developed along the following dialectics. Opponents of relationalism present the argument from hallucination; relationalists rebut that the argument relies on

¹ Phenomenal and metaphysical disjunctivism are generally thought to help the cause of epistemological disjunctivism and some have suggested that they entail it (Byrne and Logue 2008), although it is not clear that either is required to maintain a coherent epistemological disjunctivism (see Thau 2004; Snowdon 2005; 2009; Byrne and Logue 2008; Millar 2007; 2008; Haddock and Macpherson 2008 for a discussion). Some assume that phenomenal disjunctivism entails metaphysical disjunctivism (Conduct 2011), although this is not obvious. Phenomenal disjunctivists may reject the essentialist assumption implied in the metaphysical disjunctivist claim that there is a fundamental nature of experiences, or they might not identify the ‘fundamental nature’ of an experience with its phenomenal character and what constitutes it.

an unjustified assumption: that the phenomenal characters of perception and hallucination deserve the same account. But this is exactly what relationalism, with its commitment to disjunctivism, denies. So, unless the proponent of the argument provides a motivation for prescribing a non-disjunctive account of experiences, the argument begs the question against relationalism. The proponent of the argument, then, points to the fact that perception and hallucination may be introspectively indistinguishable as a motivation for their belief that perception and hallucination have the same phenomenal character.

Often it remains unclear whether, in the intentions of proponents of the argument, the appeal to the possibility of hallucinations that are indistinguishable from perception is meant to introduce an inference to the best explanation or an indubitable principle. Sometimes, disjunctivism is rejected right from the start and not even taken seriously. This suggests that the claim that the indistinguishability of experiences is to be accounted for in terms of sameness of phenomenal characters—what I will call the indistinguishability principle—is taken to be an undeniable principle (Siewert 1998; Siegel 2011; Kriegel 2013; Deutsch 2005).

However, there is very little discussion of the underlying intuitions that supposedly make this principle undeniable. Sometimes, it is motivated on the basis of a commitment to internalism (Farkas 2003)—that relationalists are bound to reject. Some relationalists have identified the source of the indistinguishability principle in an overly optimistic view of introspective powers (Hinton 1967, 1970; Martin 1997, 2004, 2006). Some

rare explicit arguments to the effect that the only coherent way to understand the indistinguishability of experiences is in terms of the sameness of phenomenal characters have been proposed (Farkas 2006), which left relationalists unmoved. The debate over relationalism may start to look like one that cannot be adjudicated on its own terms, but only by resolving much broader controversies such as that between internalists and externalists in the philosophy of mind, discussions on the nature of consciousness, and on what being self-aware of mental occurrences amounts to.

A wide-ranging look at the current debate on the philosophy of perception leaves the impression of a field divided into two hermetic camps: those who take the indistinguishability principle to be undeniable and those who ask why this should be so. The relationalist is certainly right in lamenting the drastic lack of explicit considerations in favour of the indistinguishability principle and demanding better and explicit arguments for it. However, this complaint is likely to leave those who are already committed to the indistinguishability principle's truth unimpressed. Dialectically, relationalists can say very little to convince the opponent that their commitment to disjunctivism is benign, unless they identify the exact underlying motivations of the indistinguishability principle and respond to them. The only way to advance the debate, then, is to identify these assumptions that underwrite the indistinguishability principle, decide whether they should be accepted and whether the relationalist has the resources to accommodate them.

On the other hand, some recent discussions of relationalism tend to present the alleged indistinguishability of perception from some hallucinations as a datum that any theory of conscious perception should accommodate and which relationalism allegedly fails to accommodate, presenting the argument from hallucination as relying on an inference to the best explanation (see Sturgeon 1998; 2006; Fish 2005; 2009; B. Millar 2014; Tye 2007). This might be an effect of having metabolized relationalists' complaints that not enough motivations for the indistinguishability principle are available. However, the ambiguity between these two approaches has always been present in formulations of the argument from hallucination and still is. This has pushed the debate on relationalism away from evaluation of the validity and soundness of the argument from hallucination towards a focus on establishing whether relationalism, with its disjunctivist commitment, can provide a satisfactory account of hallucination—one that accounts for the possibility that it could be indistinguishable from perception (see for instance Johnston 2004; Smith 2008; Siegel 2004, 2008; Fish 2008, 2008; Hellie 2007, 2013; Schellenberg 2014; Tye 2007; Pautz 2010; Vega-Encabo 2010; Logue 2012; 2013).

Even framed in these terms, however, the debate seems stuck at a dead end. Many considerations go into an account of perceptual experience. Some of these are phenomenological considerations; others aren't. The relationalist might contend that her theory provides an overall better account of perceptual experience, even if it doesn't provide the best explanation of the indistinguishability datum; while her opponents seem to work with the

assumption that providing the best account of the indistinguishability datum should be the priority of any account of perception.

This lead Fish (2005) to suggest that the debate between disjunctivists (and hence relationalists committed to it) and their opponents is ‘grounded in an underlying methodological disagreement’ (Fish 2005: 121), creating a situation ‘not dissimilar to that of competing research paradigms’ (Fish 2005: 125).² The opponent of relationalism reasons that, ‘even if there are *other* differences between a perception and a hallucination, as long as they are indistinguishable from the first-person point of view, then in any case there is a common ontological core—the core that the subject is in a ‘sensory state’ of the very same kind’ (Fish 2005: 122; italics in original). Relationalists, on the other hand, ‘look *past* the first-person similarities [and] find an important *difference* between perception and hallucination—the presence or absence of a seen object’ (Fish 2005: 122; italics in original).

Both diagnoses of the debate are unsatisfactory. There is something deeply troubling in the thought that the viability of an account of the phenomenal character of perception cannot be decided on its own terms or, worse, that decisions over the truth or falsity of relationalism come down to a methodological choice where no approach shows significant merits over

² A similar suggestion is made by Nanay (2014). He recommends that we frame the debate between relationalists and representationalists in terms of a disagreement about what the necessary conditions for individuating perceptual states are, where the relationalist claims that two experiences/two perceptual states must be different if their objects are different, even if the properties attributed to these objects are the same, while the representationalist claims that if two perceptual states are the same, then the properties perceptually attributed to the perceived object must also be the same.

the other. The present work aims to solve this *impasse* and provide the tools for advancing the debate, by refocusing on the argument from hallucination, identifying what one can say to make the argument compelling and what intuitions underlie its premises. This will allow us to appreciate where the fundamental disagreement between relationalists and their opponents lie and what the relationalist can say to satisfy the constraints. Understanding where the disagreement lies first requires determining where agreement stands—what constraints are imposed on any account of the phenomenal character of experience?

I will argue that at the core of the argument from hallucination—and the crucial acceptance of the indistinguishability principle—lies the intuition that phenomenal characters are superficial properties and that there is no way for the relationalist to accommodate this feature. The aim of this research is to argue that relationalists can indeed accommodate the superficiality of phenomenal characters, although this will prove more difficult than it might initially appear. The thought that relationalists are committed to violating the superficiality constraint is tied to a misleading view of superficiality, whereby superficial properties are properties for which there is no seem/is distinction. Against this account, I will argue that reflection on colour properties shows that superficial properties can appear different than they are. However, colours can be understood as superficial inasmuch as working out that something is not the colour it seems to be can only be done through a distinctive source of information—in this case vision. I then present an understanding of superficiality whereby if a

superficial property seems different than it is, this cannot come down to the property having a real nature that is inaccessible through the dedicated source of information—in the case of phenomenal character: introspection. The real nature of the phenomenal character, then, must in principle be accessible through introspection, although not at all times.

This poses a serious challenge for relationalists: how can they maintain that the phenomenal character of hallucination is different from the phenomenal character of perception without maintaining that the real nature of the phenomenal character of hallucination is hidden to introspection? *Ex hypothesis*, at least some hallucinations are indistinguishable through introspection from perception—which seems to imply that it is not possible to distinguish one from the other through introspection.

This challenge, however, holds only if certain hallucinations are indistinguishable from perception in an impersonal way, in a way that rules out that the subject's failure to discriminate can be explained in virtue of the presence of interfering factors that prevent one from picking up on a difference that would otherwise be accessible through introspection. The opponent of relationalism, then, has to prove that either there are some hallucinations that are indistinguishable through introspection from perception in the impersonal sense, or that such hallucinations are metaphysically possible. An inquiry that utilises both conceptual tools and empirical data on the psychopathology of hallucination will conclude that there is no empirical support for the claim that impersonally indistinguishable hallucinations actually exist. Moreover, the arguments

generally proposed to support their metaphysical possibility—including the argument from local supervenience and the argument from the ‘same proximate cause, same immediate effect’ principle—fail, even though ultimately I will grant that we can accept their possibility.

This challenge, I will argue, can be met by adopting a negative view of hallucination of the kind proposed by Martin (2004; 2006), whereby the phenomenal character of hallucination is solely determined in terms of its being indistinguishable from the phenomenal character of perception. Relationalists need to embrace this account of hallucination if they are to avoid committing themselves to denying that phenomenal characters are superficial.

Focusing on the idea that phenomenal characters are superficial and exploring what constraints their superficiality imposes on an account of perceptual experience provides a novel approach to the argument from hallucination and the debate over relationalism. It allows one to put some order into a debate where, as it stands, it is difficult to understand what motivations drive the proponents of competing views, where the common ground and the disagreements stand, and in which directions the various intuitions at play pull the actors in the debate. I will argue that the proponent of the argument from hallucination is motivated by an intuition that should be taken very seriously: phenomenal characters are superficial, and it is not clear that relationalists can claim that a perception and a hallucination which is indistinguishable from it have different phenomenal characters without denying their superficiality.

The present enquiry will show that the relationalist can meet this challenge and that the various versions of the argument from hallucination fail to prove relationalism untenable. Other arguments may succeed in what the argument from hallucination fails to do (see for instance Burge 2005; Pautz 2009; 2010; forthcoming; Noordhof 2002; Siegel 2011; Brogaard 2015 for some arguments that do not rely on hallucinations). However, I am confident that an examination of the notion of phenomenal character, the nature of superficial properties, the workings of introspection, and the notion of indiscriminability will prove fruitful for future work assessing further arguments against, but also in favour of, relationalism.

Chapter 1: The Relational View of Perception

In this chapter I will introduce relationalism, the view that the phenomenal character of perception is to be accounted for in terms of the obtaining of a relation of acquaintance between the subject and the object. I will identify its fundamental commitments and explain how it accommodates the way in which the phenomenal character of perceptual experience strikes us through introspection.

First, I will discuss how we should understand the notion of phenomenal character and what we should expect from a theory that aims to account for it. I consider Snowdon's scepticism about the utility of the notion of phenomenal character and the cognate expression 'what-it-is-likeness'. I will explain why I do not share Snowdon's scepticism and why I think it is valid to use the notion of phenomenal character, although only as a placeholder, to identify certain aspects of perceptual experience that call for a better understanding: the specific way in which perception strikes us introspectively. I will argue that giving a substantial answer to this question is controversial. The phenomenal character of an experience, then, is not obvious to us—contrary to what many participants to the debate hold (while

still disagreeing with each other as to what the phenomenal character of experience actually is).

In particular, I will counter the idea that it is part and parcel of the notion of phenomenal character that it is to be understood in terms of intrinsic properties possessed by the experience. This leaves it open that phenomenal characters are (in part) constituted by the obtaining of a relation to objects in the environment, as the relationalist claims.

I will then spell out how experience strikes us when we reflect upon it, and argue that relationalism seems to have an explanatory advantage over the main alternative accounts: sense-datum theory, representationalism, and qualia theories of the phenomenal character of experience. Finally, I will explore why relationalism is met with great resistance despite its apparent explanatory merits. The main problem that relationalism faces takes the form of a reduction ad absurdum. This starts with the claim that a limited range of experiences cannot be the way relationalism claims them to be and concludes that relationalism cannot be true in any case. The aim of this thesis will be to assess whether this argument is successful. The present chapter will be devoted to presenting how relationalism accounts for the phenomenal character of perception in the context of the debate of which it is part.

1.1 Trivial and Not So Trivial Claims about Experiences

Amidst the numerous controversies about the metaphysics, the phenomenology, and the epistemology of perception that occupy

philosophers, there is one claim that goes relatively unchallenged, i.e. that perceptual experiences are conscious mental occurrences such that there is something it is like for one to have them. It seems that having a conscious sensory experience amounts to having an experience that is subjectively like something for the subject who has it. This distinguishes conscious experience from unconscious mental processes and states of which one is not subjectively aware. In addition, it seems that what it is like to have a sensory experience is different from what it is like to have other mental occurrences, such as thoughts or recollections. The *locus classicus*³ for the idea of a fundamental what-it-is-likeness of conscious experience in the contemporary analytic tradition is Nagel's famous paper 'What it is like to be a Bat'. However, the consensus around the idea that there is a subjective conscious character to perceptual experience and awareness of this character is what allows the subject to distinguish her experience from other mental occurrences goes beyond the acceptance of Nagel's discussion. Nagel claims:

[...] fundamentally an organism has conscious mental states if and only if there is something that it is like to be that organism-something it is like for the organism. We may call this the subjective character of experience. (Nagel 1974: 436)

³ Nagel was not the first to talk about 'what-it-is-likeness'. As Snowdon (2010) notes in a footnote, Farrell (1950) and Sprigge (1971) had used this expression before. Nagel, however, is responsible for the fortune of both this idea and the expression.

What Nagel calls ‘subjective character’, others call ‘phenomenal character’, ‘conscious character’, or ‘sensuous character’. Here I will adopt the expression ‘phenomenal character’, most common in recent debates on perception.

Commenting on the above passage from Nagel, Snowdon (2010) makes two glosses. First, Nagel attributes conscious experiences to organisms, but Snowdon notices that the restriction to organisms should be dropped, as we do not yet know whether entities other than organisms (for instance computers) may be able to have experiences. Second, and most importantly for our current concerns, Snowdon notices that Nagel talks about the fact that there is something it is like *to be* an organism (or an entity) with conscious experience. But if this is true, this is presumably in virtue of the fact that there is something it is like to undergo a conscious experience. He then suggests that the basic claim expressed in Nagel’s passage should be ‘[...] that it is in the nature of experiences to be occurrences which it is like something to undergo’ (Snowdon 2010:11).

Nagel affirms the existence of a double implication: a subject has a conscious mental occurrence *if and only if* there is something that it is like to be that organism. Thus, it is meant to give necessary and sufficient conditions for having a conscious occurrence and thus an illuminating characterisation of its nature.

However, one might wonder whether saying that there is something it is like to have an experience is an exhaustive characterisation of what it is for an experience to be conscious, or even whether we are making any

progress at all in our knowledge of conscious experiences by pointing out that there is something it is like to have them. Snowdon expresses such scepticism and goes further still: he claims that Nagel's claim about what-it-is-likeness is either false or trivial. Implementing the remark, mentioned above, about switching the focus from the conscious creature to the conscious experience, Snowdon reformulates Nagel's biconditional as follows:

Necessarily, an occurrence E is an experience if and only if there is something it is like for an entity to undergo or enjoy E.
(Snowdon 2010: 12)

But if we read it from right to left, claims Snowdon, it is false. There are many things for which there is something it is like to undergo them even though they are not experiences, so the fact that there is something it is like to undergo an event cannot be an exclusive criterion for identifying a conscious experience. He gives the example of 'travelling on the Big Dipper', which is 'a type of thing that can happen without any experience—say, if one travelled on it asleep or unconscious' (Snowdon 2010: 15). From left to right, on the other hand, the biconditional is true, although, Snowdon says:

[...] it is more or less trivial. Experiences, like anything else, must be some way, and being events must be some way for the subject of them. Further, the slogan in no way pins down which

ways are the distinctive ways experiences are. (Snowdon 2010: 25)

As there is something it is like for the subject to undergo an experience only insofar as there is something it is like for a subject to undergo any event, and talk of ‘what-it-is-likeness’ doesn’t do anything to specify what distinctively it is like, Snowdon recommends that we should abandon it as utterly uninformative. I suspect that, as long as what is dubbed the subjective or phenomenal character of perceptual experience is identified in terms of what captures the ‘what-it-is-likeness’, he would recommend abandoning talk of ‘phenomenal character’ too.

However, pointing out, as Snowdon does, that simply saying that perceptual experience has a phenomenal character is trivial is not itself trivial, at least in the context of the current debate on perception. There is in fact a widespread tendency to take the simple trivial remark that perception has a conscious or phenomenal character for a substantive claim about certain phenomenal properties that experience allegedly possesses. And Snowdon’s invitation to abandon these expressions is primarily motivated by an attempt to avoid such a ‘major and disputable inference’ (Snowdon 2010: 27) of which the following quote from Chalmers is an example:

A mental state is conscious if there is something it is like to be in that mental state. To put it another way, we can say that a mental state is conscious if it has a qualitative feel—an associated quality of experience. These qualitative feels are also

known as phenomenal qualities, or qualia for short. (Chalmers 1996: 4)

Chalmers asserts that *any* conscious mental state has phenomenal qualities. However, while the claim that there is something it is like to have a sensory experience seems trivial and meets the general consensus of philosophers working on the subject, the idea that for *all* conscious states there is something it is like to have them is controversial (Carruthers and Veillet 2011). Since my current interest is limited to perceptual experience and I will remain neutral on the possibility that for other conscious mental occurrences there is something it is like to have them, I shall restrict Chalmers' claim to what matters to the current discussion: perceptual experiences. Chalmers assumes that saying that (a) there is something it is like to have an experience is equivalent to saying that (b) experience possesses phenomenal qualities, or qualia. Qualia are thought to be intrinsic properties of the experience in virtue of which an experience is conscious.

Snowdon correctly notices that (a) and (b) are far from being simply two different wordings for the same idea and that the inference from (a) to (b) is fallacious. To avoid the risk of erroneously slipping from (a) to (b), and because he doesn't see much use for (a) anyway, he recommends abandoning it altogether. However, we may be wary of the kind of fallacy exemplified by Chalmers, but still resist Snowdon's invitation to dismiss any talk of 'what-it-is-likeness' or 'phenomenal character'.

Although it verges on triviality, saying that perceptual experience has a phenomenal character might prove useful in formulating a problem

and giving a name to the subject matter under enquiry. Snowdon is right when he says that affirming that there is something it is like for a subject to have the experience she has (or that there is a conscious phenomenal character to her experience of which she is subjectively aware) ‘in no way pins down which ways are the distinctive ways experiences are’ (Snowdon 2010: 25). However, saying so is useful in pinning down precisely what needs further explanation, or what needs to be characterised in a more substantive way. Thus, we can take talk of ‘phenomenal character’ seriously, at least as a placeholder and insofar as it provides a useful shorthand to indicate what it is we need to account for—i.e. the distinctive way in which experience strikes one introspectively.

Saying something substantial about the phenomenal character of experiences requires answering, among others, the following questions. How is the way conscious experience strikes us best described? What are we aware of when we experience? What are we aware of when we direct our attention to the conscious character of the experience? How does conscious experience differ from other types of conscious mental occurrences (such as thought, recollection, imagination...)? Do experiences in different sense modalities have distinctive phenomenal characters?

That answering these questions is far from trivial is testified by the fact that there is a deep disagreement on all of them. The most significant disagreement concerns whether the phenomenal character of experience should be accounted for in terms of the obtaining of a relation to objects or is determined by features of the experience that the experience can possess

independently of the obtaining of a relation to the objects perceived. Among those who claim that phenomenal character is to be accounted for in terms of the obtaining of a relation, there are further disagreements about the nature of the relation and about what the objects of this relation might be. Among those who contend that it is to be accounted for in terms of features of the experience possessed independently of the obtaining of a relation, there is disagreement on what these properties are, the main options being representational properties or purely qualitative properties (qualia).

This gives us four main options when accounting for the phenomenal character of perceptual experience. Relationalism and sense datum theory both agree that phenomenal character is to be accounted for in terms of a relation to certain objects, but they disagree about the nature of those objects. Representationalists and qualia theorists agree that phenomenal character is to be accounted for in terms of properties that the experience possesses independently of the obtaining of a relation, but they disagree about the nature of these properties.

Most accounts of the conscious character of perception take either of these options, although the four options are not exclusive and do not exhaust the logical space.⁴ Mixed views are also possible. In particular, among those

⁴ Martin argues that two questions define the complete set of options for views about the nature of phenomenal character: ‘On the one hand, they are concerned with what can be present to the mind: can the presented elements in experience exist independently of our awareness of them? On the other hand, they are concerned with the manner or mode in which objects are presented to one in having experience: can the presented elements of experience be so presented as not to require their actual existence for one’s experience to be so’ (Martin 1998).

These questions generate four options, but since the type of properties constituting phenomenal character are not mutually exclusive, this generates fifteen possible accounts of perception.

who claim that phenomenal character depends on non-relational properties of the experience, some think that both types of properties (representational and purely qualitative) play an explanatory role. Moreover, there might be disagreement as to how to classify certain specific proposals that more or less explicitly aim to reconcile different views.⁵ We couldn't explain this disagreement if, as suggested by Chalmers and others, it was part and parcel of the notion of phenomenal character that it is constituted by intrinsic qualitative properties of the experience.⁶

1.2 How Experience Strikes Us

Far from being an innocuous implication of the obvious fact that experiences have a phenomenal character, then, the claim that an experience is conscious insofar as it possesses certain sensory properties, or qualia, is a substantial and controversial theory of how we should understand the phenomenal character of experience. According to the qualia theory, the phenomenal character of an experience is made up of properties of the

⁵ McDowell (2013), Schellenberg (2013; 2014), and Logue (2014), for instance, argue that perceptual experience can have both a relationalist form, as relationalism tells us, and have representational content, as representationalism tells us. Whether these proposals really manage to reconcile the fundamental commitments of both relationalism and representationalism is a controversy I will not discuss here.

⁶ For those who believe that phenomenal character is, by definition, determined by intrinsic properties of experience, the debate over phenomenal characters takes the form of knowing whether experiences have qualia in virtue of something else (typically, in these debates, this something else is the representational content of the experience); or vice-versa, whether the representational content obtains in virtue of the purely qualitative phenomenal character. Chalmers (2004) for instance takes for granted that perception has both qualitative properties and representational content and frames the debate on the nature of phenomenal consciousness in terms of a dispute over whether representational content is grounded in qualitative properties or vice versa. This way of understanding the debate, however, is partial and, as we have seen, there is no reason to restrict it in this way, because nothing in the notion of phenomenal character entails that it is to be understood in terms of qualia.

experience: when I see red, my experience has a sensational property of redness. This is made clear in the way Shoemaker defines qualia:

Qualia, if there are such, are properties of sensations and perceptual states, namely the properties that give them their qualitative or phenomenal character—those that determine ‘what it is like’ to have them. (Shoemaker 1996, 121)

Similarly, Block (1990; 1996; 2010), who introduces the term ‘mental paint’ for the qualia theory,⁷ claims that mental paints are ‘qualities of perception that are not captured by what one is directly aware of or by representational content’ (Block 2010).

The appeal of this view doesn’t entirely come down to a blatant logical fallacy, as Snowdon thinks. The notion of phenomenal character is meant to capture that sensory and conscious aspect of an experience that is subjectively accessible to one who undergoes the experience. There seems to be a contrast between the way one has access to one’s own experiences and how other people have access to them. This contrast seems to amount to the fact only the subject who has the experience has access to what it is like to have it (although other people may know what it is like to have that kind of experience). This privileged access to one’s own mental occurrences is customarily called ‘introspection’. In general terms, introspection is simply

⁷ Block proudly takes up the term ‘mental paint’ used in a derogatory way by (Harman 1990), who, using an argument that refers to the diaphanousness of experience noticed by Moore (which I will shortly discuss), claims that introspection doesn’t give us access to any ‘mental paint’, meaning by that any intrinsic property of the experience beyond those properties that are (re)presented in experience.

defined as the way one can come to know about one's own on-going mental occurrences in a way that others can't (I will discuss introspection more extensively in chapter 3).

How introspection should be characterised is very controversial. However, it might be tempting to think of introspection as a quasi-perceptual monitoring process directed inwards. Locke defines what he calls 'reflection', by which he means 'the notice which the mind takes of its own operations, and the manner thereof', as '*the perception of operations of our own minds within us*' and he notices that this faculty, 'though it be not sense, as having nothing to do with external objects [...] is very like it, and might properly be called internal sense' (Locke 1690: II.1.iv. Italics in original). Recent proponents of a perceptual model of introspection include Armstrong (1981), Lycan (1996), Nichols and Stich (2003), and Goldman (2006).

Other philosophers stress the difference between perception and introspection, where the latter is thought to be more immediate than perception. Perception, they think, is a mechanism that acts as a mediator between us and the object of perceptual awareness; but no such mediation exists between us and the conscious mental occurrence when we introspect it: I am aware of the phenomenal character of the experience just in virtue of my having it. Either way, both models of introspection invite to think that being introspectively aware of the phenomenal character of one's own experience involves attending to certain intrinsic properties that our experience has.

However, the claim that perception has qualitative properties that account for the phenomenal character of the experience and the related claim that those properties are accessed when we introspect them have been widely criticised. The difficulty for this view is that, if we reflect on our introspection, this doesn't seem to give any support to the view that there are intrinsic properties of the experience. When we try to focus introspectively on our own perceptual experience, our attentive efforts don't seem to make us aware of any intrinsic properties of the experience. What we seem to be aware of are the properties of the objects we perceive—those properties that we were perceptually aware of, even before focusing our attention introspectively on our experience.

It is now commonplace in the current literature to refer to these observations as revealing the transparency of experience, an expression whose introduction is credited to Moore. Moore, in 'The Refutation of Idealism', claims:

[W]hen we refer to introspection and try to discover what the sensation of blue is, it is very easy to suppose that we have before us only a single term. The term 'blue' is easy enough to distinguish, but the other element which I have called 'consciousness'—that which sensation of blue has in common with sensation of green—is extremely difficult to fix. [...] And, in general, that which makes the sensation of blue a mental fact seems to escape us; it seems, if I may use a metaphor, to be transparent—we look through it and see nothing but the blue. [...]

When we try to introspect the sensation of blue, all we can see is the blue: the other element is as if it were diaphanous. (Moore 1903: 446-450).⁸

The transparency of experience is invoked by a number of philosophers and used to argue for different claims, and it is not always clear whether the different descriptions of the transparency of experience describe the same feature.⁹ So we will need to carefully unpick what we can learn about how experience strikes us from these remarks.

Martin notices that, when I stare at a lavender bush, I can do two things. I can attend to the colours, shape, and other properties of the lavender bush, ‘but I can also reflect on what it is like for me now to be staring at the lavender bush’ (Martin 2002b: 380). The shift in focus of attention, however, doesn’t seem to bring about any shift in the object of one’s attention:

When my attention is directed out at the world, the lavender bush and its features occupy the centre stage. It is also notable

⁸ Moore makes reference to visual perception, but it is implicit in his discussion that he aims to extend his remarks to other sensory modalities. In this, he is representative of the predominant tendency in philosophy to take visual perception as paradigmatic of perceptual experience in general. This might be due to the greater importance that normally-abled people seem to attribute to visual perception in their daily life, or to the fact that vision seems to directly connect us to surrounding objects more than other sense modalities, and thus it is more amenable to supporting the kind of view for which philosophical reflection is supposed to cause problems. There might be issues with generalizing from visual experience to the other senses, but for the purposes of the current discussion I will follow philosophical habit and take visual perception to be at least illustrative of perception in general.

⁹ Classic discussions of the transparency of experience are to be found in Harman (1990), Martin (2002), Tye (2002), (Siewert (2004).

that when my attention is turned inwards instead to my experience, the bush is not replaced by some other entity belonging to the inner realm of the mind in contrast to the dilapidated street in which I live. I attend to what it is like for me to introspect the lavender bush through perceptually attending to the bush itself while at the same time reflecting on what I am doing. So it does not seem to me as if there is any object apart from the bush for me to be attending to or reflect on while doing this. (Martin 2002b: 380–381)

These remarks contain a positive and a negative claim. The positive claim says that, when I introspect my experience, I become aware of mind-independent objects and their properties—the lavender bush out there: the very same objects and properties that I encounter when I direct my attention to the world. There is, thus, introspective support for the claim that properties of the objects I perceive are part of what I become aware of when I introspect the conscious character of my experience.

The negative claim says that, when I introspect my experience, I don't seem to become aware of the properties of my experience independently of becoming aware of the object of my experience and its properties. Notice that this negative claim is compatible with saying that there are properties of the experience we can attend to. However, we cannot attend to them without attending to the properties of the objects we perceive. This suggests that there is no positive introspective support for the qualia theory or the idea that perception possesses intrinsic qualities.

Notice that this doesn't provide a positive ground against the qualia theory. It only registers a lack of introspective support for it. The transparency of perception is compatible with the claim that, although very difficult or even impossible to detect, those intrinsic properties exist. However, any such theory will need to explain how it is that the conscious character of an experience is constituted by the experience's possession of certain properties, even though we are not aware of them. This is a very serious difficulty, given that the phenomenal character is supposed to be the conscious aspect of experience, what is subjectively manifest to one when experiencing—and this shouldn't be something that we cannot find in introspection. So, not only is the idea that phenomenal characters are to be explained in terms of qualitative intrinsic properties of the experience *not* part and parcel of the notion of phenomenal character. There is also lack of introspective support for this view.

Moore makes another interesting remark about how our experience strikes us. He establishes that, despite being difficult to fix, there is something, which he calls 'a sensation', that is distinctively subjective and pertains to the experience one has. He defines this sensation as follows:

A sensation is, in reality, a case of 'knowing' or 'being aware of' or 'experiencing' something. When we know that the sensation of blue exists, the fact we know is that there exists an awareness of blue. When we know that the sensation of blue exists, the fact we know is that there exists an awareness of blue.

And this awareness is not merely, as we have hitherto seen it

must be, itself something distinct and unique, utterly different from blue: it also has a perfectly distinct and unique relation to blue, a relation which is *not* that of thing or substance to content, nor of one part of content to another part of content. [...] To be aware of the sensation of blue is [...] to be aware of an awareness of blue; awareness being used, in both cases, in exactly the same sense. (Moore 1903: 449)

In this passage Moore tries to go beyond the transparency of the experience and to distinguish some feature of the experience itself. This, according to Moore, is something that can be done ‘if we look enough, and if we know that there is something to look for’ (Moore 1903: 450). What he finds is compatible with the transparency of experience; it is in fact an element that is diaphanous (and indeed might explain why experience seems diaphanous), but can nonetheless be detected. Moore notices that reflection on our experience reveals something about the nature of the ‘sensation of blue’, which today we would call the sensory conscious experience: it is an awareness of blue. By attending to my experience I can become aware of the fact that experiencing is an awareness of the blueness of the lavender bush, the object of my conscious experience. By introspecting, I become aware of the perceptual relation of awareness I stand in with the property. Hence Moore’s claim that ‘to be aware of the sensation of blue is [...] to be aware of an awareness of blue’. Moore claims that, in this sentence, ‘awareness’ is used ‘in exactly the same sense’, but we might want to be more careful and draw a distinction here. When I reflect on my experience, I

find the object of the experience—in Martin’s example, the lavender bush and its properties. I also find that my experience of the lavender bush seems to amount to a relation of awareness. While I perceive the lavender bush, I do not perceive my relation to the lavender bush, but I am nonetheless aware that my experience seems to amount to the obtaining of a relation to the lavender bush. So although we are aware of the relation of awareness, we are not aware of it in the same way that we are aware of the object to which this relation of awareness relates us to (see Soteriou 2016: 7 for a similar remark).

Some of Moore’s contemporaries make a similar remark and notice that, when we reflect on our experience, we seem to become aware of the fact that our experience takes the form of a relation of acquaintance (Russell 1992: 35; Price 1932: 4). Price notes explicitly that the relational nature of the experience is supposed to be something for which there is a strong introspective support:

When I say ‘This table appears brown to me’ it is quite plain that I am acquainted with an actual instance of brownness (or equally plainly with a pair of instances when I see double). This cannot indeed be proved, but it is absolutely evident and indubitable.
(Price 1932: 63)

If we reflect on our experience, what having a conscious experience seems to amount to is being acquainted with something. Soteriou stresses this point when commenting on Moore’s passage:

One significant aspect of Moore's discussion is the way in which it brings to the fore the idea that sensory experience somehow involves some kind of psychological relation of awareness, and not simply some psychological event, process, state, or property. In this context, to say that the relation of awareness is a psychological one isn't simply to say that one of the relata of the relation is a psychological subject—a bearer of psychological properties—for there are also non-psychological relations that psychological subjects can stand in to things. The suggestion appears to be that when one has a sensory experience there obtains a distinctive psychological relation that one stands in to some sensory quality, where that sensory quality is not a quality of the psychological relation. The awareness of blue is not itself blue. (Soteriou 2013: 13)

The positive transparency claim points out that, when we reflect on our experience, we find objects of our experience. The negative transparency claim notes that there is no feature of the experience itself to which we can attend through introspection independently of attending to the object of the experience. This further remark specifies what this feature of the experience seems to amount to: a psychological relation to those objects of perception. As it is the case for any relation, its obtaining involves the existence of its terms, in this case, the conscious subject and the objects perceived. As per the negative transparency claim, this relational feature is

not something to which one can attend without attending to the object to which one is related.

1.3 Relational Views of Phenomenal Character

Some philosophers have argued that these remarks about how experience strikes us when we try to reflect introspectively support the claim that perception is indeed, as it appears to be, a relation to those objects that we appear to perceive.

Relationalism—also often called ‘naïve realism’—is characterised as follows by some of its most prominent proponents:

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

A naïve realist view of (entirely veridical) perceptual experience is as that of a relation between the perceiver and objects of perception. Taking sensory experiences to be events, these objects of perception are to be understood as constituents of the event in question. The naïve realist supposes it is an aspect of the essence of such experiential episodes that they have such experience-independent constituents. (Martin 2006: 357)

The most fundamental characterization of our perceptual relation with the physical world is to be given in terms of a relation of conscious acquaintance between perceiving subjects and the particular mind-independent objects that are presented to them in perception. (Brewer 2011: 94)

All the authors claim that the phenomenal character of experience amounts to the occurrence of a relation to some mind-independent object. Campbell doesn't mention the obtaining of any relation in the quote above; however elsewhere he says that the 'experience of an object is a simple relation holding between perceiver and object' (Campbell 2002: 115). Moreover, he claims that the phenomenal character of experience is constituted in part by the mind-independent object perceived, where the notion of constitution is to be understood as a consequence of the fact that a relation occurs: a relation doesn't obtain without all its relata being in place. Thus, we should take the idea that objects are constituents of the phenomenal character of experience to be as fundamental to relationalism as the very claim that it involves the obtaining of a relation. Because of this relation of constitution, as Martin makes explicit, 'one could not have had such a kind of event without there being relevant candidate objects of perception to be apprehended' (Martin 1997: 56).

The idea that the phenomenal character of a perception is determined by the obtaining of a relation in which the mind-independent object is 'made available' (Campbell 2002: 131), and where external objects and their properties 'shape the contours of the subject's conscious experience'

(Martin 2004: 64) explains the introspective datum of Transparency. When we introspect, we find the properties of mind-independent objects because experiencing them amounts to becoming aware of objects and their properties. So, although the redness of the tomato is not a property of the experience, it determines what it is like to have the experience (at least in part) in virtue of being the relatum of the relation of awareness that is occurring. When one introspects, one doesn't find intrinsic properties because the relation involved is, as Brewer says, fundamental: it doesn't obtain in virtue of something else obtaining (e.g. a representational state or the occurrence of a state with intrinsic properties).

By affirming that the phenomenal character of experience is partly determined by a relation of awareness to the sort of mind-independent objects of which we ordinarily take ourselves to be perceptually aware, relationalism differs from sense-datum theories. Sense-datum theories equally maintain that experience is relational and one couldn't have the conscious experience one has if the object one seems to be aware of were absent. However, for reasons that will become clear in section 1.5, they do not identify those relata of experience with the mind-independent objects of which we seem to be aware.

1.4 Open Questions and Terminological Distinctions

Invoking the obtaining of a relation to certain objects as constituent of the phenomenal character of perceptual experience doesn't exhaust the

characterisation of phenomenal character. Many questions remain open, leaving scope for disagreement among relationalists over several details.

One question concerns the kind of ontological entities, if any, that are involved in perception, beyond the obtaining of a relation. Most relationalists are elusive on this matter, although they generally have implicit underlying assumptions about the metaphysics of perception. Campbell (2002) and Fish (2009) think that perception is a mental state. Martin, as the quote above testifies, thinks perception is an event—although he remains neutral as to whether something else is involved. Brewer (2011) seems to think that the obtaining of a relation exhausts the ontological status of an experience. Soteriou (2013), on the other hand, offers a sophisticated and rich discussion of the ontology required to accommodate the conscious character of experience, in which he proposes that we see perceptual experiences as occurrent states whose obtaining over an interval of time depends constitutively upon the conscious perceptual experiences that one undergoes during that time.

Another open question concerns the nature of the psychological relation involved. Generally this is characterised as a relation of acquaintance, of awareness, of having the object presented, and so on. The notion of acquaintance provides a partial qualification of the nature of the relation. The term derives from Russell (1910) distinction between knowledge by acquaintance and knowledge by description and it thus suggests that the relation to the object in question is not discursive and is more primitive and explanatorily basic than propositional knowledge.

However, Russell maintains that the relation of acquaintance comes in different varieties (Russell 1992). So one might wonder whether there is more that one can say about the relation of acquaintance involved in perception—in case perception were not the only mental occurrence to display such a relation, is there anything that distinguishes it from other forms of acquaintance? And are there different forms of acquaintance, according to different sense modalities? One question is whether there are further relata involved in the experience, over and above the subject and the objects and their properties. Both Campbell and Brewer argue that we should think of the relation in question as a three-place relation involving a point of view and the conditions of observation. Soteriou claims that there are ‘structural features of perceptual acquaintance’ that contribute to the phenomenal character of the experience, on top of the contribution provided by the objects to which one is related. He claims that those structural features are introspectible, although they cannot be introspected ‘without focusing on the objects that one’s experience is an experience of’ (Soteriou 2013: 111), and that perceptual relations in different sense modalities have different structural features.

A terminological remark is in order at this point. The terms ‘relationalism’ and ‘naïve realism’ are often used interchangeably, and Brewer uses the term ‘object view’ for a position that, as we have seen, is importantly similar to relationalism. Despite the prevalence of the use of the term ‘naïve realism’, I prefer ‘relationalism’ and I will use this term consistently from now on. There are two reasons for this preference. One is

that the term relationalism better conveys what is crucial to the view: that the phenomenal character of perception involves the obtaining of a relation to the objects we take ourselves to perceive.

The second reason is to avoid possible confusions of the view at hand with either naïve realism understood as the purported common sense view of perception or direct realism, with which the term ‘naïve realism’ is often associated. The term ‘naïve realism’ was first introduced by proponents of sense-datum theory as a disparaging qualification for a view that they attributed to ‘ordinary folk’ (or, as Broad says, in yet another instance of philosophical condescension, ‘the unsophisticated’), and against which they framed their own positions. The term was subsequently proudly rehabilitated by proponents of relationalism such as Martin (1997, 2000), to mark a continuity with the view attributed, for better or worse, to ordinary folk, in polemic with the traditional philosophical assumption that ordinary folk can’t be right.¹⁰ However, the association with common sense realism, whatever that means, is potentially misleading.

First, it is doubtful that anything like a coherent view of perception can be attributed to common sense. If there is a common sense view of perception, it is likely to be a nebulous amalgam of possibly conflicting intuitions about various aspects of a subject matter—which is precisely how pre-theoretical intuitions about things are, and precisely why we need to theorise about them to see through them and reach a coherent picture. On

¹⁰ Martin also introduces the term ‘naïve realism’ as an alternative to the prevailing ‘direct realism’, which he finds confusing and inadequate to capture the relevant matter of disagreement among competing views (see footnote 13).

the contrary, relationalism employs sophisticated philosophical notions that are far from naïve. Moreover, the idea that one can identify some neutral pre-theoretical intuitions held by non-philosophers that are somehow not already influenced by the theoretical framework prevailing in the scientific and philosophical community and which permeate mass culture relies on a very simplistic understanding of belief-formation among the general public—a view that has long been challenged by historical and anthropological studies and by recent psychological studies showing the cultural variability of people’s intuitions.¹¹

Secondly, abiding to whatever ‘unsophisticated folk’ think about a subject matter shouldn’t be a reason to take pride (nor shame, contrary to what many philosophers, Hume and Descartes *in primis*, imply) in philosophical theories—and certainly relationalism doesn’t aim to do so. Relationalism rather aims to provide an account of experience that ‘best articulates how sensory experience seems to us to be through reflection’ (Martin 2006: 254). But providing an account that best articulates how experience strikes us should be the aim of any account of the phenomenal character of perceptual experience—this is almost tautological. The

¹¹ The cultural variability of reported intuition across different groups is often used by proponents of experimental philosophy as an argument against armchair philosophy (Machery et al. 2004; Weinberg, Nichols, and Stich 2001). For responses to this line of argument, see Sosa (2007) and Deutsch (2010). Whatever one thinks of the merits of Sosa’s and Deutsch’s defences of armchair philosophy, we shouldn’t be worried about cultural variations in intuition undermining the kind of philosophical project with which we are preoccupied here—clearly an instance of armchair philosophy. Relationalism doesn’t rely on intuitions of the kind tested by experimental philosophers, such as intuitions on Gettier cases and on reference. What relationalists take as motivations for their position is reflection about how experience strikes us, which relies on introspection, rather than intuitions.

relationalist's proposal is that it does so better than its opponents, and the arguments it uses to show this largely rely on evidence that is available to the 'unsophisticated' as well as to the smartest of philosophers. In this sense, relationalism is closer to common sense than the competing views: it is primarily concerned with getting phenomenological, intuitive, and introspective data right, rather than reflecting on alleged problems arising from philosophical reflection, and is hence it is less revisionary, avoiding the need to claim that the way our experience strikes us is fundamentally misleading with regard to its nature—although relationalism also has some counterintuitive implications, as we shall see later in this chapter and in the next.¹²

Sense-datum theorists often used the term 'naïve realism' as an equivalent of 'direct realism', the view that in perceptual experience we are directly aware of physical objects, as opposed to non-physical entities that mediate our perceptual contact with those objects. Although relationalism might be seen as committed to direct realism—provided that one accepts the terms of the debate between direct and indirect perception¹³—, this view is not

¹² Martin concedes that relationalism also has to be revisionary to some extent and claims that 'we are misled about the nature of some of our experiences by introspection' (Martin 2002: 421), but it is less revisionary because it is in a position to maintain that introspection is right most of the time and does not fail us systematically.

¹³ Austin (1962) famously argued that the contrast between direct and indirect realism is spurious and misleading and should be rejected. Many have accepted Austin's criticism and have advocated the distinction, others have kept using it as if the criticism had never appeared (and are largely responsible for the conflation of direct realism, naïve realism and relationalism. Others have tried to rehabilitate the distinction, showing that there are coherent and significant ways of understanding the contrast between direct and indirect objects of perception. Among the latter are (Snowdon 1992a) and Martin (2006, 2016). Although he argues that the debate over the direct object of perception is an intelligible and interesting one, Martin

specifically a view about the phenomenal character of perception, so it can be held by proponents of competing views of what accounts for phenomenal character, such as representationalism.

1.5 A Problematic Reasoning

Despite this initial appeal, relationalism has emerged as an articulate philosophical option only recently—while sense datum theories, representationalism, and qualia theories have dominated the debate over the last century—and still meets strong resistance. The resistance arises from what Valberg (1992a; 1992b) calls ‘the puzzle of experience’, which takes the form of an antimony. According to Valberg, there are two apparently perfectly adequate ways in which we can reflect on our experience.

On the one hand, we can be ‘open to our experience’, as we have done so far: we can reason on how experience strikes us introspectively and consider the best way to accommodate the phenomenology thus revealed. As we have seen, if we are open to how experience strikes us, we are led to think of experience as a relation to the objects that surround us: the lavender bush down the street, the computer on the desk, and so on.

does not recommend that philosophers keep using the contrast direct/indirect realism, as the same concerns at the core of that debate—how to account for the way in which experience strikes us—are better addressed in other terms. The contrast between direct and indirect objects of perception is useful as long as certain assumptions that motivate the sense-datum theory are in place (as I will discuss in the next chapter), but as soon as we question those assumptions, the contrast becomes void (Martin 2016).

On the other hand, we can also adopt another, according to Valberg, equally adequate approach to our experience. We can reason about it and the possibility entailed by what we know about experience. If we follow a certain line of reasoning, we find ourselves compelled to deny that perception amounts to a relation to the objects that we ordinarily takes ourselves to be perceiving. For the time being, we can abstract away from the specific line of reasoning that leads Valberg to this conclusion—I will say something more about it in chapter 5—and focus on the general structure of the argument, an argument that has been proposed in many versions since Hume (although precursors of it date back to Pyrronian scepticism).¹⁴

This reasoning works as a *reductio ad absurdum* of the idea that one is perceptually related to the objects one takes oneself to be perceiving. The argument has two steps: a base case and a spreading step (see Snowdon 1992). The base case establishes that, in limited range of experiences, one is not related to the object one takes oneself to be perceiving: the lavender bush at the end of the street and the computer on my desk. The range of the experiences considered are cases of illusions and hallucinations, often treated, as we shall see in the next chapter, in very similar ways (and sometimes phenomena that are not illusory at all, but which are treated like cases of conflicting appearances and phenomena of refraction, and are lumped together with illusions). The spreading step generalises this conclusion to all cases of experience, including those in which the object we

¹⁴ See Martin (2016).

take ourselves to perceive is indeed present. So, even in the case where the bush is in front of me and my experience strikes me as an experience of a lavender bush, it cannot be the case that my perception amounts to being related to the lavender bush.

As Valberg notices, what makes this reasoning so problematic is not simply that it compels us to reject a view, namely relationalism, to which we might be attracted. The problem is that leads to a conclusion that contradicts what we have always known to be true of our experience. One is faced with an argument that forces one to concede that that thing that one sees is not the lavender bush. But that one sees a lavender bush is a very basic part of what one takes experience to be, something that one has always known, even before starting doing philosophy, before reflecting on one's experience and what one can know about it through introspection.

This puts one in a paradoxical situation: if one accepts the conclusion of this problematic reasoning, one is prevented from believing what one already knows about one's experience: that it relates us to those objects that experience seems to present to us. One would be forced to accept that what one sees is not that very object that one takes oneself to be perceiving and with which one seems to interact in one's daily life, but something else,¹⁵ another object whose properties are still mysterious to one—or maybe nothing at all. But this, according to Valberg, is

¹⁵ In the version of the argument presented by Valberg, it is assumed that one is aware of mind-dependent objects. But there are versions of the argument that do not make this assumption.

unacceptable: as soon as one considers the conclusion of the problematic reasoning, one is bound to protest:

But wait, *this* object is a *book*. The object present to me when I look at the book on the table is the book on the table. There is nothing else there. Now I realize that, as a contribution to philosophy, thoughts of this sort may appear a trifle quick and simple-minded; yet it is precisely such thoughts that come over me when I reach the conclusion of the problematic reasoning. And when they come over me, they totally overcome the conclusion. The reasoning establishes that *this* is an internal object [or, in a less committal way, not the book]. But it is not. It, *this* object, is a book. What gives me the right to say this? Where is the argument? There is no argument. The arguments are all on the other side. I do not conclude that the object present to me is a book, but that is all I find—the book. (Valberg 1992a: 22. *Italic in original*).

While one accepts the conclusion, according to Valberg, one also knows that the conclusion is unacceptable, because it conflicts with what one knows. But by accepting the conclusion one is nonetheless prevented from believing what one knows: one ‘becomes closed off from what we know’ (Valberg 1992b: 37). At the same time, one both has to accept and cannot accept that what one sees is the book in front of one.

Valberg notes that there are two possible stances that one can take towards the antinomy, once it comes into light. One may try to solve it and ‘demonstrate that the conflict in which it consists is merely apparent’ (Valberg 1992b: 42). This involves exposing a fallacy in the problematic reasoning. Alternatively, one may try to explain the puzzle, in order to explain why ‘there really is a conflict such as we have described’ (Valberg 1992b: 42).

1.6 The Relationalist’s Commitment to Disjunctivism

We can see relationalists as trying to solve the puzzle of perception. Relationalists are obviously not blind to this problematic reasoning, which has played a central role in shaping the philosophical debate for centuries. On the contrary, they are aware that one implication of their view is that ‘one could not be having the very experience one has [when genuinely perceiving], were the objects perceived not to exist, or were they to lack the features they are perceived to have’ (Martin 1997: 93). It is a corollary of relationalism that the phenomenal character of perception and that of hallucination, although indistinguishable, are different. This is often called disjunctivism, as it recommends a disjunctive account of experiences: when one has a conscious experience, one is either in a psychological relation to mind-independent objects, or it only seems to one that one is in a psychological relation to the object one takes oneself to be perceiving.

According to relationalists, when a subject hallucinates, it seems to her as if she is having a conscious experience that involves a relation of

acquaintance to some mind-independent objects and their properties. However, she is mistaken: her conscious experience doesn't have the phenomenal character it seems to her to have. When one hallucinates, one is not only perceptually misled with respect to the object that perception seems to present one with. One is also introspectively misled with respect to the character that one's experience seems to have. It is in this sense that Martin acknowledges that relationalism is to some extent revisionary too. So, although relationalists may be seen as solving the antinomy by showing that we do not need to accept the conclusion of the problematic reasoning, they recognise that there is a tension in the way we think about experience. We need to accommodate both how experience introspectively strikes us and the possibility of hallucinations in which one is not related to the objects one takes oneself to be perceiving. Accommodating both intuitions cannot be done without revising some of the initial intuitions.

However, Martin thinks that the kind of revision relationalists propose is far more innocuous. The revision their opponents propose is committed to the claim that we are systematically wrong about how experience strikes us introspectively, as it exposes one to Humean scepticism. Hume argues that reflections on cases of conflicting appearances lead philosophers to 'contradict or depart from the primary instincts of nature, and to embrace a new system with regard to the evidence of our senses' (Hume 1748: 201). But when faced with the task of replacing the account immediately suggested by reflection on the senses with a new one

that takes into account these philosophical problems, philosophy finds herself extremely ‘embarrassed’, because

she can no longer plead the infallible and irresistible instinct of nature: for that led us to a quite different system, which is acknowledged fallible and even erroneous. And to justify this pretended philosophical system, by a chain of clear and convincing argument, or even any appearance of argument, exceeds the power of all human capacity. (Hume 1748: 201)

Here Hume introduces a particular form of scepticism about the senses, different to the more general Ancient and Cartesian scepticism about knowledge of the external world. If philosophical reflection shows us that introspection misleads us about the nature of perception and the kind of contact to the world that it provides, one can no longer rely on introspective reflection on one’s experience to make claims about one’s perception or the world we perceive, because the way experience introspectively strikes us has been proven erroneous.

Martin argues that the revision proposed by the relationalist is committed to rejecting the idea that *every time* experience strikes us introspectively as having a certain phenomenal character, then it has it. This means denying that introspection of phenomenal characters is infallible, and many would argue that this is not revisionary at all, because it was never part of our intuitions that introspection is infallible (I will say more about the infallibility of introspection in chapter 3).

Relationalists are primarily concerned with giving an account of genuine perceptions that best accommodates how experiences strike us. Having to recognise that such an account cannot be straightforwardly applied to hallucinations, because other extra-introspective considerations (i.e. knowing that no suitable object is present) are to be factored in, seems relatively innocuous.

1.7 Sense-datum Theories, Representationalism, Qualia Theories

Those who accept the argument from illusion and hallucination involved in the ‘problematic reasoning’ deem this disjunctive corollary of relationalism unacceptable. It is part and parcel of the ‘problematic reasoning’ that the conclusions reached concerning the base case must spread to all cases of perceptual experience. Therefore, most philosophers have sought accounts of experience that accommodate the conclusion of this problematic reasoning. Thus, they seek an account of the phenomenal character of experience whose obtaining doesn’t require the presence of the object. However, by doing so they are committed to denying that experience is as it seems to us to be when we reflect upon it.

Some seek to give an account of the phenomenal character of hallucination in terms of the obtaining of a relation to sense-data, i.e. objects that are different from the objects one takes oneself to be seeing. Although in recent debate there is an assumption that sense-data are mind-dependent entities (see Robinson 1994; Huemer 2011), when the notion was introduced by early 20th century philosophers by such as Price, Moore,

and Russell, the nature of these entities was an open question. The arguments from illusion and hallucination were meant to establish only that the objects of which we are aware are not objects such as the book, the lavender bush, and the computer, that we take ourselves to be perceiving. Whether these sense-data were mind-dependent, mind-independent, physical (as for Moore 1926, who considered them to be parts of the surfaces of objects), non-physical (Broad 1925; Price 1932), or neither (Russell 1921), was an open—often agonizing—question to be adjudicated independently of the argument from illusion and hallucination. However they conceived of sense-data, these thinkers were committed to denying that perception amounted to being related to the object we find when we reflect on our experience.

Representationalists, on the other hand, claim that the phenomenal character of the experience is to be accounted for in terms of the representational properties of the experience. Representationalists preserve the idea that, when one genuinely perceives a lavender bush, one sees a lavender bush. However, they forsake the intuition—for which I argued in section 1.2—that perception amounts to a relation to the object and as such requires the existence of the relata. According to representationalists, we are aware of the objects that we find when we introspect our experience, in virtue of the structure of perception, which has a distinctive *aboutness* or intentional directedness towards the mind-independent objects. However, the relation of intentional directedness—if it

is relation at all¹⁶—is *sui generis*, as it doesn't require the existence of one of its relata (the object of awareness) in order to obtain.

While sense-datum theorists agree with relationalists in that the phenomenal character of perception involves the obtaining of a relation, representationalists agree with relationalists that the phenomenal character of a genuine perception involves awareness of the objects we seem to perceive and nothing else.

As I briefly mentioned earlier, there are also theories that seek to explain the phenomenal character of experience purely in terms of the intrinsic qualities of the experience, or qualia. These theories forsake both the idea that when introspect we only find the objects we take ourselves to be perceiving and the idea that we seem to be psychologically related to those objects.

Relationalism, by attributing different accounts to the phenomenal character of perception and hallucination, has an explanatory advantage over these competing views. It doesn't have to claim that we are systematically wrong about how experience strikes us. It remains to be seen if this is a view that one can coherently hold or whether the arguments from illusion and hallucination succeed in proving it untenable.

¹⁶ Some intentionalists deny that perception is relational. Relations, they argue, necessarily require the existence of their relata, but intentionality allows an object to be 'present in absence'; hence it doesn't require the existence of the object presented. For instance, Crane recognises that 'the intentionalist view [...] comes with a price. For it must deny that perceptual experience is a relation' and the relationality of perception 'is one of the most obvious and common features of perception.' (Crane 2006: 141; see also pages 136–137 for an argument as to why intentionality doesn't involve a genuine relation).

Before proceeding, however, a few further terminological clarifications are in order. So far I have used ‘perception’, ‘experience’, and ‘perceptual experience’ somewhat loosely and interchangeably. After the introduction of the disjunctive view of experience this is no longer possible. To avoid confusion, I will strictly use ‘perception’ to refer to genuine, or veridical perception; and ‘experience’ to refer to any sensory experience, including perception, hallucination, and illusion.¹⁷

So far, I have also used ‘representationalism’ and ‘intentionalism’ synonymously—as they are often used in the debate. However, there is a distinction to be drawn between the general claim that experience possesses a representational or intentional content and the claim that the phenomenal character of experience depends on its intentional or representational content. This distinction is generally captured in the literature—although not consistently so—by introducing the following terminological distinction. ‘Representationalism’ is used for the claim that the phenomenal character of experience is to be accounted for exclusively in terms of the experience’s representational content. This view is also called ‘pure representationalism’ or ‘pure intentionalism’. ‘Intentionalism’ is used for the broader claim that experience possesses a representational content. For simplicity, I will stick

¹⁷ Some relationalists, following Travis (2004), deny that perception is evaluable for accuracy and hence they refuse to classify experiences as veridical and non-veridical. Doing so would be a categorical mistake. In this work, I will remain neutral as to whether it is possible to evaluate perception for accuracy. Accordingly I prefer to avoid commitments about the veridicality of perception and talk instead of ‘genuine perception’. Nothing of what I say will hinge on this, which is simply a terminological choice to avoid commitment to a claim to which I don’t need to commit.

to the prevailing convention of calling ‘representationalism’ the claim about phenomenal character and ‘intentionalism’ the more general claim.

This distinction emerged in the context of a debate internal to intentionalism. Representationalists argue that the phenomenal character of experience is exhausted by its representational content (Tye 1992, 1995, 2002; Thau 2002; Byrne 2001), while others argue that there are some aspects of the phenomenal character of perception that are not determined by the representational content that should be accounted for in terms of the experience’s possession of qualia (Peacocke 1983; Block 1996, 2003, 2010; Shoemaker 1990, 2001). However, it is worth noticing that relationalism is compatible with intentionalism: relationalists may accept that representational content has some explanatory role to play in perception. This explanatory role might contribute to other aspects of experience, like its functional role. It is also in principle possible to maintain that representational contents contribute to some extent to the phenomenal character of the experience, provided one preserves the main relationalist claim that the phenomenal character of an experience is partly constituted by some mind-independent object and would thus not be what it is in the absence of suitable objects. It is not clear, on the other hand, that relationalism is compatible with representationalism, as the latter seems committed to the idea that the phenomenal character of experience is solely determined by its representational content, in such a way that two experiences can have the very same phenomenal character whether or not the relevant object is present.

1.8 Conclusions

I have argued that relationalism provides a straightforward and non-revisionary account of certain features of perceptual experience, while the main competing views have to provide a revisionary account of some aspects of how experience strikes us. This doesn't mean that they fail to accommodate it—although they ought to provide an explanation of why experience seems to be systematically different to how it is and why reflection on experience throws us into the paradoxical situation described by Valberg, once we accept the consequence of the problematic reasoning. On the other hand, the viability of relationalism will depend on their ability to resist the conclusion of the problematic reasoning.

In this chapter I have only sketched the general lines of this problematic reasoning: since there are certain experiences for which relationalism is false, and we have compelling reasons to apply the same account across all cases of perception, relationalism should be rejected. As we shall see in the next chapter, proponents of the argument do not always present it in a detailed and clear manner. Reconstructing the forms of reasoning and the assumptions underlying the problematic reasoning and identifying a version of the argument that really poses some threat to relationalism will occupy us for the entirety of the next chapter.

Chapter 2: The Argument from Hallucination

In the previous chapter we have seen that, despite its initial appeal, relationalism is often considered a non-starter in virtue of what Valberg calls the ‘problematic reasoning’, featuring one or another version of the argument from illusion or hallucination. Towards the end of the last chapter I observed that, although the argument has been presented in various forms throughout several centuries, it is often difficult to understand precisely what the premises involved are supposed to be and what form the reasoning takes, so much so that Dummett suggested that the so-called ‘argument’ is not an argument at all, but rather a starting point indicating the need to explain how illusions are possible, how it is possible for things to appear other than they are.¹⁸

However, although often incomplete and unsatisfactory, reconstructions of the argument abound both among early modern proponents of mental representations (Hume 1748) and the sense datum theorists (Russell 1912, Ayer 1940, and more recently Robinson 1994 and Foster 2000), and among those who have reconstructed the argument in

¹⁸ ‘We commonly employ a distinction between how things appear and how they really are; and it is therefore natural to push this distinction to its limit. This seems to me the best way in which to view the so-called ‘argument from illusion’. If this is regarded as an argument, properly so called, with premisses and a conclusion, it is difficult to make out what are the premisses and what the conclusion. Rather, it is a starting-point’ (Dummett 1979: 2).

order to respond to it (Snowdon 1992; 2005a; ; French and Walters forthcoming among others).

In this chapter I will unpack the argument, by spelling out its structure and its not always explicit premises. As we have seen in the previous chapter, the argument has two steps: a base case and a spreading step. Those steps need to be carefully examined separately. First, I will ask whether we can establish that, in some cases, a relationalist account cannot be true, and if so, what the cases are for which this can be established. Then I will consider the arguments put forward to spread this conclusion to all cases.

Before starting assessing the argument, a clarification is in order. Traditionally, the argument was not directed against relationalism, rather direct realism. Because of the prevalence of epistemological concerns in those debates, asking whether perception really puts us in contact with the objects that we take to exist in the world around us had priority over the question of whether objects play a role in determining the phenomenal character of perception. This is not to say that the two questions are not connected, and that concerns about the characterization of the conscious character of experience did not feature in these debates. Early sense-datum theorists, at least, were very concerned with how experience strikes us, as my discussion in the first chapter, largely indebted to Moore, shows. However, the argument can be turned against relationalism. If we are not directly aware of physical objects, it cannot be the case that the phenomenal character of perception is constituted by those objects. This explains why

the relationalist account of experience failed to emerge until very recently, when many objections to this traditional argument from illusion emerged.¹⁹

Moreover, the argument didn't only aim to show that we do not directly or immediately perceive material or physical objects around us. It also concluded that instead, we are only aware of mental representations of them (Hume) or sense-data (Moore, Russell, Broad, Price). So there is a negative and a positive conclusion to the argument. The negative conclusion claims that the direct object of perception is not what we initially take it to be. The positive conclusion claims that what we are in fact related to are sense-data or mental representations. For our present purposes, we are merely interested in the negative conclusion.

The aim of this chapter is to formulate a compelling argument that targets specifically relationalism. Unpacking the structure of the argument to make it valid and unpacking what underwrites the different assumptions will require a considerable amount of exegesis. However, I will conclude the chapter with a formulation that, it seems to me, seriously threatens to overturn relationalism. Whether the assumptions involved really are to be accepted will be a question for the rest of this thesis.

¹⁹ Early criticism of the argument, pointing out a number of problems in the structure of the argument and its implicit and problematic premises, can be found in Prichard (1906), Hicks (1912, 1938 Chapter 2), Paul (1951), Barnes (1945), Austin (1962), Anscombe (1965) Mainly under the influence of Austin, from the Seventies onwards philosophers became quick in dismissing the argument from illusion (see for instance Pitcher 1971: 13-20, Jackson 1977: 107, Searle 1983: 61, Harman 1990, Tye 1992, Millikan 1991).

2.1 The Base Case: Conflicting Appearances, Illusions and Hallucinations

The first part of the argument seeks to establish that, for some base case, the phenomenal character of experience cannot be accounted for in terms of the obtaining of a psychological relation to the objects we seem to perceive. Traditional presentations of the argument lump together different cases, to which the term of ‘illusion’ is often applied indiscriminately: (1) cases of conflicting appearances, (2) optical illusions, and (3) hallucinations.

I will consider these three groups separately and examine how, for each case, the argument attempts to demonstrate that relationalism is untenable.

2.1.1 Conflicting appearances and Paradigmatic Looks

Most of the examples featuring in presentations of the argument from illusion by sense-datum theorists are cases of conflicting appearances: cases in which an object appears to have different properties to different subjects or to the same subjects in different conditions. These experiences are very common and involve ordinary optical phenomena such as refraction, perspectival changes, and changes in lighting conditions. For instance, in his presentation of the argument from illusion, Ayer mentions: a coin which looks circular from one point of view and elliptical from another; a straight stick which appears bent when it is seen in water; mirror images; mirages; change in the experienced taste of food due to the condition of the palate; subjective change in perceived temperature; and a coin which seems larger

when it is placed on the tongue than when it is held in the palm of the hand :
3).

Hume also presents the argument from illusion using a case of perspectival change:

The table, which we see, seems to diminish, as we remove farther from it: but the real table, which exists independent of us, suffers no alteration: it was, therefore, nothing but its image, which was present to the mind. (Hume 1748: 201)

For simplicity, and to avoid problems that are not central to the argument from illusion, we may replace Hume's dynamic property with a static property, such as having a certain colour.²⁰ So we can use the example of seeing a white table under a light that makes it appear yellow. I will also focus only on the negative conclusion that we do not perceive the table, disregarding the positive proposal that Hume aims to establish: i.e. that what we perceive instead is a mental image. This gives us the following, very simple, argument (see Snowdon 1992: 70):

1. Whatever I perceive is yellow
2. The physical table is not yellow
3. (from 1 and 2) What I perceive is different from the physical table.

Robinson claims that the argument is valid, as the conclusion follows from the two premises in virtue of Leibniz's Law (Robinson 1994:

²⁰ As Snowdon notices, the assumption that there is a persistent external object seems to depend on the thesis that we directly perceive an external object—a thesis that the argument is designed to reject (see Snowdon 1992b: 74).

32). However, in a recent paper, French and Walters (forthcoming), drawing on a remark by (Snowdon 1992: 74), argue that the argument, as it is presented, is invalid. The conclusion doesn't follow from the two premises only in virtue of Leibniz's law. What follows from Leibniz's Law is simply that the yellow object is not the physical table, but this is compatible with one seeing the yellow object *and* the table. In order for the argument to be valid, one needs to add a further premise, which Snowdon calls the 'Uniqueness Assumption'. The 'Uniqueness Assumption' claims that:

(UA): 'there is, in a particular direction of attention, as it were, a unique, single, [directly] perceivable thing' (Snowdon 1992: 74).

Snowdon finds the principle quite plausible, but French and Walters argue that the Uniqueness Principle, as it stands, is clearly false, because it entails that if one seems to see a wall behind a table that appears yellow when it is in fact white, one should conclude that one doesn't see the wall either, which seems hard to defend. The Uniqueness Principle also claims more than what the base case aims, and needs, to establish, for the base case seeks only to establish that one is not aware of the object one putatively sees as yellow (the physical table), not that one is not aware of *any* physical object. Craig and Walters replace the Uniqueness Principle with a principle in the vicinity, the 'Exclusion Principle':

(EH): If in an illusion S is aware of an object which is non-identical to the ordinary object O that the subject is putatively perceiving in an illusory way, then S is not aware of the ordinary object O.

Contrary to the Uniqueness Principle, this principle is not obviously false, but—Craig and Walters argue—it is not obviously true either. They consider possible justifications for the Exclusion Principle, including the idea that the object one directly perceives spatially excludes or occludes the physical object, but they conclude that none of these reasons is conclusive.²¹ In this context we don't need to establish whether the 'Exclusion Principle' should be accepted. It is enough to notice that the argument, as it is generally presented, is invalid, and that the further premise that needs to be added is far from obvious and a satisfactory justification for it has yet to be provided. Due to the widespread failure to recognise the need of such a further premise little or no attempt has been made to justify it.

Even if one accepts the Exclusion Principle, one might still take issue with the first premise. It is not obvious—a relationalist may say—that the first premise is true. It is not clear that, when I look at the white table showered with yellow light, I see something yellow, and I do not see something white. Austin (1962) famously pushed this line. He argued that describing phenomena of refraction, perspectival changes and changes in illumination as 'illusion' is misleading. The ordinary notion of 'illusion'

²¹ See (O'Shaughnessy 1984; O'Shaughnessy 2000) for another argument for the claim that seeing sense-data doesn't exclude seeing other entities.

suggests the idea of being misled by experience,²² while most of these cases are ‘too familiar’ (Austin 1962: 3-26) to be misleading. Beyond the critique of the terminological choice, there is a substantive point in Austin’s critique. The argument describes the experience of seeing the white table under a yellow light as a case where one sees a yellow table and doesn’t see a white table. But this is not how things look to us. This description is misleading, as it ignores completely the very well established phenomenon of perceptual constancy. Throughout the day, I see my white desk under all sorts of lighting conditions: under the bright white daylight, or under a yellow light of my desk lamp, in the dusk, and so on. On all occasions, the table looks white to me. Also, when the table is unevenly illuminated, all parts of it still look white. Additionally, it is a well-known fact that coloured light casts shadows of a complementary colour.

The same phenomenon of perceptual constancy applies to many other properties: size, shape, volume of sounds, etc. Consider the example used by Hume. When I look at the table from different distances, the table appears the same size, even though the retinal image of the table varies in size as I move back and forth. It doesn’t look as if the table changes size as I

²² Austin is not committed to the claim that the ordinary notion of illusion *entails* the idea of perception deceiving us. On the contrary, he criticises the extensive use of the metaphor that perception misleads us, while stressing that it is a misleading metaphor, because the senses do not tell us anything, either true or false (see Austin 1962: 11 and Travis 2004 for a discussion). The point here is that the proponents of the argument not only take the metaphor of being misled by the senses too seriously, but they also trade on an ambiguity, applying it to cases where the metaphor certainly doesn’t apply. Paradigmatic cases of appropriate use of the metaphor of being misled by the senses are cases of optical illusions and illusions produced by professional illusioners (Austin 1962: 22). Austin sees the attempt to extend the (already possibly misleading) metaphor to all cases of conflicting appearances as a disingenuous rhetorical device.

move around it. If it did, the world would appear a buzzing, blooming, incoherent flux of ever changing colours, shapes and sounds, and it wouldn't seem to us in the first place, as Hume admits, that we see a table, an object that seems to be independent of our perceiving it and which doesn't constantly change size. So there wouldn't be a tension between how experience strikes us—as an experience of the table, which doesn't constantly change size—and what we learn about it after reflection on perspectival variations—that it is an experience of an ever-changing storm of properties, something different from the table we seem to perceive.

Thus, relationalists claim that they can provide an alternative and better account of these experiences than that suggested by the proponents of the argument. Brewer, for instance, claims that perceptual experience is a triadic relation where the third relatum is made up by the circumstances involved in experience, including the subject's point of view. As we have seen in the previous chapter, this is a very widespread way of conceiving the relation in place in perception among relationalists. For instance, although Campbell (2002) doesn't explicitly engage with the argument from illusion, he equally commits to a similar triadic relation. This triadic relation accounts for the fact that in some sense the table appears yellow, and still appears white in some other sense. The way the table appears is a function of the colour property possessed by the object, the colour of the light, and my position relative to it.

Brewer's account of illusions builds on the idea that there are looks that are attributed to objects relatively to the conditions of observation.

When one experiences a table as yellow, when it is in fact white, the table one is perceptually related to presents the look of paradigmatic yellow things, given one's viewpoint and relevant circumstances. A white ball in red lighting conditions looks red because it looks the way paradigmatic red objects look in ordinary conditions. A straight stick halfway submerged in water looks crooked because it looks the way paradigmatic crooked objects look in ordinary conditions (Brewer 2008; Brewer 2011).

Martin gives a similar account of the immersed stick that looks bent:

In a circumstance in which one did confront a bent stick in good lighting, a paradigmatic circumstance for encountering the shape of being bent, one would be inclined to recognize the object as being bent; it would strike one as similar to bent things, and one would find in it an obvious similarity with bent things and a contrast with other shapes. If the psychological situation a subject is in when he or she truly utters ['The stick looks bent to me'] is relevantly similar to this paradigm kind of circumstance, then the subject is inclined to find the shape before one as similar to the paradigm of bent things—as more like being bent than anything else. (Martin 2010: 214-215)

For Martin, when we see a straight stick submerged in water, the visually basic properties (what he calls its look) that the stick really possesses are manifest to us. The stick looks bent because, in the current circumstances (when the stick is half-submerged in water), the visual properties that the stick possesses are similar to the characteristic look of

bent things. The stick looks like straight sticks look in water, which is in many respects like bent things look. However, they do not look exactly like typically bent things.

2.1.2 Optical Illusions and the Phenomenal Principle

At this point the proponent of the argument from illusion may concede that the cases presented above are not cases of illusion, and that by saying that one doesn't see the whiteness of the table, but sees the instantiation of yellowness, one misdescribes the phenomenon. However, there are genuine cases of illusion, cases where something really looks to have a property it doesn't have and doesn't seem to have the properties it actually has. This is something even Austin acknowledges: there are genuine cases of illusions, such as optical illusions (Austin 1962: 22). These cases might involve the same kind of phenomena involved in the cases mentioned above, such as perspectival change, but are situations artfully created in order to mislead the observer. An example is the class of anamorphic *trompe-l'oeil*: anamorphic bidimensional pictures that, seen from a certain angle, look like tridimensional objects. Another example is the class of anamorphic sculptures, where a tridimensional object looks to have a certain shape that in fact it doesn't have (say, that of a cube, or a human face, or a telephone) only when seen from a certain perspective.²³ Another example is the class of optical illusions like the Müller-Lyer diagram, in which there are two lines of the same length, one with hashes at the end of the line directed inwards

²³ Among the many artists who exploit anamorphosis in their sculptural work, see the work of Jonty Hurwitz and William Kentridge for examples.

and the other with hashes directed outwards. The line with the hashes directed outwards looks longer than the one with the hashes directed inwards.

In cases of optical illusion, it is not possible to reject the first premise by saying that it relies on a misdescription of the phenomenon. The argument can then be restated with the proper kind of case:

1. When I look at the Müller-Lyer diagram, I am aware of some things that are different in length
2. In the Müller-Lyer diagram the lines are not different in length.
3. What I am aware of are not the lines of the Müller-Lyer diagram.

However, it has been argued that the first premise is still suspect. Why should one accept that I am aware of things that are different in length? As we saw in the previous section, it seems possible to account for something's looking F without assuming that something actually instantiates the property F. Some relationalists apply the account sketched above to cases of optical illusions. Brewer, for instance, applies his account to optical illusions like the Müller-Lyer diagram. Here one cannot say that the lines in the Müller-Lyer diagram seem to be in one sense of the same length and in another sense of different lengths. One line just looks longer than the other, even if we know that the lines have the same lengths. But for Brewer this misleading appearance can be explained, like the cases discussed above, in terms of a triadic relation, where the way the object look is a function of the property of the object, the subject perceiving it, and the subject's position

relative to the object and other factors concerning the conditions of observation:

The [Müller-Lyer] diagram is visually-relevantly similar to a pair of lines, one longer and more distant than its plane, one shorter and less distant—a paradigm of inequality in length. It is therefore perfectly intelligible how someone seeing it might take that very diagram as consisting of unequal lines, regardless of whether she does or not. In this sense: they look unequal in length. (Brewer 2008: 172).²⁴

But why, many have wondered, should we accept that there are some things that have different lengths? All we know is that the two lines look of unequal length. But things, the relationalist can say, can look a certain way without being that way. The relationalist can claim that

²⁴ It should be noticed that not all relationalists think that this type of relational account can be applied to illusions. Some relationalists suggest that we should understand illusions in the same way as hallucinations, as experiences that doesn't possess a relational nature, but which only seem to have one. In an early paper, for instance, Martin seems to be committed to the idea that relationalism cannot apply to illusions, when he claims that:

naïve realist account of perceptual experience [...] cannot be directly applied to any case of delusive experience, such as illusions where one does perceive an external object, but misperceives it as other than it really is. (Martin 1997: 95)

The disagreement entails different ways of classifying experiences on one or the other side of the disjunction. Whether one embraces a perception/illusion *versus* hallucination or perception *versus* illusion/hallucination version of disjunctivism largely depends on how one spells out the relationalist commitments, for instance whether one allows other factors beyond the object and its properties to contribute to the phenomenal character of experience, or whether one conceives of the relation of acquaintance involved in perception as multi-dimensional (see previous chapter for a discussion of some open questions). If one has a definition of relationalism that rules out its application to illusions, one should accept the base case for illusions and try and block the argument at the level of the spreading step. In this case, illusions will pose the same problem that hallucinations pose to relationalism, and should be treated in a similar way.

accepting the first premise relies on a fallacious inference from ‘The lines seem to differ in length’ to ‘there are two things which differ in length’. This is a very familiar objection to the argument from illusion. Many philosophers have attributed the acceptance of the first premise to a failure to recognize the intentional nature of perceptual experience, which can present properties without them being instantiated in any object (see in particular Barnes 1944; Anscombe 1965; Harman 1990). This objection led to abandonment of the sense-datum paradigm in favour of an intentional paradigm in the second part of the last Century, together with the dismissal of the argument from illusion.

However, Robinson argues that it would be wrong to treat acceptance of the first premise ‘as little better than a howler’ (Robinson 1994: 54). The first premise relies on a principle that its proponents find very intuitive and phenomenologically well-motivated. Robinson calls this principle the ‘Phenomenal Principle’:

(P): If there sensibly appears to a subject to be something which possesses a particular sensible quality then there is something of which the subject is aware which does possess that sensible quality. (Robinson 1994: 32)

Martin—although he doesn’t subscribe to the Phenomenal Principle—also notices that the Phenomenal Principle shouldn’t be dismissed as the sign of the failure of the sense-datum theorists to realise that mental states are intentional and thus do not require the existence of the object they refer

to is explanatorily inadequate. For it doesn't take into account the fact that these sense-datum theorists recognise the intentionality of beliefs and the like, while denying that, on introspective ground, for the case of perception (Martin 2000: 208). Sense-datum theorists think that the Phenomenal Principle is required as part of an account of the phenomenal character of perceptual experience that best accommodates how experience strikes us introspectively.

So, although some proponents of the argument present the phenomenal principle as obvious, something that 'cannot indeed be proved, but [which] is absolutely evident and indubitable' (Price 1932: 62), it is best understood as motivated on the basis of its explanatory power (see (Snowdon 1992b): 73; Martin 2000: 209). The aspect of experience that calls for an explanation is tied to an aspect of perceptual experience that I have considered in the previous chapter as a motivation for relationalism. I noticed that perception strikes us an occurrence that relates us to what we seem to perceive, and that the obtaining of a relation involves the existence of the terms involved. Sense-datum theorists notice that experience strikes us as relational in this way even when it is illusory.

In the passage by Price that I quoted in the previous chapter, Price is implicitly defending not only a relational account of perception, but also the Phenomenal Principle:

When I say 'This table appears brown to me' it is quite plain that I am acquainted with an actual instance of brownness (or equally plainly with a pair of instances when I see double). *This* cannot indeed

be proved, but it is absolutely evident and indubitable. (Price 1932: 63)

Whenever something appears brown to me, there is something which is brown that I am aware of, irrespective of whether the object that seems brown to me is really brown. The argument against relationalists here could take the form of a *tu quoque*: how can relationalists, of all people, deny the Phenomenal Principle, when they are the first to insist that experience strikes us as relational and that this is best explained in terms of experience involving the obtaining of a relation?

The relationalist's response to this objection could be twofold. First, as we have seen in the previous chapter, relationalists admit that they cannot provide an account that is entirely non-revisionary. When one undergoes an illusion, one precisely has an experience that seems to be different than it is: it seems to present an object that instantiates a property, which in fact it doesn't have. But this is precisely what makes them illusions. The relationalist, however, preserves the idea that we are perceptually related to the object we seem to perceive: in the case of the Müller-Lyer diagram, the two lines drawn on the piece of paper in front of me. However, the two lines do not instantiate the property that they seem to instantiate.

Second, the relationalist can argue that, if the sense-datum theorists claim to defend the Phenomenal Principle on the grounds that it provides an entirely non-revisionary account of illusions, they are deluded. When they say that we ought to understand illusion as a case of being related to an object that instantiates the property that the object doesn't have, they might

well preserve the apparent relationality of illusion (something that, despite denying the Phenomenal Principle, the relationalist accommodates too). However, they commit themselves to the conclusion that what I am aware of when I see the Müller-Lyer diagram are not the lines I see in front of me, but some other mysterious entity—be it mind-dependent, mind-independent, physical or not. With this commitment, they violate another fundamental aspect of how experience strikes us: it is *this* object (these lines jotted on the piece of paper) that I seem to be related to, not something else, not a sense-datum. So, at best, they are on a par with relationalists in terms of the explanatory power of their respective theory. They are at least equally revisionary—although denying that the object I seem to see is the object I am really aware of is arguably more revisionary (and more dangerously so) than denying that the object I see instantiates the property it seems to instantiate.

Much more needs to be said about a relationalist account of illusions in order to make it compelling. The aim of the present discussion is not to provide a relationalist account of illusions, but rather to identify how the opponent can establish the negative conclusion for the base case. It seems that there are a number of problems with trying to establish it for illusion, even if cases that are unambiguously illusory, such as optical illusion, are selected. The structure of the argument is invalid unless we introduce certain assumptions. But these assumptions, namely the Exclusion Principle and the Phenomenal Principle, are far from obvious.

This weakens significantly the force of the argument. As Martin (2000: 211) notices, the suasive force of a *reductio ad absurdum* is conditional on the fact that the assumptions made in the argument are more convincing and appealing than the claim to be rejected, in this case relationalism. So far, it has emerged that, in order to establish the negative claim about the base case, a much more complex argument than that initially proposed is required, and that this more complex structure involves commitment to assumptions that do not have nearly as much suasive force as the thesis the argument is designed to overturn.

Moreover, with its commitment to the Phenomenal Principle, the argument depends on acceptance of the sense-datum theory, a view that is largely discredited. We want an argument against relationalism that can also be accepted by those who explicitly reject the Phenomenal Principle, such as representationalists.

2.1.3 Hallucinations

When we turn to hallucinations, on the other hand, the base case seems very easily established. We do not need the Uniqueness Assumption or the Phenomenal Principle to establish it. In the case of hallucination, there is no need to introduce any argument at all to establish the negative claim of the base case. It is a definitional feature of hallucination that one is not aware of a mind-independent object, as hallucinations are conscious sensory experiences occurring in the absence of the relevant object. Based

on this premise, one can also easily establish the truth of a negative claim targeting relationalism specifically (NC):

(NC): The phenomenal character of one's hallucination is not (even partially) determined by the obtaining of a psychological relation to the mind-independent object one seems to see.

This simply follows from the fact that, when one hallucinates, there is no suitable object in the environment for one to be related to. The problematic reasoning, then, should focus on the argument from hallucination, as it is the only version of the argument for which the base case can be straightforwardly established. Since the base case is so easily established for hallucination, while there are so many pitfalls in trying to establish it for illusions, it seems that the real threat for relationalism, if any, should come from an argument from hallucination, which is what we need to focus on here. Contrary to the argument from illusion, the relationalist cannot hope to block the argument at the base case: she can only hope to block the spreading step, which is, thus, what we need to focus on now.

2.2. The Spreading Step

The spreading step in the argument from hallucination aims to extend the negative conclusion for the base case (that in hallucination the phenomenal character of one's experience is not partly constituted by the mind independent object one seems to be presented with) to cases of genuine perception. How can the proponent of the argument secure this move?

The motivation that is generally given for the spreading step is that perception and hallucination are phenomenologically alike: both are sensory conscious experiences that seem to present objects as if they were in front of one. Ayer, for instance, claims:

When people whose legs have been amputated continue to feel pressure upon them, their experience is qualitatively as if pressure really were being exerted upon their leg. But, it is argued, if, when our perceptions were delusive, we were always perceiving something of a different kind from what we perceived when they were veridical, we should expect our experience to be qualitatively different in the two cases. We should expect to be able to tell from the intrinsic character of a perception whether it was a perception of a sense-datum or of a material thing. But this is not possible, as the examples that I have given have shown. (Ayer 1940: 6)

We can apply Ayer's reasoning to spreading the negative interim claim targeting relationalism. If the phenomenal character of a hallucination were constituted by completely different things than the phenomenal character of a genuine perception, one should be in a position to notice. But one is not: when one hallucinates, one has the same kind of experience one would have if one were having a corresponding genuine experience. The idea here is that perception and hallucination are phenomenologically alike and this phenomenological similarity is supposed to ground a unique account of their respective phenomenal characters.

However, Ayer doesn't explain why, in the absence of a phenomenologically noticeable difference between perception and hallucination, one should assume that they merit the same account. As common sense tells us, and as (Austin 1962: 50) pointed out, mere indistinguishability doesn't entail identity. A lemon and a lemon-shaped soap might be indistinguishable without being identical. Why should it be otherwise, when it comes to events like perceiving a lemon and hallucinating one?

Proponents of the argument from hallucination are not so unwary as to miss the obvious fact that indistinguishability doesn't entail identity. Broad, for instance, grants that it would be logically possible to admit that hallucinations do not put one in a relation to any physical, mind-independent objects and yet to maintain that in the perceptual case a relation of acquaintance to a physical, mind-independent object obtains—which wouldn't be a possibility if the entailment held. However, he still maintains that 'such a doctrine would be utterly implausible and could be defended only by the most desperate special pleading.' (Broad 1952: 9). Unfortunately, however, when it comes to explaining why this view is implausible, he doesn't provide any further reason than their resemblance and a 'continuity between the most normal and the most abnormal cases of seeing' (Broad 1952: 36-37).

Because the simple indistinguishability of perception and hallucination is not enough to justify the idea that perception and hallucination involve the same kind of experience and deserve the same

account, there must be some other assumption that motivates the spreading step. (Hinton 1967b; Hinton 1967a; Hinton 1973) argues that the idea that perception and hallucination deserve the same account derives from a misunderstanding of the implications of the quite trivial linguistic fact that there are what he calls ‘perception-illusion disjunctions’.

Although Hinton uses the term ‘illusions’, what he really has in mind are hallucinations, his favourite example being that of phosphenes, which he defines as ‘an illusion of a flash of light, which you get when an electric current is passed through your brain in a certain way’ (Hinton 1967b: 65). To avoid confusion, then, and to conform to the discussion of the argument from hallucination, I will systematically replace ‘illusion’ with ‘hallucination’ when reporting Hinton’s view.

Perception-hallucination disjunctions are sentences like:

- (a) I see a flash of light or I (only) seem to see a flash of light;
- (b) I see a flash of light or I have a hallucination of doing so;
- (c) I see a flash of light or I have a hallucination of seeing one.

(a), (b), and (c) are equivalent.

According to Hinton, that there are perception-hallucination disjunctions is ‘certain’ (Hinton 1973: 37) and innocuous. Hence, Hinton accepts that the truth of ‘I have an experience of x ’ is entailed by both the truth of ‘I see an x ’ and ‘I hallucinate an x ’. Perception-illusion disjunctions,

like any disjunctive claim, are true if either one or the other of the disjuncts is true.

This gives them a relative incorrigibility that the two disjuncts taken in isolation do not have. We can say for instance:

d) I am inclined to believe that I see a flash of light.

By this I mean that I believe that I see a flash of light but it is possible in fact that I am hallucinating one.

We can also say:

e) I am inclined to believe that I hallucinate a flash of light.

A subject undergoing an experimental test in a psychological laboratory during which alternatively she is presented with flashes of light and electric currents are passed through her brain, might claim (e), implying that she thinks she now has a phosphene, but without excluding the hypothesis that a real flash of light is being presented to her. By contrast, it doesn't make sense to claim 'I am inclined to believe that (a) I see a flash of light or I (only) seem to see a flash of light'. I can doubt whether I am in the situation described by either the first or the second disjunct, but I cannot doubt whether I am in either one case or the other.

That perception-illusion disjuncts are relatively incorrigible, Hinton notices, is a platitude. However, many philosophers take the relative incorrigibility of these disjunctive claims to show that, whenever one perceives, the same type of psychological event is going on: an experience, whose occurrence makes true the disjunctive claim. Many philosophers assume that whenever a perception-illusion disjunction is true, something

happens, something that occurs when I either perceive or I hallucinate: I have an experience whose occurrence makes true the disjunctive claim.²⁵

2.3 The Indistinguishability Principle

Hinton doesn't clearly explain why he thinks that the relative incorrigibility of a perception-hallucination disjunctive claim is taken as a reason to postulate the occurrence of a common event. Obviously, the simple fact that we accept a perception-hallucination disjunction shouldn't commit us to accepting it. In order to make sense of a disjunctive claim, one does not need to assume that one single kind of happening makes the two disjuncts true: two different kinds of event can make a disjunctive claim true. 'Either Tom is having breakfast or he is jogging' is not made true by a special 'having-breakfast-or-jogging' event. It also doesn't seem that the idea that a common event makes the disjunctive claim true arises from the relative infallibility of such a disjunctive claim. Many other disjunctive claims have the same relative infallibility (such as 'Either it is raining or it isn't raining'), but we do not take them to be made true by the same kind of event.

²⁵ Hinton also notices that this way of using the notion of experience as a shorthand for a perception-illusion disjunction is a piece of philosophical jargon that differ significantly from the ordinary use of the term 'experience', which normally stands for either one of two things:

- a) 'the transparent, colourless, or empty verb, akin to 'undergo' or 'do'' (Hinton 1973: 13), as in 'Mary experienced many difficulties in her life', or 'I will never forget the experience of swimming with dolphins', where one is the subject of the event that one experiences;
- b) 'a verb of which a very rough paraphrase would be 'witness'' (Hinton 1973: 13), where one is not the subject of the event or object one experiences, rather a 'detached observer'. This is when one, for instance, experiences the great firework display, the carnival flowers, a murdering (when one is not involved in the murdering) (Hinton 1973: 16).

However, I think that what Hinton has in mind is the following. ‘Either it rains or it doesn’t rain’ is different from ‘I see a flash of light or it only seems to me I see a flash of light’ in the following respect. In the case of the perception-hallucination disjunction, the event that makes the first disjunct true is introspectively indistinguishable for the subject from the event that makes the second disjunct true. By contrast, if one utters ‘Either it rains or it doesn’t rain’, in ordinary circumstances, one is able to distinguish whether the event that makes the disjunctive claim true is an instance of raining or of not raining.

The introspective indiscriminability of what makes the two disjuncts true calls for an explanation. The sameness of their phenomenal characters provides such an explanation. Notice that this is more plausible than the initial suggestion we found in Ayer. Ayer claimed that, if perception and hallucination involved different kinds of experiences, we should be able to notice. Since we do not notice, we should conclude that they are the same kind of experience. As we have seen, it is simply false that we should expect to notice a difference between any two things different in kind, and thus that if two things or events are indistinguishable that they must be the same kind of thing or event.

The reasoning suggested by Hinton is an inference to the best explanation, and it doesn’t rely on the fallacious inference from indistinguishability to identity. It states that an obvious explanation for the fact that a perception and a hallucination are indistinguishable is that they share certain properties. The lemon and the lemon-shaped soap are not the

same kind of object. However, they are visually indistinguishable, and this is so because they share certain observational properties (colour, shape, texture). A perception and an introspectively indistinguishable hallucination must equally be accounted for in terms of their sharing certain properties. Those shared properties are their phenomenal characters.

This puts the finger on what seems *prima facie* a forceful intuition to support the spreading step. It is plausible to think that indistinguishability of two things or events needs to be explained in terms of their sharing certain properties. It is also plausible that those shared properties, in the case of perception and hallucination, are their phenomenal characters. What else could they be? But if this is so, this is conclusive reason for accepting the spreading step and the success of the argument against relationalism: if the indistinguishability of a perception and an hallucination must be explained in terms of their sharing the same phenomenal character, the phenomenal character of the former cannot be accounted for in terms of being related to a mind-independent object, as this is not available in the latter case.

Following Martin (1997), I will call this assumption IND:

IND: If two experiences are indistinguishable for the subject of them, then they have the same phenomenal characters.

It is because we think that the indiscriminability between perception and hallucination should be explained in terms of their possession of the same phenomenal character, that we think that there is an event that occurs both in perception and hallucination: this event is the occurrence of an experience with a certain phenomenal character.

We are now in a position to reformulate the argument from hallucination in standard form, spelling out all the implicit premises, as follows:

1. Relationalism claims that the phenomenal character of conscious perceptual experiences is to be accounted for in terms of the obtaining of a psychological relation with the mind-independent objects one is perceptually presented with;
2. In some cases of hallucination, one has an experience which is introspectively indistinguishable from an experience one may have in a genuine perception;
3. In the hallucinatory case, the object one seems to perceive doesn't exist;
4. (from 3) In the hallucinatory case, the phenomenal character of one's experience is not accounted to in terms of the obtaining of a psychological relation with the mind-independent objects one is perceptually presented with;
5. If two conscious experiences are introspectively indistinguishable, they have the same phenomenal character IND;
6. (from 1, 4, 5) Relationalism cannot be true even in the genuine perceptual case.

It should be clear by now that the argument hinges on IND and that this is the premise the relationalist needs to challenge if she wants to block

the argument. Premise (1) simply states the relationalist position. Premise (2) and (3) seem undeniable. Is (5) equally compelling?

I said that the sameness of phenomenal character is introduced by proponents of the argument from hallucination as an inference to the best explanation. It better explains an aspect of how experience strikes us: perception and hallucination strike us as indistinguishable and this datum calls for an explanation. However, the relationalist can make a similar objection to that addressed to the proponent of the Phenomenal Principle. Accepting IND might look like the best explanation for the local datum that perception and hallucination can be indistinguishable. However, it commits one to denying that the phenomenal character of genuine perception is what it seems to be. So in this respect, it doesn't provide the best explanation. At best, relationalism and the proponent of IND are on a par.

If she wants to avoid this obvious objection, then, the proponent of IND should argue for something stronger than the claim that IND provides the best explanation for the indistinguishability of perception and hallucination. She should argue that their indistinguishability can *only* be explained in terms of the sameness of their phenomenal character. The argument is secured only if the proponent can claim that IND is a necessary principle: that no other coherent explanation of the indistinguishability is available. This also explains the ambivalence in the recent debate surrounding disjunctivism mentioned in the introduction, between presenting IND as resulting from an inference to the best explanation and presenting IND as an undeniable principle.

But if one frames IND as a necessary principle, it becomes less plausible. Why should an appeal to the sameness of phenomenal character of perception and hallucination be the only plausible explanation for their indistinguishability? Hinton expresses his bewilderment about it:

Why [...] should events have to have properties in common in order to be mistaken for one other? Why should it not just seem as if they had properties in common? Seeing a flash of light and having that illusion seem, but only seem, to have in common the property ‘when x occurs a flash of light occurs’. (Hinton 1967b: 225).

Hinton challenges the idea that if two experiences are indistinguishable, they must have a ‘worn-on-the-sleeve property in common’ (Hinton 1967b: 225). Although in certain occasions two things might be indistinguishable because they share some properties, like in the lemon and the lemon-soap case, it is not obvious that this must always be the case. In some cases, it might merely seem as if they have the same properties. In many cases we take one thing for another because they share certain properties. But in some cases, we take one for the other only because they *seem* to share certain properties. Two towers seen from a distance might seem to be qualitatively identical, in as much as they seem to have the same shape and size. In fact it might well be that one is closer than the other and one has a round base, and the other an octagonal base. Although they seem to have the same size and shape, they do not have any such shape in common. The colours of two objects might be indistinguishable for a subject, they might both, for instance, look red to one. However, it might

just be the case that one is red, while the other only seems to be red in virtue of being showered with red light. Equally, with experiences, it is not clear that if a perception and a hallucination are indistinguishable this is because they share the same phenomenal character. It might just be the case that one is a relation to an object, while the other only seems to relate one to an object.

This helps us identify exactly what claims the relationalist needs to challenge and hence what disjunctivism is committed to. In some sense, the relationalist can accept that the term ‘experience’ captures something that is common to perception and hallucination. The problem arises from the way we conceive of this common experience. There is a common event in perception and hallucination only in a ‘very unambitious sense’ (Hinton 1973: 46), inasmuch as this common event is understood in disjunctive terms, by what is picked up by the perception-hallucination disjunction. Nothing more specific than a perception-illusion disjunction can ‘answer the question as to what is happening to the subject’ (Hinton 1973: 46) when one either perceives or hallucinates a flash of light. The problem arises from thinking that the only way to make sense of the idea that both in perception and hallucination an experience is going on is to think that this experience is individuated by the possession of certain properties, i.e. by a given phenomenal character.

As Hinton puts it:

Nothing prevents us from introducing some such general term as ‘flash-experiencing’, stipulatively defined as ‘perceiving or being

illuded that you perceive a flash of light’, and nothing then stops us from saying that the statement, ‘Bob flash-experienced’, or ‘Bob had the experience of a flash’, or whatever form of words we choose to define in that way, ‘answers the question as to what happened to Bob’. It does so in the sense that it reports an event of which Bob was the subject. (Hinton 1973: 46).

According to Hinton, what perception and hallucination have in common is only that one is what the other merely is like. One is the reality and the other one is the illusion of that reality. The fact that one is the reality and the other its illusion doesn’t need to be explained in terms of shared phenomenal properties. Or at least it isn’t immediately evident why it should be so. Hinton’s strategy consists in shifting the burden of the proof onto the proponent of the argument, who should provide a more compelling motivation for accepting IND. This helps us understand better that disjunctivism is better understood as a purely negative claim. It denies that we ought to explain the fact that a hallucination can be indistinguishable from a perception in terms of sameness of phenomenal character.

The disjunctivist strategy for resisting the sameness of phenomenal character of perception and hallucination is to point out that no sufficiently compelling motivation has been provided for it yet. Hinton does not provide any positive argument to reject IND, but he contends that it is often simply taken to be obvious, (‘in a sense, so obvious that it is part of our basic understanding of the world within which [...] our concepts are embedded’, as says (Snowdon 2008: 54) in his discussion of Hinton’s book

Experiences). So Hinton sets a challenge for the proponent of the argument from hallucination: she must provide justification for IND.

In the next chapter I will discuss what one might say to support IND. In particular, we will have to make clearer what makes experience so special with respect to entities for which the inference from indistinguishability to identity is clearly invalid. In order to do so, we will need to have a closer look at the notions of indistinguishability and phenomenal character.

2.4 Conclusions

As we have seen in this chapter, the problematic reasoning that is supposed to reject relationalism is a *reductio ad absurdum* of the view, and requires two steps. First one needs to prove that, for a certain range of base cases, the relationalist account is not applicable. Second, one needs to show that the unsuitability of the relationalist account spreads from the base case to genuine perceptions.

In its traditional form the argument uses as base cases quite different phenomena, ranging from cases of perspectival change and refraction, understood as cases of conflicting appearances to complex hallucinations, and tends to treat them indiscriminately as cases of illusion. I argued that the base case does not succeed when we focus on illusions and cases of conflicting appearances. I indicated two contentious assumptions that are required for its success: the Uniqueness Assumption and the Phenomenal Principle. I have argued that the Uniqueness Principle is clearly false, while another principle in the vicinity of it—the Exclusion Principle—is not

obviously false, although it is also not clear that it is true. More importantly, I have argued that relationalists can resist the Phenomenal Principle. The principle cannot be presented as the best way to account for how experience strikes us in the case of illusion, because it forces us to reject a very important aspect of how experience strikes us: that what we seem to see is *that* object that looks to me yellow, not another object that instantiate the property of being yellow that the object doesn't have.

While adjudicating the base case for illusions proves problematic, it is straightforward and uncontroversial that relationalism cannot apply to hallucinations. However, it is not easy to understand what motivates the spreading step. Proponents of the argument motivate it in virtue of a phenomenological similarity between perception and hallucination—or, in some versions, their indistinguishability. However, this is puzzling because in the case of ordinary objects indiscriminability doesn't entail identity. I argued that what really motivates the spreading step is not a fallacious inference from indistinguishability (an epistemic claim) to identity (a metaphysical claim), but rather that the indistinguishability of perception and hallucination calls for an explanation and this explanation is given in terms of the occurrence in both cases of an event (an experience) with the same phenomenal properties.

Relationalists can retort that IND cannot be the best explanation of how experience strikes us, because it commits us to a revisionary account of genuine perception. The proponents of the argument must then aim for a

stronger claim: that accepting IND is the *only* coherent way to explain the indistinguishability of hallucinations from perception.

Hinton challenges this assumption, but doesn't provide motivations for rejecting it either. In subsequent chapters, we shall explore considerations both for and against IND. In the next chapter I will explore what one may say to motivate IND.

Chapter 3: How to Be Superficial

In this chapter I will explore the rationale for subscribing to the Indistinguishability Principle. The Principle runs as follows:

(IND): If two experiences are introspectively indistinguishable, they have the same phenomenal character.

In the previous chapter, I identified IND, following Hinton and Martin, as the key and most vulnerable premise of the argument from hallucination. Hinton shows no less than bemusement at the thought that so many philosophers are persuaded that if two experiences are indistinguishable, then they have a ‘worn-on-the-sleeve property in common’, as Hinton (1967: 225) puts it. When it comes to ordinary objects, the fact that they are indistinguishable for someone does not need to be explained in virtue of some property that they share. Of course, this is sometimes the case: a lemon and a lemon-shaped soap can’t be told apart because they have certain properties in common—superficial properties such as shape and colour. However, their sharing those properties does not entail their identity, as they can have *other*, more fundamental properties that distinguish them from one another. Sometimes objects ‘just seem, but only seem, to have [properties] in common’ (Hinton 1967: 225). Two

towers might look equally round from a certain distance, when in fact one is square. Properties can look different than they are and this explains why things can be indistinguishable without having to share certain properties.

In this chapter I will identify what the proponent of the argument needs to commit to in order to justify her acceptance of IND. I will suggest that the most compelling motivation for IND relies on the intuition that phenomenal characters are superficial. There is one way of thinking about superficiality that would make IND necessarily true. According to this understanding, superficial properties are properties for which there is no seem/is distinction. I will argue that there are no reasons in favour of—and in fact some reasons against—this way of understanding superficiality. I will then present an alternative way of understanding superficiality. According to this account, if a property is superficial with respect to a certain source of information, its nature must be accessible—at least under certain conditions—through the relevant source of information. This way of understanding superficiality doesn't entail the failure of the seem/is distinction. Thus, IND doesn't logically follow from the superficiality of phenomenal characters. In fact, it is possible for a superficial property to appear to be different than it is if there are interfering conditions that prevent the relevant source of information from functioning correctly.

However, the superficiality of phenomenal characters poses a serious challenge for the relationalist. If the nature of phenomenal characters must be accessible through the relevant source of information (in this case introspection), the relationalist needs to explain how she can maintain that a

perception and an indistinguishable hallucination have different phenomenal characters without being committed to denying that phenomenal characters are superficial. This leads us to the problem that the remainder of this thesis will be devoted to answering. In this chapter I will examine in detail the nature of this challenge and what motivates it.

3.1 The Purported Infallibility of Introspection

We can understand Hinton's bemusement at the widespread acceptance of IND: there is no general requirement to explain indistinguishability in terms of the sameness of certain properties. As we have seen, one might be unable to distinguish cotton muslin from dupioni silk, but this doesn't mean that the two types of fabric have relevant properties in common. For a colour-blind person, two apples might be indistinguishable with respect to their colour. This doesn't mean that the two apples have the same colour.

However, there is widespread conviction that there is a significant difference between two introspectively indiscriminable experiences and two perceptually indiscriminable apples. Two apples are indiscriminable for a subject with respect to their colour because they seem to be of the same colour to the subject. Precisely because indiscriminability is epistemic, the reason for the indiscriminability is explained in terms of facts that have to do with the subject: to the subject they look the same. There is a compelling intuition that, when it comes to explaining the fact that two experiences seem the same with respect to their phenomenal character, explaining the indiscriminability in terms of what's going on with the subject only serves

to support the claim that they have phenomenal characters, for the notion of phenomenal character is meant to capture how experiences strike us introspectively. Deutsch somehow voices this widely accepted, but often implicit, assumption:

What it is like to perceptually experience something that is perceptually indistinguishable from something else is just what it is like to perceptually experience that something else. Our notion of phenomenal character seems essentially tied to our notions of appearing the same as, or being perceptually indistinguishable from. Things that are perceptually indistinguishable cannot give rise to phenomenally different perceptual experiences—this ought to strike one as a conceptual truth. (Deutsch 2005: 9)

In this passage Deutsch doesn't actually claim that indistinguishable experiences ought to be explained in terms of experiences with the same phenomenal character, but he clearly assumes this. His claim is that the experience of indistinguishable objects should be explained in terms of experiences with the same phenomenal character. Thus, Deutsch seems to skip the second step (probably because it seems to him so obvious and indubitable) in the reasoning mentioned above: (1) the perceptual indiscriminability of two objects should be explained in terms of the indiscriminability of their experience; (2) the indiscriminability of two

experiences should be explained in terms of the identity of their phenomenal characters.

Beyond begging the question against any externalist account of phenomenal character, Deutsch's claim doesn't take into account the trivial fact that indistinguishable objects are seen from different angles, giving rise to experiences that are very much distinguishable with respect to how they strike one subjectively. But it is fair to assume that Deutsch is just being elliptical and has in mind experiences of indistinguishable objects seen in sufficiently similar conditions. What is relevant here is the claim that it 'ought to strike us as a conceptual truth' that when two things are perceptually indistinguishable to one, the experiences of the two things have the same phenomenal character.

It remains unclear, however, why we should take this to be a necessary truth. Deutsch appeals to the idea that 'our notion of phenomenal character seems essentially tied to our notions of appearing the same as, or being perceptually indistinguishable', which is true: these are clearly related notions. But Deutsch doesn't specify the nature of this tie and yet assumes it amounts to identity. The task we face as philosophers of perception is precisely to understand how these notions all relate to each other. Underwriting Deutsch's—and many others'—assumption is clearly a cluster of intuitions about the nature of phenomenal character, its relation to introspection, and to perception. It is unclear what exactly these intuitions amount to. We need to unpack them and see whether any of them should be accepted and can support IND.

A relationalist can say that the experiences of two objects (seen from the same point of view and under relevantly similar conditions) are phenomenally the same, because the notion of phenomenal character captures how the experience strikes one. But this doesn't entail that any two indistinguishable experiences should have the same phenomenal character. We have seen that perceptual indiscriminability doesn't entail sameness of properties in the objects perceived because one can explain the fact that two objects with different qualities are not discriminated in terms of something going wrong with the perceptual discriminatory capacities. One could say the same for introspective indistinguishability: two experiences seem the same with respect to their phenomenal character because something goes wrong with the introspective capacities of the subject. That is, unless one is committed to infallibilism about introspection.

Could a commitment to the infallibility of introspection be what underwrites IND? This is unlikely. Infallibilism about (certain) introspective judgments may have been popular in the past,²⁶ but nowadays it is out of favour and the recent literature on introspection abounds with counter-examples to infallibility. It is quite common, for instance, to realise that an episode of our childhood that we think we are recollecting is in fact an act of imagination, as our parents have told us that this event never actually

²⁶ The origin of the infallibility thesis is sometimes attributed to Descartes (1637; 1641). However, he explicitly commits to indubitability, while the attribution of the infallibility thesis is debated—although some textual evidence, especially in the Second meditation, may support his subscription to an infallibility thesis (see Schwitzgebel 2010). The infallibility thesis is also attributed to Locke (1690) and Hume (1748) and, more recently, has been defended by Ayer (1963) and the early Shoemaker (1963).

occurred. Very often our own emotions, motivations, and intentions are not transparent to us. For instance, we take ourselves to be angry with someone, when in fact we feel guilty and frustrated with ourselves, and so on. Counter-examples to the infallibility thesis also include introspective judgments about one's own experience. Empirical evidence suggests that when a perceptual input is presented to us for a very short duration of time, mistakes are inevitable (see Churchland 1984 for a discussion of these data).²⁷ Schwitzgebel notices that even outside of experimental settings, in very ordinary experiences when we have all the time we want to reflect on our experience, introspective judgments are not immune to error and uncertainty either:

Consider now your visual experience when you are looking directly at a canonical color but rather your visual experience, in an ordinary scene, of the region ten degrees or thirty degrees away from the center point. How clear is that region? How finely delineated is your experience of the shape and color of things that you are not looking directly at? People give very different answers to this question—some say they experience distinct shape and color only in a very narrow, rapidly moving foveal area, about one to two degrees of arc (about the size of your thumbnail held at arm's length); others claim to experience shape and color with high precision in thirty or fifty or one-

²⁷ See Bayne and Spener (2010), Shoemaker (1996), Rosenthal (2005) for other counterexamples to the infallibility of introspection.

hundred degrees of visual arc; still others find shape imprecise outside a narrow central area, but find color quite distinct even twenty or thirty degrees out. And furthermore, people's opinions about this are not stable over time. In the course of a conversation, they will shift from thinking one. (Schwitzgebel 2012b: 32)

As Schwizgebel notices, there is a great deal of disagreement about, for instance, how we experience objects in the peripheral areas of our visual fields, and about the boundaries of the visual field—where they are located, whether we can see them or whether we rather experience *within* boundaries without really experiencing them. Schwizgebel's explicit target here is indubitability, but since infallibility is stronger than indubitability, we don't seem to be infallible about those aspects of our phenomenal character if we tend to hesitate and further revise our judgment on how experience strikes us.

However, one might be tempted to argue that, although these examples show that introspective judgments can be wrong, these errors are due to misclassification or a lack of the appropriate concepts. Alternatively, it might be that one doesn't pay enough attention to how experience strikes one, but if one paid enough attention, one would be in a position to formulate correct judgments. Thus, there is a pre-judgmental sense in which introspection gives us infallible access to the phenomenal character of our experience. It is not clear that we can really interpret the scenarios described by Schwitzgebel as cases where, if only one could find the right concept, or

tried a bit harder to focus on her experience, one would be able to formulate true judgments. Moreover, to defend this view one would have to explain how a capacity may be pre-judgmental yet infallible. But what is more important here is to understand what one needs to commit to in order to claim that there is some form of non-judgmental introspection that is infallible.

In a recent paper, Giustina and Kriegel (forthcoming) defend precisely this view. They claim that there are two forms of introspection, in the same way that, according to Dretske, there are two forms of seeing: ‘epistemic seeing’ and ‘non-epistemic seeing’ or ‘simple seeing’ (Dretske 1969; 1981; 1995). Epistemic seeing involves facts, takes the logical form ‘S sees that p’, and requires one to possess and employ the relevant concepts and cognitive capacity. Non-epistemic seeing involves objects and is pre-judgmental. Giustina and Kriegel suggest that, along the lines of this distinction regarding perception, we should postulate two forms of introspection: fact-introspection and thing-introspection.²⁸ The latter ‘does not involve *classifying* or *characterizing* one’s conscious experience. It is a more direct awareness of the experience itself, in its pure phenomenal appearance so to speak’ (Giustina and Kriegel forthcoming). Others share the idea that we have some ‘immediate’ pre-judgmental access to the phenomenal character of our experience on which introspective judgments rely. Gertler, for instance, claims that ‘my judgment registering the relevant aspect of how things epistemically seem to me (*this phenomenal property is*

²⁸ They note that the term ‘thing’ in ‘thing-introspection’ is not meant to restrict the ontological category that one can introspect.

instantiated) is directly tied to the phenomenal reality that is its truthmaker' (Gertler 2012: 111; italics in original).

Giustina and Kriegel maintain that thing-introspection is infallible, but this is not due to any special cognitive achievement. It is *de jure* infallible: infallible in virtue of the fact that the domain of phenomena it concerns 'have no independent existence' and are 'facts fixed by how they are represented by the relevant representation' (Giustina and Kriegel forthcoming). In order to defend this limited version of infallibilism, one needs to commit to controversial claims regarding the nature of phenomenal character and its relation to introspection, starting with the idea that there is indeed a pre-judgmental form of introspection. This requires an account of introspection that not everybody would be willing to accept, plausibly either a quasi-perceptual model of introspection or a model where one is introspectively aware of the phenomenal character of one's experience just in virtue of having the experience (see the first chapter for a discussion of these views).²⁹ By contrast, the notion of judgmental introspection doesn't require one to commit to any particular view of introspection. In this sense, introspection is whatever cognitive faculties, processes, and states give rise to introspective judgments, i.e. judgments that the subject formulates, from a special first-person perspective, on her own mental occurrences.³⁰ These

²⁹ Both Kriegel (2009: Chapter 4) and Gertler (2012) defend versions of the latter.

³⁰ Accepting that there is a special way of accessing one's own mental occurrences—a way that is not available to third parties and that is some sense privileged with respect to the latter—doesn't require one to commit to the idea that a special mechanism or process is involved in introspection. Even a pluralist who maintains that introspection deploys a plurality of faculties, states, and processes that are also employed when performing other cognitive tasks, including knowing

faculties, processes, and states might not be exclusive to introspection; they might also be deployed in other cognitive tasks (see Schwitzgebel 2011, 2012; and Butler 2013 for examples of this pluralist view of introspection).

What is important to notice here is that even if we accept that there is a pre-judgmental form of introspection and that it is infallible, IND doesn't need to be true. Discriminating is a cognitive act: it is a mental activity involving cognitive capacities. Thus indiscriminability is the impossibility of performing such a cognitive act. One might not be in a position to discriminate the distinct phenomenal character of two experiences because of failures in one's judgmental capacities, even if one is pre-judgmentally aware of the phenomenal character for what it really is. We need to look elsewhere to motivate IND.

3.2 Superficial Properties

Giustina and Kriegel's restricted infallibility thesis relies more on an assumption about the relation between phenomenal characters and introspection than on a commitment to some special cognitive achievements that are possible through introspection. It is not clear what sort of special relation they think this could be. However, it is certainly part of the notion

other people's minds, can maintain that we still have special access to our own mental occurrences. This could be due to the a greater availability of evidence or to the way these multiple faculties, processes, and states can be combined when one introspects, rather than when one performs other tasks concerning the external world and other minds. Some theories of introspection explicitly commit to a parity between the first- and the third-person perspective on one's mind (see, for instance, Bem 1967 and Gopnik 1993), thereby denying that there is an introspective (and somehow privileged) and a non-introspective way of accessing one's own mind. However, apart from the parity theories, which are very contentious, it seems that all agree that introspection preserves some degree of privilege.

of phenomenal characters that there is a special relation with introspection. As we have seen in earlier chapters, the very notion of introspection is introduced to capture that aspect of experience that is accessible through introspection. But this doesn't yet say what kind of relation ties phenomenal characters and introspection together. Understanding the nature of this relation might help us make progress with deciding whether there is a valid motivation for IND.

One intuitive way of cashing out this special relation is to say that phenomenal characters are superficial with respect to introspection. Superficial properties differ from what we might call 'deep' entities in the following respect. Deep entities like the ordinary objects we find around us have a multitude of properties, accessible through a multitude of sources of information.³¹ A lemon has gustatory properties, shape, surface texture, colour, a biochemical composition, and so on. Some of these properties are accessible through taste, some by touch, by vision, or by sophisticated experimental enquiry. They have, so to speak, multiple dimensions and not

³¹ It is an open question that I will not address in this thesis whether the mark of deep entities is merely multiple sets of properties accessible through one source of information, rather than that they have multiple properties where some of them are other than the basic perceptual ones, like chemical structure—which are in some sense more fundamental. Choosing between the two options depends on one's ontological commitment regarding the existence of essential or fundamental properties. I am inclined to believe that we can draw a working distinction between deep and superficial properties just in terms of the number of sources of information through which their properties are accessed. This has some implications for what entities count as superficial. In particular, this entails that if observational properties such as shape and size are to count as superficial, the relevant superficial property is not the shape and size that one can both see and touch. Instead we would count as independent superficial observational properties the visible shape and size and the touchable shape and size.

all of these dimensions are accessible through the same source of information.

Some entities are not like this. Phenomenal characters seem to have only one dimension. It seems that all there is to a phenomenal character is what can be introspected of it. Phenomenal characters seem superficial with respect to introspection. Observational properties, or at least some of them, seem superficial: a smell can only be known through olfaction; a colour can only be seen.³² Both phenomenal characters and these observational properties are superficial in the sense of being tied in some special way to a specific source of information: introspection in one case and perception in the other. Given this initial understanding of superficiality, it is possible that, besides observational properties and phenomenal characters (which are superficial with respect to introspection), there might be other superficial properties with respect to other sources of information, and that these share interesting properties with observational properties and phenomenal characters.

There has been little discussion of the nature and features of superficial properties. The discussion about superficiality mainly takes the form of wondering whether this or that property is superficial, or whether superficiality is a coherent notion at all, but surprisingly little can be found in the way of defining what superficiality is. Deciding whether this or that property is superficial is generally made on the basis of a quite widespread,

³² Whether other properties that might be considered observational, such as shape and size, are superficial is more controversial (see also previous footnote). Thus I will only focus on observational properties such as colour, smell, and taste here.

although little argued for, conception of what superficiality amounts to. Williamson voices this widespread way of understanding superficiality as follows:

Things of most kinds have depth; they can be indiscriminable without being identical. It is nevertheless tempting to believe that things of some kinds are purely superficial; for them, indiscriminability and identity would coincide. (Williamson 1990: 48)

Clearly, this understanding of superficiality would make IND true by definition, unless one denies that phenomenal characters are superficial.

Williamson here reports, without subscribing to, a widespread and tempting form of reasoning: if deep entities can be indiscriminable without being identical, superficial properties are properties for which indiscriminability and identity coincide. They are properties for which there is no seem/is distinction. The idea here is that what accounts for the failure of the inference from indiscriminability to identity in the case of things with depth is that they have multiple properties accessible through multiple sources of information. So one might be unable to distinguish through perception a lemon from a lemon-shaped soap, because the lemon and the lemon-shaped soap share certain properties—those that are accessible through perception. However, they have many other properties that differentiate them, but that are not accessible to the subject (in given circumstances). Superficial properties do not have depth; they are available

entirely through one source of information, having only one dimension. If this is the case, the explanation for the failure of the inference from indistinguishability to identity that we just gave for deep entities fails. It cannot be the case that one fails to discriminate between two properties because one has partial access to them. If there is access to the superficial entity at all, it must be through the elective source of information, and this gives complete access to the entity. Or so the argument goes.

However, this understanding of superficiality is deeply unsatisfactory. First of all, it seems to rely on a fallacious argument:

1. Deep entities can be indistinguishable without being identical;
2. No superficial property is deep;
3. Therefore, no superficial property can be indistinguishable without being identical.

But this argument is invalid because the major premise is distributed in the conclusion, but not in the minor premise. Moreover, when it comes to superficial properties, although one of the explanations for the failure of the inference from indistinguishability to identity normally available for deep entities is not available (one cannot appeal to the multiple properties of an object that are not accessible through the currently exploited source of information), another explanation—which we have already considered, when discussing Hinton in the previous chapter—is. Two objects might be indistinguishable with respect to their superficial properties because they *just seem* to have the same properties, due to inattention, an unfavourable point of view, impairments of the subject's cognitive or perceptual

capacities, and so on. I mentioned earlier the case of two apples that are indistinguishable to a colour-blind person, although one is green and the other is red. But colour is precisely one of those properties that we are inclined to think of as superficial. So it doesn't seem true that superficial properties are properties for which indistinguishability entails identity.

Finally, if it were true that phenomenal characters are superficial and that for superficial properties indiscriminability entails identity, this would entail the infallibility of introspection—and not just in its restricted sense, as involving pre-judgmental introspection, because, as we have seen, introspective discrimination seems to entail cognitive capacities. But we have already seen that there are many reasons to resist the infallibility claim. Thus, if one needs to commit to this view of superficiality in order to justify IND, the proponent of IND is in a very unstable position. For the relationalist can easily dismiss the assumptions on which it relies.

3.3 Observational Properties

Although understanding the notion of superficiality as entailing an inference from indiscriminability to identity for superficial properties is problematic, the initial intuition that certain properties are superficial remains. Thus, there might be an alternative way of understanding the notion of superficiality that proponents of IND might use to support their view and show that the disjunctivist commitment to relationalism is untenable. I do not aim to provide a full account of superficiality here. I will merely identify some criteria that allow us to draw an interesting distinction between deep

and superficial properties. Then I will consider whether these criteria impose constraints on what can be said about superficial properties—constraints that the proponent of IND could use against relationalism.

In order to find such criteria, I will primarily turn to those observational properties that seem superficial, and in particular to colour, rather than phenomenal characters. This is because, as saw in the first chapter, it is difficult to spell out what phenomenal character is without committing to some controversial and philosophically loaded claims. Modelling an account of superficiality on phenomenal character runs the risk of incorporating independent and controversial assumptions about phenomenal characters into the notion of superficiality. The nature and metaphysics of colours is very controversial too, but at least it seems possible to agree on what we are talking about, that is, on the phenomenon we are trying to describe, without making metaphysical assumptions.

Not everybody would agree that colours are superficial: that they are tied to vision in an interesting way, which tells them apart from other things that can be perceived. Reductive physicalists about colours, for instance, would argue that colours have depth, since their nature is defined by properties that are hidden to perception, such as their spectral reflectance profile. For our present purposes, I don't need to delve into the dispute over the metaphysics of colour or any other observational properties. What I propose here is a conditional claim: if colours are superficial, there are certain constraints that are imposed on how we should conceive them. These

constraints can then be extended to all superficial properties, and once we have these constraints and criteria for superficiality in place we can apply them to phenomenal characters too—whose superficiality is far less controversial.

Moreover, when people deny that colours or other properties are superficial and claim that they have no less depth than objects like lemons and pieces of soap this is partly motivated by the difficulty of giving a coherent account of superficiality. Previous attempts to identify a criterion that distinguishes observational properties from other properties and properties failed this task. Many of these accounts talk of being superficial with respect to perception, in the yet indeterminate sense of having a special tie to perception that I have introduced here. However, they fail to frame the nature of this tie in a satisfactory and informative way. Take for instance this proposal from Langsam:

Properties are observational in so far as they are presented or represented in perceptual experience. Properties are represented in experience by means of appearances; thus we can say that observational properties are those properties that objects can appear to have when we perceive them. (Langsam 2000: 69)

This definition, however, is far too inclusive and fails to single out observational properties from any other entity that can be seen. The realm of properties that objects can appear to have when we perceive them is far broader than the class of properties we want to call observational (typically,

as we saw, colour, shape, sound, smell, taste...). Being a lemon is a property that a lemon can appear to have when we look at it. Emotions and physical states can be presented in experience too: it is by looking at someone's face and how it appears that we can tell that she is irate, sad, or out of breath.

It may be tempting to mark the difference between the property of being a lemon and being yellow by noting a difference in the way that their respective concepts are acquired. In order to possess the observational concept of yellow, it may seem, I need to have seen something yellow. Not so for an x-ray tube: many bits of knowledge may go into one's concept of an x-ray tube, but having seen one doesn't seem a necessary requirement for knowing what they are. As a matter of fact, I take myself to possess a (rudimentary) working concept of 'x-ray tube', although I have never seen one. It is enough for me to know what it is used for and how it works, and maybe to have heard descriptions of what it looks like. This means that perceptual acquaintance with a superficial observational property is *necessary* for the possession of its respective concept, while it is not for the possession of concepts of deep entities.

Despite its initial appeal, even this proposal is inadequate. In fact, the requirement of a perceptual encounter with the property in order to possess the respective concept is neither necessary nor sufficient for something to be a superficial observational property. This condition discriminates between recognitional and non-recognitional concepts, but we can have recognitional and non-recognitional concepts of both superficial observational properties

and deep entities. For instance, one can have a recognitional concept of a lemon, which requires a perceptual sensitivity to what lemons look like, and have a non-recognitional concept of colour. Therefore, the requirement of perceptual encounter over the possession of a concept only marks a distinction between the types of concept one may have and cannot be extended to the type of property that the concept is a concept of.

One might challenge the suggestion that one can have a non-recognitional concept of colour: after all, how can one really possess a concept of red if one hasn't seen red before? One may use the term 'red', but unless one has seen red, the use of the term doesn't signify that one knows what red, the superficial, visible property is. So the concept one possesses must grasp something else: not the superficial property of being red, but rather the reflectance property of materials, or the wavelength range corresponding to red. Thus, one might insist, we can indeed learn something about properties themselves by looking at the kind of concepts we employ to refer to them. One could then derive the following definition of superficial observational properties: they are properties for which one can only have a recognitional concept.

However, a famous thought experiment may help us appreciate that one can have a non-recognitional concept of colours. Jackson describes a woman, Mary, who was born and raised in a black and white room, and trained there—always through black and white media—to be a colour scientist. Mary doesn't leave the colourless room until she has learned all the physical facts about colours and colour perception. Only then does she

leave the room and see red for the first time. We can suppose that, even before seeing red, Mary had a superficial observational concept of red, although not a recognitional one. Mary might, for instance, have been able to entertain thoughts such as ‘The visible red light has a wavelength of about 650 nm’. Or, upon reading black and white versions of Goethe’s *Farbenlehre* and Kandinsky’s *Du Spirituel dans l’Art*, she might be in a position to claim that orange is more similar to red than it is to blue. When she utters these sentences, she is not (only) referring to physical properties. She is referring, it seems, to the observational property of red, the ‘visible red’.³³ Thus one can have a concept of an observational property without having had the relevant perceptual contact with such a property. Ordinary examples may illustrate this point too: blind people have concepts of colours: they know that lemons are yellow and tomatoes are red. The difference between Mary’s situation before and after leaving the room concerns her ability to use recognitional concepts for the observational property redness, not her ability to think about redness.

³³ Of course, a reductive physicalist would reject the suggestion that colours are superficial. However, as mentioned earlier, my aim here is not to defend a specific view of colour metaphysics, or a view where colours are superficial properties. As mentioned, my aim is to explore whether a satisfactory and coherent notion of superficiality is available—that is, whether one who wants to claim that colours are superficial can explain their superficiality coherently. For our current purposes, then, I do not need to worry about disagreement regarding the observability of colour, as long as the view is coherent. Whether this is the case is exactly what I assess here.

3.4 The Superficiality Constraint

I propose that we look instead at the difference between superficial observational properties and non-superficial entities in epistemic terms: there are different patterns of epistemic possibilities for claims concerning observational properties compared to claims about non-observational properties. This epistemic approach draws from Peacocke's (1983: 87–106, see also Peacocke 1986; 1992 and Wright 1982) discussion of observational concepts, although I shall not embrace all aspects of Peacocke's account. Peacocke suggests that observational concepts have a pattern of epistemic possibilities that is different from that of non-observational concepts, such as concepts of natural kinds, in the following respect: 'It is epistemologically possible that something which is tomato-like is not a tomato.' (Peacocke 1983: 94). Here Peacocke is interested in drawing a distinction between concepts such the concept of a natural kind 'tomato', which possesses an experiential component (tomato-like) that could be shared with something that is not a tomato, and observational concepts such as 'square', for which:

[t]here is nothing which is both the experiential component of the visual concept of being square and which could also be enjoyed by someone who does not possess all the concepts and the theories possessed by someone who has the concept of being square. (Peacocke 1983: 93)

For our present purposes, we can use this distinction to look at the conditions under which one might fail to distinguish between two things, for deep and superficial properties respectively, and what would allow one to distinguish them. For deep entities, as we have seen, it is possible to fail to discriminate two objects because we are accessing them through a certain source of information and the properties available through that source of information are identical. For example, a lemon and a lemon-shaped soap might be indistinguishable with respect to their visible properties. The failure to discriminate here is due to the fact that the piece of soap shares certain properties with the lemon. It might be the case that, even if one looked very carefully at the lemon-shaped soap and the lemon, one would keep failing to discriminate between the two, because the visible properties of each are exactly the same. However, it would be enough to explore the piece of soap through other sources of information to find further properties of the object that are not shared with lemon. Trying to eat it would unpleasantly reveal that it possesses a taste that lemons do not possess; trying to wash it under running water would reveal a capacity to dissolve and produce bubbles that lemons lack; and so on.

When it comes to superficial observational properties, this scenario is not possible: it is not possible to explain the failure to discriminate two distinct properties in terms of the properties being the same with respect to what is accessible through a certain source of information, but are different through another source of information. They don't seem to be accessible through other sources of information other than the designated sense-

modality. However, we have already seen that it is possible to fail to discriminate between observational properties. If I look at two vases in the dark, they might be indistinguishable with respect to their colour. What I ought to do in order to put myself in a position to distinguish between the two colours is not to look for further aspects of colours that are not accessible to perception. I wouldn't find any: all I would find, if colours are superficial as they seem, would be properties of the object that are related in some way to the superficial property of being coloured—such as a spectral reflectance profile—but I wouldn't find what colour the vases are. I might be able to infer that the colour of one vase is blue by learning about the reflectance properties of its surface, because I know how colours along the visible spectrum correlate with certain spectral reflectance profiles. But if I have never seen blue before, knowing that the vase has some reflectance properties won't be of any use in discovering the colour of the vase. There is, however, a very simple and efficient way to find out what colours the vases are—one that only involves perception. I just have to take the vases into the daylight, or turn the light on and look at them in better conditions. For observational properties, then, what motivates the failure of discrimination is the presence of some conditions that interfere with one's perceptual discriminatory capacities. These interfering conditions might be of different sorts: they might have to do with the context in which the perception occurs (the distance, the lighting conditions, and so on), with the perceptual system (the subject might have a sight deficit, or some neurological condition), or one might not be attentive enough to pick up a

difference with respect to observational properties that would otherwise be noticeable.³⁴

We can derive the following constraints about what can account for the failure to discriminate superficial properties:

The inability to distinguish among what are in fact different superficial properties must be due to the occurrence of interfering conditions that prevent the source of information with respect to which the properties are superficial to detect the difference.

This excludes that one could account for the failure to discriminate between superficial properties in terms of the presence of certain distinguishing features that are not accessible through the elective source of information with respect to which the entity is superficial. The converse of

³⁴ The sorites puzzle has been often used to argue against the existence of a coherent notion of superficiality (see for instance Goodman 1951; Travis 1985 and Dummett 1975). It might be thought that it presents a counterexample to the present proposal too. Ex hypothesis, sorites cases are cases where everything seems perfectly in order with the perceiver's perceptual system and the viewing conditions. Take the classical example of a series of patches of colours gradually changing in hue, where patch *a* is indistinguishable from *b*, *b* is indistinguishable from *c*, but *a* is distinguishable from *c*. All the successive patches are seen in a neutral, unchanging light, the subject is attentive and nothing is wrong with her sight. Yet she fails to see a difference between the colour of the first and the second patch, the second and the third one, the third and the fourth one, and so on. Although no interfering condition can be invoked here to explain the indistinguishability of non-identical superficial properties, what it takes to realise that the patches are different doesn't go beyond the ken of vision. We can tell that patch *a* is not identical with patch *b*, despite their perceptual indistinguishability, because we see an overall progression in hue across all the patches. This requires one to use some reasoning, but the reasoning is not the source of information: it only helps one to exploit at best the information provided by perception. Also, it might be possible to see immediately that any two adjacent patches are distinct by contrasting them with other background colours, which—thanks to the colour-contrast effect—might make the difference more salient.

this gives us a constraint about what may constitute the real nature of a superficial property:

What determines the nature of a superficial property must be accessible through the relevant source of information.

If we apply these constraints to phenomenal characters, the following constraint applies to any account of phenomenal character that seeks to preserve the intuition that phenomenal characters are superficial:

What determines the phenomenal character of an experience the nature of a phenomenal character of an experience should be accessible through introspection.

I shall call this the ‘superficiality constraint’. Notice that saying that what determines the nature of a phenomenal character should be accessible through introspection doesn’t mean that it needs to be accessible through introspection at any time. In the same way as one can fail to appreciate the real colour of something because of interfering conditions, one might fail to appreciate the phenomenal character of an experience because of interfering conditions—lack of the appropriate concepts, simple failure to focus attentively enough on one’s experience, a cognitive impairment that interferes with one’s introspective capacities, and so on. Moreover, it doesn’t mean that what determines the nature of phenomenal character is *only* accessible through introspection: the superficiality constraint is not necessarily incompatible with claiming that one can know about the

phenomenal character of one's own experience or of others' through non-introspective means—for instance by testimony.

3.5 Impersonal Indiscriminability

Can this constraint support IND? Well, it depends how one interprets 'indistinguishable'. As Williamson points out, indiscriminability is a relative notion: relative to a context, a subject, and the source of knowledge concerned. When we say that *a* is indistinguishable from *b*, this is always for a subject at a certain time. While introducing Williamson's notion of indiscriminability, Fish makes the following observation:

The subject relativity is there to account for the fact that you may be able to distinguish between two things where I cannot—perhaps because you have sharper eyesight than I do. The temporal relativity is there to account for the possibility that I may be unable to distinguish between two things at one time, yet be able to distinguish between them at a different time. For example, if I am looking at two people at a distance or in half-light, I may well be unable to make a discrimination. According to this definition, those individuals are therefore indistinguishable (for me, at that time). But bring them closer to me, or turn up the lights, and I may then be able to tell one from the other. So at this later time when the conditions under which I am attempting the discrimination have improved, the individuals

are now distinguishable (again, for me, at that time). (Fish 2008: 146)

Fish notices that *a* and *b* can be indistinguishable for one and the same person at one time and distinguishable at another time, depending on contingent conditions of observation. Fish mentions distance and lighting, but we can include other factors, including subjective ones, such as variability of one's eyesight (due, for instance, to surgery, corrective lenses, or decline in sight due to illness), variations in level of attention, etc. If we understand indiscriminability in these relative terms, a modal claim about the indiscriminability of two things turns out to be true every time one fails to discriminate them: they are indiscriminable for that subject, in those conditions, at that time, etc. This is manifest in the current debate on the argument from hallucination, where, in order to support the claim that hallucinations are indistinguishable from perception it is often thought to be sufficient to show one case in which one fails to discriminate it from genuine perception.³⁵ If we take indiscriminability claims to be made true by any instance of failure to discriminate, we have already seen that the relationalist has many reasons to reject IND: accepting the claim would be to accept that one can never fail to discriminate two distinct phenomenal characters.

³⁵ One of many examples is this quote from Fish, where he notices that 'the evidence for [the claim that hallucinations can be indistinguishable from perception] lies in the annals of literature, psychopathology, and in some cases even personal experience. There we find numerous cases in which subjects have sincerely taken themselves to be having an experience just like that of seeing something (we will focus only on the visual case) in situations where the thing apparently encountered was not there to be seen' (Fish 2008: 144).

However, we use indiscriminability claims in all sorts of ways: depending on the context, different factors to which the indiscriminability can be relative to will or will not be taken into account. Sometimes, if we are wondering whether something is indiscriminable to someone, we are not interested in knowing whether it is indiscriminable for her on this particular occasion, but rather whether she is unable to discriminate them in a set of conditions. We can thus make those conditions relative to which two things are discriminable for one more or less wide. For instance, we can wonder whether the sounds made by two birds are indistinguishable to a subject when they are heard for a long enough period of time, and the source of the sounds are close enough to the person. Even understood in these terms, it is still possible to admit that two indistinguishable sounds are nonetheless distinct, because interfering conditions intervene. It is possible that the auditory acuity of the subject is liable to declining with tiredness and the sounds are heard when the subject is tired, or that they are systematically inattentive to birds' sounds, that they have a hearing impairment, or that, from the position the subject listens to, there are always other noises that partly cover the birds' sounds and make them indistinguishable.

Also, notice that even if the subject is completely unable to distinguish between two properties at any time and in any conditions, claiming that the two observational properties are nonetheless different doesn't entail a violation of the superficiality condition. For my colour-blind friend, my scarf and my skirt are indistinguishable with respect to their colour; they are indistinguishable for him in any conditions, even the most

favourable ones. The same goes for someone who doesn't have access to the relevant source of information at all, such as a blind person. What makes the property nonetheless superficial is that, although the blind person personally has no means to perceptually distinguish two objects with respect to their colour, what he needs to do in order to know that they are distinct is to ask other subjects with normal sight. The blind person will have to defer to people who can see in order to distinguish the respective colours of the objects, but deferring to other seers still makes the nature of the properties accessible to perception, although not the perception of the person in question. The fact that *someone*, under some circumstances, could distinguish between the two colours is enough for them to be different, even if they are indistinguishable for the subject in question.

This means that one doesn't have to deny that phenomenal characters are superficial if she claims that two experiences have different phenomenal characters, although they are indistinguishable to the subject who has the experience: it might be that the subject has some impairment that prevents her from noticing a difference in phenomenal characters. But as long as the difference in phenomenal character is one that could be discovered through introspection, the superficiality condition is not violated. All that is required by the superficiality condition here is that if there is a difference, it must fall under the scope of introspection.

This suggests that the proponent of IND can hope to justify the acceptance of IND only if she conceives of the notion of indistinguishability in an impersonal sense.

We use indistinguishability in this sense in ordinary language. We might, for instance, say that two shades of blue are discriminable under these lighting conditions at this time, even though a given subject fails to discriminate them. If I need to patch my dress and I walk into a shop asking for a piece of fabric of a colour that is indiscriminable from that of my dress and I am given a piece of fabric that only the clerk fails to discriminate (because, say, she is colour-blind) or which is indistinguishable for me only in the lighting conditions of the shop, which turn out to be very poor, I won't take my request to be satisfied. I will think that the colour of the patch is not indiscriminable from that of my dress in the relevant sense, despite the fact that at a certain time and place I fail to discriminate them, or that a given subject is constitutively unable to discriminate them.

Martin (2004) uses a similar case to introduce an impersonal notion of indiscriminability: often we want to use 'indiscriminable' in a way that is not relative to the inability to discriminate of a particular subject. We instead want a notion of indiscriminability such that two things can be claimed to be indiscriminable only if *nobody* can distinguish between them. This 'nobody' needs qualification, though. As said, indiscriminability is relative to the source of information used, and it is likely that in this case we are interested in (standard and non-augmented) human sight (as we are interested in the sensitivity to colours of the potential members of the extended social community of those who can express judgments about the quality of my dress). So even if it turned out that Martians or other non-human animals with a different sensitivity to light wavelengths than us—or

even humans with exceptional sensitivity to certain wavelengths generally unseen by standard human sight, such as tetrachromats—turn out to be able to tell the colour of the patch from the colour of my dress, or if there were a special device capable of extending the range of wavelengths to which we are sensitive, we would still consider the patch and the dress indistinguishable by human sight with respect to their colour.

The proponent of IND can argue that (at least some) hallucinations are indistinguishable through introspection from perception in an impersonal sense. For brevity I will from now on call these kinds of hallucinations ‘IIH’: Hallucinations that are Impersonally Indistinguishable through Introspection. In these hallucinations, it is not the case that there is some interference or disabling condition that prevents one from discriminating a difference that would be otherwise accessible through introspection. If this is the case, it is difficult for the relationalist to claim that the IIH has a different phenomenal character than perception without violating the intuition that phenomenal characters are superficial. It seems that the impersonal notion of indiscriminability rules out the possibility of explaining the introspective indiscriminability of a hallucination from a perception without assuming that the hallucination has a real nature that is inaccessible through introspection—which would amount to violating the superficiality constraint.

But as we have already seen, it is not clear that one can deny the superficiality of phenomenal character without departing from what the notion of introspection is designed to capture in the first place: how the

experience introspectively strikes the subject of the experience. Thus, unless the relationalist explains why it is possible for the phenomenal character of an IIIH to be different from that of a genuine perception without committing to the idea of phenomenal characters, one should accept IND.

Reflecting on the nature of superficial properties thus poses a serious challenge for the relationalist. This challenge seems difficult to avoid, because it doesn't rely on any controversial claim. It doesn't rely on the idea that introspection is infallible, nor on any controversial assumption about superficial properties being such that for them indiscriminability entails identity. All it relies on is what seems a very minimal requirement imposed on the notion of superficiality, a requirement whose denial would make it difficult to preserve any intuitive distinction between superficial and non-superficial properties. This requirement holds that it is not possible for a superficial property, to have its nature determined by something that is inaccessible through introspection. But the relationalist is committed to claiming that hallucination does not have the same phenomenal character as perception. In the case of an IIIH, there is no way to distinguish hallucination from perception through introspection, and it is not possible to attribute the subject's inability to discriminate to interfering conditions. It thus seems that the relationalist is committed to claiming that what makes the phenomenal character of hallucination different from the phenomenal character of the genuine perception from which it is indistinguishable depends on some aspects of it that are inaccessible through introspection—an obvious candidate for this difference would be facts that have to do with

the world around one: the fact that the relevant object is not present and doesn't determine the phenomenal character of the experience.

One fails to understand what is really at stake in the disagreement between relationalists—with their commitment to disjunctivism—and their opponents unless one focuses on the intuition that phenomenal characters are superficial. Often opponents lament that disjunctivism is committed to 'absurdities', although it is not always clear what this absurdity is supposed to amount to. Knight, for instance, claims that the 'disjunctivist strategy requires that we adopt an extreme form of skepticism with respect to our knowledge of our mental states' (Knight 2013: 3). Knight seems to suggest that disjunctivism is unacceptable because it requires the assumption that we can be introspectively wrong about our mental states in a very radical way. However, it is not clear why admitting that introspection can go radically wrong is unacceptable, unless we subscribe to some form of infallibilism. But most opponents of disjunctivism would not be outraged by the admission that introspection can go radically wrong with respect to all sorts of self-attributions of mental states. So why is maintaining that the phenomenal character of hallucination is not what it seems considered so absurd?

The discussion of this chapter explains why. Disjunctivism seems unacceptable because it seems committed to violating the superficiality of phenomenal characters, the intuition that whatever determines the nature of a phenomenal character must be accessible through introspection. We should understand Knight's accusation of a too extreme scepticism in this

sense: disjunctivism requires one to claim that what is, by definition, necessarily accessible through introspection (i.e. the phenomenal character of an experience) is in fact not accessible through introspection. This explains why what Knight calls an extreme form of scepticism is considered absurd—while all kinds of extreme sceptical claims have been made in the history of philosophy without generally being deemed absurd, only undesirable. What's really at stake here is not the disjunctivist's commitment to scepticism about introspection, but it's violation of the superficiality constraint.

Martin (2006) also presents opponents of disjunctivism as committed to a too immodest view of introspection. In this way he frames the disagreement between proponents and opponents of disjunctivism in terms of a disagreement about the powers of introspection. Although, locally, this might be what motivates some opponents of disjunctivism, this cannot be where the fundamental source of disagreement really lies, as I have argued in this chapter. What fundamentally motivates the rejection of disjunctivism is the intuition that phenomenal characters are superficial and disjunctivists cannot account for the phenomenal character of hallucination without violating this intuition.

Disjunctivists are either seen as committing to a blatantly incoherent view—they claim that phenomenal characters are superficial, yet they give an account of the phenomenal character of hallucination that is incompatible with their superficiality—or they surreptitiously change the rules of the game and the meanings of the terms used in the debate. The relationalist

strategy might seem somewhat similar. First the relationalist claims that the phenomenal character of genuine perception is constituted by the mind-independent objects we seem to perceive. She motivates this claim by appealing to introspection. She thus seems to comply with the intuition that phenomenal characters are superficial. Then the opponent points out that in hallucination, where there is no suitable object that can constitute the phenomenal character of hallucination, experience also strikes one introspectively exactly in the same way: mind-independent objects seem to be presented. To this the relationalist replies that the phenomenal character of perception is not what it seems to be through introspection. The opponent retorts that maybe the relationalist doesn't appreciate the full extent of the consequence of her claim. Since *ex hypothesi*, hallucinations are impersonally indistinguishable from phenomenal characters, she commits herself not only to the claim that the phenomenal character of hallucination seems introspectively different to what it actually is (which could be acceptable), but also to the claim that the phenomenal character of hallucination really is inaccessible through introspection, which violates the intuition that phenomenal characters are superficial. At this point the relationalist might reply that, if she is committed to denying the superficiality of phenomenal character, *tant pis* for the superficiality. What *she means* by 'phenomenal character' is the property of relating the subject to certain objects in the environment and their features. If this violates the superficiality, well... phenomenal characters are not superficial after all. If the relationalist adopts this strategy, she can be rightly accused of trading

with the terms of art used in the debate. Everybody assumes that it is part of the notion of phenomenal characters that they are superficial, because this follows from the idea that phenomenal characters just capture what it is like to have that experience. The relationalist may well stipulate that the phenomenal character doesn't need to be superficial. But then it seems they are talking about something else, which leaves it open what do they have to say about the phenomenal character (understood as superficial) of perception and hallucination, respectively.

The tenability of relationalism depends on whether relationalists find a way to account for the difference in phenomenal character between an IIIH and perception that doesn't violate the superficiality constraint. However, it is not obvious how this can be done. In order to make this difficulty more explicit, we can formulate these worries in the form of an independent argument against relationalism. I will call this argument 'argument from superficiality':

1. Some hallucinations are IIIH;
2. An IIIH is phenomenally conscious;
3. (from 2) IIIHs have a phenomenal character;
4. Phenomenal characters are superficial.
5. If phenomenal characters are superficial, the phenomenal character of an experience is accessible (at least in principle) through introspection (superficiality constraint);
6. Relationalism claims that the phenomenal character of an IIIH is different from the phenomenal character of a perception;

7. (from 6) For the relationalist, the phenomenal character of an IIIH is different from how it introspectively seems: its real nature is hidden to introspection;
8. (from 2 and 7) Relationalism is committed to the claim that the phenomenal character of IIIH is not superficial;
9. (from 4 and 8) Relationalism is untenable because it denies the superficiality of phenomenal characters.³⁶

What strategy could the relationalist adopt to resist this argument? We have already seen that trying to resist (4) and reject the idea that phenomenal characters are superficial is not a viable option. A relationalist could adopt different strategies to counter this argument. She could try to deny the first premise. If she shows that there is no IIIH, all cases of indistinguishable hallucinations could be explained in terms of interfering conditions preventing one from introspectively discriminating difference that are indeed accessible through introspection, although not on that particular occasion. I will explore this strategy in the next chapter.

However, in Chapter 5 I will show that, although there is indeed no empirical evidence that any IIIH exist, the relationalist is not in a position to deny the metaphysical possibility of IIIHs. Another option is to deny the third premise: that hallucinations have any phenomenal character. This

³⁶ Soteriou (2013: 198–206) presents a similar challenge for the relationalist. However, the challenge for Soteriou doesn't arise from the assumption that phenomenal characters are superficial or from a consideration of the conditions imposed on superficiality. For Soteriou the problem arises from what we can take to be the ground for claims about phenomenal characters (see also Soteriou 2005: 183).

strategy is actually adopted by some disjunctivists, in particular Fish (2008; 2009), but is sometimes also erroneously attributed to Martin (see for instance Knight 2013). In Chapter 6 I will show why this strategy is unsatisfactory. Ultimately, in Chapter 6, I will argue that the strategy that the relationalist should pursue to block this argument is to deny premise (7), showing that it doesn't follow from (6).

3.6 Conclusion

In this chapter I explored what the proponent of the argument from hallucination needs to commit to in order to support IND. I considered two assumptions from which, if accepted, IND would logically follow. These assumptions are a commitment to the infallibility of introspection, on one hand, and to the idea that phenomenal characters cannot seem different than they are. I argued that both these assumptions are ill motivated and can be resisted by relationalists. However, although the assumption that there is no seem/is distinction for phenomenal characters should be rejected, we can see it as a (unsatisfactory) way of cashing out a valid intuition: phenomenal characters are superficial, in the sense that their nature seems entirely determined by what is accessible through introspection. This assumption seems to be a constitutive aspect of the notion of phenomenal character, which is introduced precisely to single out those aspects of experience that are accessible to the subject who undergoes the experience through introspection.

I have argued that, without making further controversial assumptions (such as the claim that introspective judgments can never be false, or that nothing can interfere with the way a phenomenal character introspectively appears to one), the superficiality of phenomenal character provides the proponent of IND with a solid argument in its favour. The superficiality of phenomenal characters imposes that, if there is a difference between the phenomenal character of perception and hallucinations that are impersonally introspectively indistinguishable from perception (IIH), it must be in principle detectable through introspection. But it is part and parcel of the notion of impersonal indiscriminability through introspection that no difference can be detected through introspection. So the relationalist seems committed to denying that the phenomenal character of an IIH is superficial: she seems committed to saying that an IIH seems to have a phenomenal character that is not accessible through introspection. What we access through introspection is only the phenomenal character that the IIH seems to have. But *real* phenomenal character, whatever this amounts to, is hidden to introspection—which seems absurd.

The only controversial assumption that this argument makes, it seems, is that there are indeed IIHs. The existence of IIHs cannot simply be motivated by the platitude that we often fail to tell that a hallucinatory experience is not a genuine perception: as we have seen, a simple failure to discriminate two things doesn't entail that the two things are impersonally indistinguishable. So we need further evidence for the claim that there are IIHs. In the next chapter I will explore whether there is any such evidence.

Chapter 4: Hallucinating, for Real

In the previous chapter I argued that what underwrites the acceptance of IND is the idea that phenomenal character are superficial and that, in case of hallucinations that are impersonally indistinguishable from perception (IIHs), there is no way for the relationalist to claim that the phenomenal character of that experience is different from that of perception without committing to the idea that what determines the reality of the phenomenal character of the hallucination is inaccessible through introspection—which would amount to denying the superficiality of phenomenal characters.

This argument against relationalism, however, works only if it is possible to show that there are indeed IIHs. In all other cases of hallucinations, it is open to the relationalist to explain the difference in phenomenal character between hallucination and perception by claiming that interfering conditions prevent one from noticing the differences in phenomenal characters that would in principle be accessible through introspection.

In this chapter, I will explore whether there are indeed any IIHs. If it turns out that there aren't any, the argument from hallucination and the related argument from superficiality are unsound, as one of their premises is

false. I will examine both theoretical considerations—in order to understand what it takes for a hallucination to count as an IIIH—and empirical considerations, to establish whether any of the hallucinations reported in the literature meet these conditions. I will conclude that it is not possible to find empirical support for the claim that IIIHs actually exist.

4.1 Inistinguishable from What?

If my analysis of the current debate on disjunctivism and the argument from hallucination is correct, both parties tend to overlook the fact that IND holds only if indistinguishability is understood in an impersonal way, and the failure to notice this results in little attention being paid to what we mean when we accept that hallucinations are indistinguishable from perception—as well as what follows from this.

As a result, people accept as obvious premise (2) in the argument from hallucination (that some hallucinations are indistinguishable from perception) without discussing in what sense they are or can be indistinguishable. However, premise (2) is obvious only in the weak sense that sometimes people who hallucinate think that they are perceiving. But if we understand (2) in this sense, the argument is a *non-sequitur*. The proponent of the argument from hallucination, therefore, has to show that the following revised version of premise (2) is true:

(2*): Some hallucinations are *impersonally* indistinguishable through introspection from genuine perception (some hallucinations are IIIH).

Since proponents of the argument from hallucination have so far failed to realise that it is vital for the success of the argument that hallucinations be impersonally indistinguishable from genuine perception, there is currently no discussion of whether (2*), rather than merely (2) (which does not specify in which sense the notion of indistinguishability is used), obtains.

Establishing whether (2*) is true is far from easy. It is not even clear how we should approach this issue, what we need to look for, or what sort of phenomena might confirm the claim that hallucination is impersonally indistinguishable through introspection from perception. A preliminary question we need to answer to work towards an answer is the following. What is the logical structure of the relation of indistinguishability that is predicated in this claim?

The go-to reference for discussion of the structure of indiscriminability is Williamson. He claims that:

a is indiscriminable from *b* for a subject at a time if and only if at that time the subject is not able to discriminate between *a* and *b*, that is, if and only if, at that time the subject is not able to activate (acquire or employ) the relevant kind of knowledge that *a* and *b* are distinct. (Williamson 1990: 8)

Williamson stresses that this is not a definition, in the sense that it doesn't provide necessary and sufficient conditions for the correct use of the ordinary English word 'indiscriminable'; it merely allows us to identify

paradigmatic cases of indistinguishable entities (which might be objects, events, or properties). Around these paradigmatic cases, there might be less clear cases for which we might wonder whether it is appropriate to use the notion of ‘indiscriminability’. The paradigmatic case that Williamson has in mind here are cases in which indistinguishability takes the form of a one-to-one relation between particulars. Williamson notices that, in this paradigmatic form, the relation of indiscriminability has the following properties:

- (1) *Reflexivity*. Since a cannot be distinct from itself, and since knowledge is factive, it is impossible to activate the knowledge that a is distinct from itself, a is necessarily indiscriminable from itself.
- (2) *Symmetry*. If one is unable to activate the knowledge that a is distinct from b , one is also unable to activate knowledge that b is distinct from a .
- (3) *Non-transitivity*. It is a well documented phenomenon that there is what is called a phenomenal sorites: there can be a series of items $s_1 \dots s_n$, presenting a smooth, continuous, transition with respect to some perceptual property (colour, shape, pitch, taste...) such that s_1 is indistinguishable with respect to some perceptual dimension from s_2 , s_2 is indistinguishable from s_3 , but s_1 is distinguishable from s_3 .

However, as Williamson himself remarks, the relation of indistinguishability can take a different form and relate more than two terms and terms that are not particulars. The question we need to ask is what form the relation of indistinguishability predicated in (2*) takes. The superficial structure of the sentence suggests that this is a one-to-one relation between a

hallucination and a perception. But what are the hallucination and the perception in question? What the hallucination is supposed to be is fairly clear: it is whatever hallucination one is undergoing. But what is the perception?

It cannot be a genuine perception that one is having, as one does not perceive and hallucinate the same thing at the same time. Maybe cases of polyopia³⁷ could be interpreted as cases where one can directly compare a genuine perception and a hallucination, but subjects undergoing polyopia systematically describe the copies of the object that they experience as different in shape, size, and definition from the original object (Jones, Waggoner, and Hoyt 1999). That is to say, in such cases the hallucinated copy *is* indeed distinguishable from the real object (and if it is indistinguishable, it is not impersonally indistinguishable).

It is not even the case that, when we talk about hallucinations that are impersonally indistinguishable through introspection, we are comparing a current hallucination with a veridical perception from the past. Although there might be cases of hallucination where one hallucinates scenes from the past, these cannot be cases of impersonal indistinguishability, since it is not possible to exclude that the inability to discriminate is attributable to memory failure rather than introspective failure (see Farkas 2006: 211–213).

Dorsch claims that we should understand the relation predicated in (2*) as a one-to-many relation (the label is mine), in the sense that it claims

³⁷ Polyopia is the phenomenon of seeing two or more images (up to hundreds) arranged in ordered rows, columns, or diagonals after fixation on a stimulus. It may occur both in monocular and binocular vision.

that the hallucination is indistinguishable from each of the members of the class of perception. He notices that we use indiscriminability in this sense in many ordinary situations, in sentences like: ‘I cannot tell from his looks whether he belongs to that community’ (Dorsch 2013: 183). In both cases, indiscriminability turns out to be a relation between more than two entities. And in the latter case, it stops being symmetrical in any meaningful sense. For it is not necessary that any member of the comparison group is itself indiscriminable from anything other than the entity originally compared with the group. In particular, the members of the group need not be indiscriminable from each other (see Dorsch 2013: 184).

Dorsch is certainly right to point out that when we talk about indistinguishability of hallucination we mean something along the lines of not being able to tell that the hallucination does not belong to the class of perceptions. However, the way Dorsch spells it out in terms of the hallucination’s not being distinguishable from any of the perceptions in the class of perceptions (a one-to-many relation) doesn’t seem to me like the most helpful way to spell this out. The problem with this proposal is that it is not clear what the class of relevant perceptions is supposed to be. Does it contain perceptions the subject has had in the past, or any perception that human beings have ever had? Or does it contain possible perception? I do not see any reason for restricting this class, or any way of deciding non-arbitrarily how to restrict it, so I suppose the class should include all possible perceptions. This would mean that the relation is to be understood in counterfactual terms: if one were to compare one’s actual hallucination

with any possible perception, one would not be able to distinguish it through introspection from any genuine perception. But this is clearly false. A hallucination of a hammer is indeed distinguishable from a perception of a squirrel or a chair. One might try to solve the problem by circumscribing the class of perceptions from which hallucination is indistinguishable to the class of perceptions of a certain kind of objects *F*—a hallucination of a hammer may be indistinguishable from a perception of a hammer. But perceptions of hammers can differ greatly from each other phenomenologically and so it cannot be the case that the hallucination in question is indistinguishable from each one of them. One could then propose that when we say the hallucination of an *F* is indistinguishable from perceptions of an *F*, ‘an *F*’ stands for a more detailed specification than something like ‘a hammer’: the distance and perspective from which it is seen, the colour, the shape, the background, and all other relevant details should be specified. However, even if one adds many specifications to the description of the *F*, several perceptions with different phenomenology can fit the same description: they may have slightly different shades of the same colours or be more or less illuminated.

It seems that the only way to narrow down the members of the class of perceptions of an *F* to only those from which the given hallucination is impersonally indistinguishable would be to take only the perceptions whose phenomenology matches the hallucination. This would amount to assuming that we should understand the relation of impersonal introspective indiscriminability between hallucination and perception in terms of

sameness of phenomenal character.³⁸ But we should avoid a characterization that entails IND.

I think a better way of understanding the relation of impersonal indistinguishability is to say that the hallucination is indistinguishable from any perception with respect to the property of belonging to the class of perceptions. This doesn't present the problem of having to narrow down the relevant class of perceptions and better captures what we have in mind when we say that one is unable to tell that a hallucination is indistinguishable from perception. As Martin notices, saying that a hallucination of a picket fence is introspectively indistinguishable from perception means that the hallucination 'is such that it is not possible to know through reflection that it is not one of the veridical perceptions of a white picket fence' (Martin 2006: 364).

To say that a hallucination is impersonally indistinguishable from perception is to say that one is unable to tell from introspection that the hallucination does not belong to the class of perceptions. Understood in this sense, the relation of indistinguishability is not a one-to-many relation, as the hallucination doesn't need to be indistinguishable from any individual perception. It only seems to belong to the same class as they do. The relation is rather a relation between the occurring hallucination and the class

³⁸ This view is largely assumed by opponents of disjunctivism, but it is rarely defended. See Farkas (2006) for a defence of the idea that the only way to make sense of indistinguishability is in terms of sameness of phenomenal character. The discussion in this chapter can be seen as an indirect refutation of her argument. For one thing, I argue that there is indeed an alternative way to understand the relation of indistinguishability holding between perception and hallucination. Moreover, Farkas fails to take into account various aspects of the relation of indistinguishability that my discussion highlights, such as its being a relation holding between an individual and a class.

of (all) perceptions: the relation is one of not being able to tell that the former doesn't belong to the latter. The claim that a hallucination is indistinguishable from perception, when properly understood, amounts to the inability to correctly categorize the experience as a hallucination rather than as a perception. There does not need to be any specific perception from which the hallucination is indistinguishable in order for it to be considered a member of the class of perceptions. This makes the relation asymmetric: the hallucination is indistinguishable from perception in the sense that one cannot tell that one's experience doesn't belong to the class of perception, but there isn't any sense in which the class of perception is indistinguishable from the hallucination.

4.2 Hallucination: A Fuzzy Notion

The proposed reading of indistinguishability is close to certain characterizations of hallucination to be found in clinical and experimental psychology, where the notion of indiscriminability is entirely absent, but instead hallucinations are mainly classified on the basis of the degree of success or failure in reality testing and presence or lack of an insight into the unreal nature of the hallucination.

Looking at attempts made by psychologists to define hallucination is instructive for philosophers, who tend to underestimate the variety of hallucinations, treating them as a monolithic mental kind identified with a simple definition. Aleman and Larøi (2008: 16) indicate the following as the most accurate definition of hallucination, although they acknowledge that

this definition also has its limits and that it might not be possible to reach an exhaustive definition:

A sensory experience which occurs in the absence of corresponding external stimulation of the relevant sensory organ, has a sufficient sense of reality to resemble a veridical perception, over which the subject does not feel s/he has direct and voluntary control, and which occurs in the awake state.

(David 2004: 108)

This definition only captures paradigmatic examples of hallucination, leaving out certain phenomena that are considered hallucinations both in ordinary language and for diagnostic purposes, and including phenomena that are not hallucinations. Among the weaknesses of this definition are the following:

1. The clause ‘in absence of corresponding external stimulation’ is supposed to distinguish hallucination from illusion. However, many cases of hallucination are triggered by an external stimulus. For instance, someone distraught might hear a menacing voice in the blowing of the wind. Depending how we understand the ‘corresponding external stimulation’, these phenomena will count as illusion or hallucination: illusion if we take them to be misperception of the external stimulus, or hallucination if we consider that the external stimulus doesn’t correspond to the experience (because too different from what is presented). Similarly, cases of polyopia might

equally well be considered as either hallucinations or illusions, depending on whether one considers that a real object triggers the pattern of multiple copies (illusion), or that there are multiple visible objects, while there is only one real object (which makes the various ‘copies’ of the object hallucinatory) (see ffytche 2013: 59).

2. The reference to sufficient resemblance to perception is meant to distinguish hallucination from imagination.³⁹ However, it is not clear that hallucinations always bear a stronger resemblance to perception than imagery does. A well-known fact among psychologists since the pioneering work of Francis Galston in the nineteenth century is that different subjects present a striking variety in their respective ability to visualize (Thomas 2009). Some, whom Phillips (2014) calls ‘super-imagers’, report to be able to easily imagine scenes with the apparent

³⁹ The definition refers to resemblance with respect to the sense of reality. However, the notion of a ‘sense of reality’ is as ubiquitous as elusive, so much so that philosophers (Austin 1962: 59) and psychiatrists (Sarbin 1967) alike have at times recommended abandoning it. Even Jaspers (1912), who first stresses the importance of the sense of reality of some hallucinations for diagnostic purposes, laments the lack of clear criteria for the notion. However, Aggernaes (1972: 222–227) systematised the various criteria proposed in the literature in the following criteria: (1) *sensation* (‘if it seems to the subject that the experience is due one’s sensory organs, and not merely imagined or thought’); (2) *behavioural relevance* (‘the patient or subject experiences this something as being of relevance for his emotions and/or needs, and/or actions, in the actual or in some potential situation’ 223); (3) *publicness*: (‘if the experiencer is aware [...] that anybody else possessing normal sensory faculties would be able to perceive this something with his senses if he were within reach of it’); (4) *objectivity questioned* (‘if the experiencer feels that under favourable circumstances, he would be able to experience the same something with another modality of sensation than the one giving the quality of sensation’); (5) *existence* (‘if the experiencer is aware [...] that he feels certain that this something also exists when nobody experiences it at all’); (6) *involuntarity* (‘if the experiencer feels that it is impossible or extremely difficult for him to alter or dismiss the experience simply by wishing it to be altered or dismissed’); (7) *independence of a quite unusual mental state in oneself* (‘if the experiencer is aware [...] that this experience is not simply the result of his being in a quite unusual state [...] such as] psychosis, ‘bad nerves’, very intense emotional state, and drug-withdrawal state’).

richness and vividness of normal vision.

3. The aspect of voluntary control is meant to add a further distinction from imagination, but here too we find that not all hallucinations are completely beyond voluntary control. On the contrary, teaching patients—especially with schizophrenia—how to control their hallucinations is a coping strategy widely adopted by practitioners (see for instance Jimenez et al. 1996; Badcock and Hugdahl 2014). On the other hand, patients with anxiety disorders, depression, eating disorders, and psychosis frequently report the intrusion of extremely vivid, detailed, and uncontrollable visual imagery (see Tudor et al. 2013).
4. The condition of wakefulness is meant to exclude dreams from the definition. However, episodes like hypnagogic and hypnopompic hallucinations—which occur respectively at sleep onset and at awakening—could be considered either hallucinations or dreams.

The unclear boundaries between hallucinations and many other mental occurrences suggest that at least some hallucinations could be not sensory experience, but rather instances of other types of mental occurrences that one mistakenly takes for sensory experiences. One of the most credited theories of verbal hallucination in schizophrenia, for instance, explains hallucination in terms of misattribution of inner speech to an external agency (see for instance Allen, Aleman, and McGuire 2007; Moseley, Fernyhough, and Ellison 2013). In all these cases, the indistinguishability from perception simply amounts to a meta-cognitive

error: the subject self-ascribes a kind of mental state, while in fact she has another kind of mental state. None of these cases count as impersonally indistinguishable cases that can support IND.

This doesn't exclude, of course, that *some* hallucinations are genuinely experiential. We need now to consider whether, among these, there are some that are introspectively indistinguishable from perception.

4.3 What Counts as Impersonally Indistinguishable from Perception

All sorts of different phenomena go under the name of 'hallucination'; from very simple phosphenes and tinnitus to the complex multi-modal scenes experienced by some psychotic patients, hallucinations vary greatly in their aetiology, phenomenology, and population affected. When researching for IIHs among these multiform phenomena, we can proceed by exclusion.

A first notable fact is that not all hallucinations are confused with perception by the subject who undergoes them.⁴⁰ Many subjects preserve

⁴⁰ One might contend that failure to discriminate *tout court* doesn't exclude that the hallucination in question is *introspectively* indistinguishable from perception. As far as we know, one might preserve insight into the unreal nature of their experience in virtue of extra-introspective clues, rather than on the basis of introspection. One might for instance reason that the content of one's hallucination is very odd and incoherent with one's previous experiences and conclude that one must be hallucinating, although, solely on the basis of introspection, one wouldn't be able to do so. Or one might be told that one is hallucinating. However, since we are not interested here in deciding whether this or that hallucination count as introspectively indiscriminable, but only with whether there is any hallucination that is impersonally introspectively indistinguishable from perception, and since there are many cases of hallucination in which the subject believes she is perceiving, we had better focus on the latter, rather than looking at those hallucinations in which one preserves some insight into the hallucinatory nature of one's experience, and examine whether the ability to discriminate is due to introspective or extra-introspective clues.

the insight that what they are experiencing is not real.⁴¹ We thus need to focus on hallucinations in which the subject doesn't preserve insight into their unreal nature. What would it take for a hallucination to count as *impersonally* indistinguishable through introspection from perception? As per the notion of impersonal indistinguishability defined in the previous chapter, it must be the case that the difference between the hallucination and perception, if any, is such that it doesn't fall under the scope of introspection.

Many cases of hallucinations fail to meet this condition. It seems that, if only the subject were more attentive, for instance, she could discriminate her current hallucination from cases of perception. For instance, all the cases of polyopia, palinopsia—including afterimages⁴²—, phosphenes, and hallucinations in migraine aura⁴³—although they might

⁴¹ At the onset of the systematic scientific study of hallucination in the 19th century some thinkers tried to systematise the notion of hallucination and restrict it to those experiences where the subject lacks insight into the hallucinatory nature of her experience, calling experiences where one lacks such an insight 'pseudohallucinations'. However, since its introduction, the meaning of the notion has always oscillated between two uses: that of a hallucination with insight and that of vivid imagery (See Berrios 1996: 49–59 for an overview of the history of the notion). The notion of pseudohallucination is still occasionally used in the clinical practice with both meanings. However, the current consensus among neuroscientists and clinicians is that the notion of pseudohallucination 'is of little theoretical and clinical importance' (Aleman and Larøi 2008: 22), 'gives rise to much confusion, and is weak because it is defined negatively' (Berrios and Dening 1996), and has 'low construct validity and [is] accordingly clinically ambiguous' (Van der Zwaard and Polak 2001). Thus, nowadays most clinicians and psychologists apply the term 'hallucination' irrespective of whether the object experienced is located externally or is taken to be real.

⁴² Palinopsia is the persistence or recurrence of visual images after the stimulus has been removed. It is a non-specific term that describes multiple types of visual symptoms, including afterimages, light striking, and visual trailing, and is linked to a variety of aetiology.

⁴³ An aura is a perceptual disturbance that often precedes the onset of the migraine or seizure (although sometimes the aura is not followed by a manifest seizure or migraine). Perceptual disturbances include the occurrence of non-

occasionally be taken by the experiencer to be cases of genuine perception—present introspectible features that are not to be found in genuine perception: many hallucinations differ significantly from perception with respect to vividness, sense of reality, complexity, etc. The objects presented in these hallucinations do not appear to be material objects and do not seem to occupy the space in the same way as material objects do; they sometimes superpose them without occluding them entirely, they do not exhibit size, colour, and other forms of perceptual constancy and, in the case of phosphenes, afterimages and some hallucinations in migraine aura, they remain visible even when one closes one's eyes.⁴⁴ A subject who fails to realize that she is hallucinating would have been able to do so if she had paid attention and reflected more on her experience, for example like someone who fails to distinguish a mending done poorly but would have noticed it if she had looked more carefully.

Many cases of more complex hallucinations have a dream-like character or an otherwise fuzzy, confused, incoherent, indeterminate character. Epileptic patients, for instance, often describe the hallucinatory experiences that precede a seizure as dream-like or memory-like, and these episodes are called a 'dreamy state' by some clinicians (Vignal et al. 2007). Hallucinations induced by LSD and other psychotropic drugs, but also in

complex hallucinations in different modality, such as a strange smell, a flash of light or visual patterns, the feel of a pressure, and noises. It also includes deformed experience, such as visual alteration in the contours of objects or a change in the experience of environmental light.

⁴⁴ See Phillips 2013 for a discussion of afterimages and how the relationalist can accommodate them; see Sacks (1985; 2012: 122–132) and Schott 2007 for a description of migraine aura; Gersztenkorn and Lee 2015 for a literature review on palinopsia.

some psychotic states, are described as more vivid and bright than any genuine perception ever experienced (see Huxley 1954; Sacks 2012: 90–121; Duppen 2015).

How about hallucinations that don't display these phenomenal differences and seem to have all the traits that we generally find in perception—experiences where the object presented seem constant, persistent, perceived as existing in the external environment, that present the same vividness and coherence as genuine perceptions? It is not clear that the presence of all these conditions in a hallucination indicate that it is an IIIH.

The problem is that it is not clear what would count as evidence for this claim. In order to prove that certain hallucinations are impersonally indistinguishable from perception, we have to show that, in such cases, there isn't anything that interferes with the subject's capacity to tell that she is not in fact perceiving. Many hallucinations are vivid, described by the subject as entirely similar to perception, and believed to be genuine perception, but there is no way to verify that nothing is going wrong with some cognitive faculty of the subject or with some other condition that, if absent, would enable the subject to distinguish her experience from perception.

On the contrary, there is extensive evidence that some cognitive impairment or some state of distress can influence the subject's insight on the hallucinatory nature of her experience. The incidence of hallucinations where the subject lacks insight into the unreal nature of her experience is much higher in the clinical population with psychiatric or neurological disorders such as schizophrenia, Alzheimer's syndrome, or peduncular

hallucinosi⁴⁵, or in altered states due to psychoactive drugs, where the hallucination is associated with cognitive and behavioural abnormalities, and emotional distress—which are likely to impair one’s ability to tell a hallucination apart from normal perception. On the contrary, in hallucinations in non-psychiatric populations such as patients with sensory deficits, e.g. Charles Bonnet syndrome (Eperjesi and Akbarali 2004),⁴⁶ in the large majority of cases the patient is fully aware of the hallucinatory nature of her experience.

This suggests, albeit in a non-conclusive way, that cognitive impairments may, unbeknownst to the subject, interfere with her ability to discriminate her hallucination from perception.

Even in cases where the non-psychiatric patient takes herself to be perceiving, we are not in a position to know whether some factor interferes with her capacity to judge about her current state and whether, were this factor to be removed, she would still not be able to tell her hallucination

⁴⁵ Peduncular hallucinosis is a rare neurological disorder that causes vivid and very realistic visual hallucinations that typically occur in dark environments and last for several minutes. The hallucinations often involve people and environments that are familiar to the affected individuals, and patients can rarely distinguish between the hallucinations and reality. Peduncular hallucinosis occur in patients with serious lesions in the frontal cortex and are often associated with dementia, Alzheimer’s syndrome, and delirium tremens (see Benke 2006).

⁴⁶ Charles Bonner syndrome is a common condition among people who have lost their sight entirely or to a great extent and consists in the occurrence of visual hallucination (very rarely accompanied by sound) of simple repeated patterns or complex hallucinations, often with repeated figures (animals, people, faces—often deformed—landscapes, buildings) scrolling across one’s visual field repeatedly. The hallucinations are typically well-defined, organized, and clear, and the patient has little or no control over them. They are usually pleasant or emotionally neutral, but may sometimes cause distress if the figures appearing are frightening, such as scary animals or deformed figures. It is believed that they are due to spontaneous activity in the visual cortex that hasn’t yet accommodated to the lack of stimuli (See Manon et al. 2003 for an overview).

apart from perception. Take a case of hallucination in a patient with Charles Bonnet syndrome, where the subject thinks she is genuinely perceiving. It could be the case that she fails to recognize herself as hallucinating because she is, for instance, surprised or scared by her experience and isn't paying sufficient attention to its character. There is no way to exclude any such interfering condition.

While there is sufficient empirical evidence to claim that in many cases an interference disrupts the subject's ability to discriminate her hallucination from perception, no clear evidence currently seems to be available for the existence of cases where no such interference is at play. Hence, so far, premise (2*) is not empirically supported. Since premise (2*) is meant to be an empirical claim, it cannot be accepted unless further empirical studies show that in certain cases of hallucination we can rule out the intervention of an interference of some sort as the cause of the subject's inability to tell her experience apart from genuine perceptions.

The only way in which we can introduce the case for the possibility of hallucinations that are impersonally indistinguishable from perception is via a stipulation: one should simply assume that some hallucinations are impersonally introspectively indistinguishable from perception. In the next chapter I will consider whether good reasons exist for making such an assumption. In the remainder of this chapter I will consider how these considerations bear on the way we should think of the challenge that hallucinations pose to relationalists.

4.4 The Problem of Hallucination is Not a Phenomenological

Problem

Whether or not this stipulation is acceptable and well-grounded, the present chapter has shown that the opponent of relationalism should be doing much more work in order to make the argument from hallucination go through. The spreading step requires IND and IND has some plausibility only if we employ the notion of impersonal indistinguishability. Therefore, the argument must also show that some IIIHs exist. In this chapter, I have shown that there is no conclusive empirical evidence to support that claim. This needs to be introduced through stipulation and the validity of the stipulation must be motivated on theoretical grounds.

This confirms what I argued in Chapter 2. The argument from hallucination arises from a conflict between the way perception strikes us, on one hand, and a reasoning that emerges when we start to think problematically and philosophically about experiences, which allegedly compels us to give the same account to both perception and hallucination, on the other. As Valberg puts it, this is ‘problematic reasoning’, not a phenomenological datum (as transparency and the apparent relationality of experience are). As we will see in the next chapter, it is because of certain things we know about experiences, such as what brings them about and the sufficient conditions for their occurrence, that we suppose that a certain type of hallucination (IIIH) is possible. And only the possibility of IIIHs sets the spreading step in motion.

Not all proponents of the argument from hallucination, however, are as clear as Valberg about the theoretical nature of the problem. There is widespread temptation to think of the problem posed by hallucination as a phenomenological one. Some philosophers explicitly present the argument from hallucination as a phenomenological problem for relationalists.

Crane, for instance, claims that ‘[t]his problem is a phenomenological problem, in the broad sense of having to do with how things appear to us’ (Crane 2005b: 242: see also Crane 2005a). The problem, according to Crane, arises from the incompatibility of four assumptions that are equally phenomenological, in the sense that they are motivated by how experience strikes us. These assumptions are the mind-independence of the object of perception, what he calls the ‘openness of perception’—‘when a subject has a perceptual experience of an object, the nature of the experience is partly determined by the nature of the actual object currently being perceived’ (Crane 2005b: 231)—, the possibility of hallucination, and the ‘identity of subjective indistinguishables’—what I have called IND.

As we have seen, IND is certainly not grounded in how experience strikes us, and is far from obvious. Interestingly, when it comes to explaining and motivating the assumptions composing the antinomy, Crane indicates two possible motivations. One is what he calls a Cartesian view of the mental—recently defended by Farkas (2003). This view commits to some version of the infallibility of introspection, which I rejected in Chapter 3. The second motivation he considers is local supervenience, which I will

consider and reject in the next chapter. Irrespective of whether they are valid motivations or not, none of these is a phenomenological motivation, so it is puzzling and misleading that Crane calls the problem ‘phenomenological’. This feeds into confusion regarding the structure and the motivations underlying the argument from hallucination and it makes it look easier than it is to mount a seriously threatening argument.

Millar (2014) states even more clearly that the problem hallucinations pose for relationalism is phenomenological. He argues:

I suffer a hallucination, some object seems to be immediately present to me even though there is no such object—but perceptual acquaintance can’t explain the phenomenological directness of such an experience since one can only be perceptually acquainted with objects that exist. Thus, naïve realism would seem to have difficulty handling what we might call the phenomenological problem of perception (‘the phenomenological problem’ for short): how to develop a theory of perception that explains the phenomenological directness of perceptual experience given that illusory and hallucinatory experiences possess this phenomenological feature, *and assuming that such a theory should provide a unified explanation of why veridical, illusory and hallucinatory experiences possess this phenomenological feature.* (Millar 2014: 626, my italics)

Millar embeds among the allegedly obvious assumptions about how experience strikes us the assumption that a unified explanation should be provided for the phenomenal character of perceptions, illusions, and hallucinations. But there is no phenomenological datum that supports this assumption. If anything, in this chapter, we have seen that there are reasons to suppose that it is not even possible to achieve a unified account of hallucinations, let alone of perceptions, illusions, and hallucinations. There is no clear phenomenological characteristic that all hallucinations share and that they also share unequivocally with perceptions.

Moreover, even if we grant that at least some hallucinations have certain phenomenological features in common with perception, this is no reason to suppose that they are the same kind of mental occurrence whose phenomenal character is to be accounted for in the same way. The fact that such similarities exist is not something that the relationalist overlooks, as we saw in Chapter 2. However, relationalists reject the spreading step and idea that we should account for them in the same way. Invoking the similarities between perception and hallucination is not enough to compel one to accept the spreading step, just as invoking the similarity between a lemon and a lemon-shaped soap is not enough to prove that they are the same kind of thing. What is needed to secure the spreading step is IND. But, as we have seen, IND has some plausibility only if introspective indistinguishability is impersonal, and thus the spreading step can obtain only if one can show that hallucinations are indistinguishable in this sense. As we have seen in this chapter, no phenomenological evidence is available to support this claim.

Therefore, there is no valid phenomenological argument from hallucination against relationalism.

There is a tendency to evaluate relationalism on the basis of its ability to provide a valid and credible account of hallucination, and a presumption that it can't. However, relationalists can say whatever they want about the phenomenal character of hallucination, as long as they don't thereby establish that the spreading step is successful. They can say that the phenomenal character of hallucination is constituted by sense-data, qualia, intentional properties, relations to spatial locations, relations to universals, or anything else. It is also possible that different hallucinations deserve different accounts.

If the spreading step fails, no constraint is imposed on the relationalist as to what account can be given for hallucination—apart from the constraint that it cannot be a relationalist account. Of course, for any account of hallucination the relationalist proposes, it remains to be seen whether it is satisfactory, but this should be decided on its own merits and terms, not with respect to the consequences this could have for the relationalist account of genuine perceptions.⁴⁷ One may lament that relationalists so far haven't said enough about how we should account for hallucinations. But this would simply amount to expressing distaste for one's chosen research topic: complaining that the relationalist is interested

⁴⁷ In Chapter 6 we will see that constraints are indeed imposed on the account that relationalists should give of hallucinations. But these constraints do not arise from the acceptance of IND, but from a problem identified by Martin (2004; 2006), which he calls the screening-off problem.

in accounting for the phenomenal character of genuine perception, rather than that of hallucination.

Moreover, the discussion in this chapter shows that, even if it turns out that we have to stipulate the possibility of impersonally indistinguishable hallucinations, for those cases of hallucination that are not impersonally indistinguishable—as far as we can tell, *all* documented hallucinations—can still be accounted for in terms of whatever one deems fit. The success of the spreading step would impose constraints only on the account to be given of impersonally indistinguishable hallucinations. The failure to realise this is largely due to the lack, in the current debate, of serious discussion on the notion of indistinguishability and of what is required for IND to be plausible, among both proponents and opponents of relationalism. Failing to appreciate the distinctions between different notions of indistinguishability and that impersonal indistinguishability through introspection is required to (tentatively) motivate IND leads many to think that the simple failure to tell that one is not genuinely perceiving imposes constraints for the relationalist with respect to what she can plausibly say about the phenomenal character of hallucinations.

4.5 Conclusions

In this chapter I explored whether there is any evidence to conclude that at least some hallucinations are impersonally indistinguishable through introspection from perceptions. I argued that the structure of the relation of indistinguishability with which we are concerned holds between an

occurring hallucination and the whole class of perceptions: it amounts to saying that it is impossible to tell through introspection that one's current experience is not a member of the class of perceptions.

I then argued that, in virtue of the superficiality constraint that emerged from the discussion in the previous chapter, in order for a hallucination to be an IIIH, it must be the case that what distinguishes the hallucination from perceptions is not accessible through introspection. In order to show that this is the case, one should prove that there is no condition (altered state of mind, or cognitive deficits) that may impair one's introspective capacities and prevent one from noticing through introspection certain phenomenological differences that would be otherwise detectable. I argued that, when considering hallucinations, it is not possible to exclude the occurrence of any of these interfering conditions. Indeed, it is not clear what kind of empirical evidence could prove the absence of such conditions.

The proponent of the argument from hallucination can still maintain that it is *possible* to have hallucinations that are indistinguishable from perception in the impersonal sense, even though we do not have conclusive empirical evidence to say so. She would have to resort to introducing via stipulation the possibility of certain hallucinations that meet this criterion. In the next chapter I will explore whether this stipulation is legitimate, how it should be thought of, and, if we accept it, what follows from this.

Chapter 5: Philosopher's Hallucinations

We have seen that there is a sense in which premise (2) in the traditional argument from hallucination is obviously true: some hallucinations are such that one may fail to tell that they are not a case of genuine perception. However, if we understand premise (2) in this way, even if we accepted IND, the argument from hallucination would be a non-sequitur. For IND requires impersonal indistinguishability through introspection, while the sense in which (2) is trivially true only delivers mere failure to discriminate on a given occasion, or if it involves a modal notion of indistinguishability at all, it is one that is relative to the subject and her current situation, not the impersonal notion of indiscriminability required.

In order for the argument to succeed, thus, one needs to show that a modified version of premise (2) is true. This modified version reads as follows:

(2*) Some hallucinations are impersonally indistinguishable through introspection from genuine perception (IIHs).

But in the previous chapter we saw that there is no empirical evidence available to support this claim. What the proponent of the argument can do, then, is to show that IIIHs are metaphysically possible, although there is no evidence that any actual hallucination meets the requirements for being an IIIH. In other words, (2) should be further replaced with the following modal claim:

(2#) IIIHs are metaphysically possible.

In this chapter, I shall consider whether we have any grounds for making such a modal claim. I will consider three principles that might be seen as viable candidates for supporting (2#): experiential naturalism, local supervenience, and the same proximal cause, same immediate effect principle. I will argue that the first is insufficient to demonstrate the possibility of IIIHs; while the other two principles are problematic and beg the question against relationalism. In the process, I will also argue that two alternative versions of the argument from hallucination based on the latter two principles, which aim to reject relationalism without assuming IND, fail.

I will also introduce and dismiss a further argument against relationalism, aiming to show that, if we consider—as we should—the object of perception as causally related to the perception, it cannot also play a constitutive role in determining the nature of perception, since cause and effect are metaphysically independent.

However, at the end of the chapter I will present a non-conclusive reason for accepting that IIIHs are possible, based on (1) the metaphysical possibility of stimulating the brain in such a way that the subject is in the same neural state as in a genuine perception and (2) the fact that, in such a case, we lack reasons for assuming that any condition might interfere with one's capacities to introspect.

5.1 A Metaphysical Possibility

An obvious reaction to the failure to find empirical evidence for the actual existence of IIIHs is to say that there is no need to establish that any actual hallucination is an IIIH. It is enough to establish that IIIHs are metaphysically possible. This metaphysical possibility, it is widely believed, is as telling of the nature of experience as the actual existence of IIIHs.

Hence, one might argue, we can replace the existential claim in (2*), saying that that there are hallucinations that are impersonally indistinguishable through introspection from veridical perception, with a modal claim:

(2#) IIIHs are metaphysically possible.

Many recent versions of the argument, indeed, feature such a modal claim, although as we saw in the previous chapter, it is often suggested that the argument can run on the simple assumption that actual hallucinations are indistinguishable (in a subjective, not impersonal, way) from perception. Robinson (2013), for instance, claims that philosophers are generally not (or should not be) interested in hallucinations as they actually occur, but in metaphysically possible IIIHs, that he calls 'philosophers' hallucinations'.⁴⁸ The revised argument from hallucination, with the modified second premise, runs as follows:

⁴⁸ Like most philosophers, Robinson doesn't make the distinction between different notions of indistinguishability, so he only says that philosopher's hallucinations are subjectively indistinguishable from perception, not that they are impersonally indistinguishable.

1. Relationalism claims that the phenomenal character of conscious perceptual experiences is to be accounted for in terms of the obtaining of a psychological relation with the mind-independent objects with which one is perceptually presented;
- 2#. In some cases of hallucination, one has an experience which is introspectively indistinguishable from an experience one may have in a genuine perception;
3. In the hallucinatory case, the object one seems to perceive doesn't exist;
4. (from 3) In the hallucinatory case, the phenomenal character of one's experience is not accounted for in terms of the obtaining of a psychological relation with the mind-independent objects with which one is perceptually presented;
5. If two conscious experiences are introspectively indistinguishable, they have the same phenomenal character—IND;
6. (from 1, 4, 5) Relationalism cannot be true even in the genuine perceptual case.

We need to assess whether there is any reason to accept (2#).

The first thing to notice is that in the previous chapter I argued that when we assess whether an actual hallucination is indistinguishable from perception, we are assessing whether the hallucination in question can be (or could be, were the conditions more favourable) said through introspection not to be part of the class of perceptions. (2#) claims that a certain type of hallucination is possible. Since it states the possibility of a scenario, we

don't have an actual hallucination to be compared with the class of perceptions. But we have a counterfactual scenario and instructions to imagine such a scenario. The instructions go roughly like this: take whatever perception you are having now and suppose that the object seen is removed while you keep having an experience introspectively indistinguishable from the perception of the object.

In the previous chapter, we argued that, when it comes to ordinary hallucinations—however we understand the relation of indistinguishability—the relation holds between an occurring hallucination and a class of perceptions. Since the relation holds between an individual and a class, it is not symmetrical. The relation of indistinguishability holds between a hallucination and a class in the sense that it is not possible for the person to tell that her hallucinatory experience is not a member of the class of genuine perceptions. This relation is asymmetrical because the reverse cannot be true: it doesn't make sense for a class to be indistinguishable from an individual.

On the contrary, in the case of philosopher's hallucinations, the relation of indiscriminability holds between a genuine perception chosen at will and a hypothetical hallucination. This is a one-to-one relation involving particulars, and we saw in the previous chapter that this kind of relation is, in its typical form, symmetrical. What might prevent the relation from being symmetrical in any meaningful sense, however, is the fact that the hallucination is a hypothetical construct and the perception provides the model for imagining such a hallucination.

While it makes sense to claim that a counterfactual hallucinatory situation is indistinguishable from a genuine perception, it is not so clear what it would mean for an actual situation to be said to be indistinguishable from a counterfactual one, all the more so if the counterfactual situation is designed to be indistinguishable from the target perception. If it makes sense at all to say that the perception is indistinguishable from the hallucination, it can only be in a trivial, non informative way: the perception is trivially indistinguishable from the counterfactual hallucination, since it was part of the instructions to imagine the latter such that is indistinguishable from the perception. If there isn't a structural asymmetry, there is at least an important asymmetry in the order of explanation of the direction of the relation of indiscriminability. Being indistinguishable from a definite perception was part of the instructions given for constructing the counterfactual hallucination. It is an open question whether or not these instructions can be followed. Whether or not they can successfully be followed—i.e. whether or not it is possible to have a hallucination that is indistinguishable from the target perception—will tell us something about the hallucination, but it is unclear whether it will tell us anything about the perception. Even if we can grammatically say that the perception is indistinguishable from the hypothetical hallucination (which has been designed to be indistinguishable from the perception), this doesn't say anything informative about the perception itself, if not that it can be used as a model for conjuring up a counterfactual situation.

The asymmetry of this relation will become relevant in the following chapter, in which I will defend a negative view of hallucination as the only way to respond to the superficiality constraint and to other challenges arising for relationalist accounts, where this negative account of hallucination would be threatened if the relation of indistinguishability were symmetrical. But it is also relevant for the purposes of the current chapter, i.e. for deciding whether we should accept (2#). For it is important to understand the nature of the relation between a hallucination and the perception from which it is meant to be impersonally indistinguishable.

Many philosophers hold (2#) with extreme confidence. Most often, the possibility of IIIHs, often called 'perfect hallucinations' or 'corresponding hallucinations' is simply stated, as a fact that doesn't require further questioning. Farkas notices that many 'introductions to the problem of perceptual knowledge, or epistemology, or scepticism, or the nature of perception' contain something like the following paragraph:

Suppose I now see a teacup in front of me. Wouldn't it be possible that everything seems the same, and yet the teacup I take myself to be perceiving is not there? Wouldn't it be possible to have a hallucination which is subjectively indistinguishable from my present experience? (Farkas 2006: 207)

It is clear that these are rhetorical questions, as becomes manifest in Farkas' next sentence, where she abandons the questioning and switches to

talking about 'the *fact* that hallucinations are possible'.⁴⁹ It should be noticed that in this passage, as in those I will quote later and in most discussions on the argument from hallucination, there is no mention of the hallucination being impersonally indistinguishable from perception, because of a lack of awareness of the relevance to the debate of different notions of indistinguishability. However, the scenario that Farkas and others introduce as possible here is compatible with the conditions imposed by the impersonal notion of indistinguishability that I discussed in Chapter 3.

Similarly, in their introduction to their edited volume on disjunctivism, Byrne and Logue start by proposing the following thought experiment:

Imagine that you are looking at an ordinary lemon in good light. Your vision is good: you see the lemon, and it looks yellow and ovoid. Now suppose that, unbeknownst to you, some minor deity removes the lemon, while preserving its proximal neural effects. Your brain is in the same local physical states as it was in when the lemon was there: the neurons in your visual cortex, for instance, are firing in the same pattern. After the removal, you do not see the lemon, because the lemon is not around to be

⁴⁹ By 'hallucinations' she does not mean any kind of hallucination, but 'subjectively indistinguishable hallucinations', as she says in the preceding passage. However, this is an indication of the widespread tendency that I signalled earlier among proponents of the argument from hallucination to switch from presenting the argument as relying simply on the trivial *fact* that we have hallucinations to the possibility of a certain kind of hallucinations that is supposed to pose special problems to relationalists. This adds to the confusion due to the lack of discussion of different notions of indistinguishability.

seen. Yet—we can all grant—you notice nothing amiss. Questioned after the removal, you claim that you have been looking at a lemon for the last few minutes. (Byrne and Logue 2008: vii)

For Smith (2002) there is indeed no room for doubt on this matter:

It is surely not open to serious question that [...] if the activity of your optic nerve when you are genuinely perceiving something green is precisely replicated artificially, you will, other things about you being normal, seem to see something green *in a genuinely sensory manner*. (Smith 2002: 203)

Byrne and Logue, as well as Smith, mention the conditions in which we have reason to believe that an IIIH could be produced: i.e. upon activation of the subject's neurons in a way that matches the neural state she is in when having a genuine perception. I will call this situation a 'neurally matching situation'. Thus, we can say that (2#) relies on the following assumption:

- (a) It is possible to have a neurally matching situation: i.e. it is possible to reproduce the neural state one is in when perceiving, in the absence of the object of perception.

Robinson (1994), however, notices that this assumption alone is not sufficient to motivate (2#). A further assumption is needed. In his words, this is the idea that:

(b) 'our mental life is strongly tied to what happens in the brain'

(Robinson 1994: 152).

The thought is that, in virtue of the relation between psychological occurrence and neural state stated in (b), if it is possible to stimulate the brain as stated in (a), the result of (a) would be an IIIH. For Robinson, both these claims are uncontroversial, as is their ability to grant (2#). But we should pause and ask whether or not this is true.

One could object that (a) is far from obvious. Our knowledge of the minute functioning of the brain is as yet far too unsophisticated to allow us to freeze the brain in a state specific to a given time. As far as we know, there might be insurmountable limitations making it impossible to ever do so, no matter how far the evolution of neuroscience and technology goes.

However, Robinson points out that (a) should be understood as a claim about what is compatible with known natural laws. It doesn't matter whether freezing the relevant brain state is or will ever be practicable. What matters is the fact that such a neurally matching situation is physically possible; that is to say possible without breaching the causal laws.⁵⁰

Such a fixing of brain state is conceivable, not only in the sense that it is logically possible; it is empirically possible, in the sense that we may conceive of ways of doing it which involved no breach of natural law. This is compatible with its not being

⁵⁰ By noting that this scenario is physically possible, I do not commit to the claim, weaker than metaphysical possibility, that the scenario is possible *only within* the current physical laws, but that it is possible without breaching physical laws, which is compatible with it being possible in a world governed by different physical laws.

practicable and even with our having reasons for thinking that it might never be practical. It might even happen by a remarkably improbable coincidence (given quantum theory) or there might be a drug which held constant the relevant areas of the brain.
(Robinson 1994: 152)

Thus understood, there doesn't seem to be any reason to reject (a): the possibility of neural replication, although potentially impracticable, doesn't seem to breach natural laws.

5.2 Experiential Naturalism

Whether the possibility of a neurally matching situation entails the possibility of IIIHs as stated in (2#) depends on how we understand the nature of the 'strong tie' stated in (b). Robinson talks vaguely about a 'strong tie' between mental life and the brain. What is this tie supposed to be? Is it causation? Is it mere correlation? If there is a causal law, what sort of causal law links them? Talk of a strong connection, as noticed by Bergson, is as obvious as uninformative:

That there is a close connection between a state of consciousness and the brain we do not dispute. But there is also a close connection between a coat and the nail on which it hangs, for, if the nail is pulled out, the coat falls to the ground.
(Bergson 1988: 12)

We cannot hope in this context to define the exact nature of the relation between mental occurrence and neural states. Doing so would require taking a stance in the debate over the controversial mind–body problem, which is far too deep and broad to be satisfactorily tackled here. However, the acceptance of (2#) should not require committing to any controversial claim. Robinson must conceive the acceptance of (b) as being independent from commitments on the mind–body problem, since there is notoriously very little agreement on the latter, while Robinson presents (b) as widely agreed upon.

One uncontroversial idea that might motivate (b) is what Martin (2006) calls ‘experiential naturalism’. This is the idea that ‘our sense experiences are themselves part of the natural causal order, subject to broadly physical and psychological causes’(Martin 2006: 357). Martin explains what he takes to be involved in being part of the natural causal order by underlying what it would take to reject experiential naturalism:

Where there is a genuine causation, then there is a pattern of causes and effects that has an implicit generality. If a given causal condition has produced a certain kind of effect in one context [and another kind of effect in another], then there will be some general condition which differs between the two. It seems that to deny this would be either to claim that in the case of psycho-physical causation there need be no such determinacy of cause to effect, or to embrace in general a radical singularism about causation according to which the fact that particular

causes were related to the effects that they had implied nothing at all about the general patterns of causation. (Martin 2004: 288)

Experiential naturalism affirms that the generality and regularity of causal processes applies to experiential phenomena. As Martin notices, this is a 'regulative assumption of both empirical work on sense experience and philosophical discussion of it' (Martin 2006: 360). This principle is general enough to be accepted by proponents of all sorts of metaphysical views and empirical hypotheses about the relation between mental occurrences and brain states.

But is this reading of (b) enough to support (2#)? Valberg seems to think so. Valberg doesn't talk of 'experiential naturalism', but rather of what he calls a 'causal picture of experience', which consists in the idea expressed by 'experiential naturalism' complemented with some more or less rudimentary notion of the specific causal laws concerning 'the basic facts about the transmission and reflection of light, the nature of the eye, and the process in the nervous system and brain which results or culminates in visual experience' (Valberg 1992b: 24). Valberg claims that the idea that the same brain activity could continue to occur even in the absence of the external stimulus is 'implicit in our everyday knowledge about the causal dependence of experience on objects and processes in the world (implicit, i.e., in the causal picture of experience)' (Valberg 1992b: 25). We have already granted that we can suppose that it is possible to maintain a brain state unvaried in the absence of the object experienced. But Valberg says something more than that: he claims that 'if the activity in my brain were to

continue on as it is, my experience would also continue on as it is' (Valberg 1992b: 25) and 'it would be in some sense be true that how things are in my experience would remain the same' (Valberg 1992b: 28). This further step is more problematic. It is far from obvious that this would suffice to maintain an experience that is indistinguishable from the original perceptual experience—let alone the far stronger claim that it would be the very same experience. So we need to consider more carefully what 'experiential naturalism' entails. Here Valberg seems to conflate the claim that a given neural state is a necessary condition for a certain mental occurrence with the claim that the brain state is sufficient for the mental occurrence.⁵¹

The principle of 'experiential naturalism' states that if we have two situations in which the same cause produces a certain kind of effect in one situation and another kind of effect in the other situation, this must be due to some difference in the background conditions or antecedents of the causes in the two situations. For instance, the striking of a match generally generates fire; but this effect doesn't occur if the match is struck in an atmosphere that doesn't contain oxygen, or if the match is immediately immersed in water after the striking. In the case of our neurally matching situation, the principle predicts that the effect of the brain activity in the hallucinatory case can differ from the effect of the brain activation in the

⁵¹ By committing to this sufficiency claim, as a matter of fact, then, Valberg commits to the something along the lines of the 'same proximate cause, same immediate effect' principle that I will discuss later in this chapter. However, I have opted to discuss his proposal separately, as relying on a more general commitment to the simple idea that there is a some causal link between neural state and conscious experience, because he seems to believe that the possibility of IIIHs derives solely from the general and pre-theoretical idea that there is some sort of psycho-physical causal law, without the need to commit to any particular view about causality.

perceptual case only if different contextual conditions are in place. But it is no mystery that there are indeed different conditions in the perceptual and hallucinatory case: the presence or absence of the external object. Since the brain states in perception and in the neurally matching situation are identical, it is possible that the resulting effects will differ in all sorts of ways. As far as what experiential naturalism dictates, it is possible that the absence of the object would act like the absence of oxygen in the match scenario: in its absence there might not be an experience with a phenomenal character, let alone one indistinguishable from that of a genuine perception.

This is not to deny the intuition that neural replication would result in an IIIH. For all we know so far, it might or might not be true, but certainly it doesn't follow from experiential naturalism. Sure, it would be disconcerting and disappointing to be forced to reject the powerful intuition that neural replication, if possible, would give rise to IIIHs. But precisely because this intuition is so forceful and widespread, it is important to understand what exactly underwrites it. Unfortunately, assumptions about the link between mind and brain underwriting (2#) are rarely discussed and often conflated.

In what follows, I will identify two assumptions that are often more or less explicitly offered to motivate (2#) and are often used interchangeably, although they have different implications and demand different commitments. These are the principle of local supervenience and the 'same proximal cause, same immediate effect' principle. Both principles are often used to mount independent arguments against relationalism—arguments

that are allegedly more compelling than the traditional one (including in the revised form we are considering here), since they do not require the acceptance of IND, which (as we have seen) is the most speculative and controversial premise in the argument. I will argue that neither local supervenience nor the 'same proximal cause, same immediate effect' principle are well motivated. Therefore, they cannot be used to motivate (2#) in the revised argument from hallucination, nor they can feature in independent arguments against relationalism.

5.3 Local Supervenience

One tempting way to think of the 'strong tie' between conscious experience and brain state is in terms of supervenience. It is often claimed that the mental supervenes on the brain, in the sense that no two mental occurrences can differ qualitatively without there being an underlying difference in the properties of the brain state. Typically, supervenience claims are used in philosophy of mind as part of attempts to define the commitments of physicalism and, although different supervenience claims might entail stronger or weaker versions of physicalism, even the weakest version of the supervenience thesis entails a commitment to at least a very minimal physicalism. Thus, supervenience, at least in some version, can be accepted by any philosopher with physicalist inclinations, as it remains neutral on all sorts of questions about the nature of mental phenomena. Supervenience states a connection of sufficiency between brain and phenomenal character, but it doesn't specify the nature of such a relation, or what explains it, and is

therefore compatible with multiple ways of construing the relation between the mind and the brain, including ways that allow other factors to influence the relation of dependency between the brain state and phenomenal character of the experience.

Despite the pervasive success of physicalism among current philosophers, there are still some who resist it—for instance Robinson, who is committed to a version of substance dualism.⁵² Therefore, for Robinson it cannot be a supervenience claim that underwrites (2#). However, this seems to be exactly what many other philosophers who equally take (2#) to be undeniable have in mind and is thus worth exploring.

Although supervenience is generally taken to be the minimal commitment of physicalism, the literature is replete with ‘an unlovely proliferation’—as Lewis puts it (Lewis 1986: 14)—of brands of supervenience claims: general and local, weak and strong supervenience claims being the most popular versions. These are not equally minimal and non-committal and it is a matter of debate which brand of supervenience one should accept. Also, supervenience claims often leave unqualified what properties of mental occurrences supervene on the brain state, probably under the assumption that *all* properties of a mental occurrence do so.

But it is important to make distinctions. For instance, an externalist about mental content might maintain that the mental content of an

⁵² Some maintain that some versions of property dualism are compatible with physicalism. For instance, primitivist emergentists claim that the mind supervenes on the body insofar as the latter is a precondition for emergence. Whether these are coherent views and whether they really fulfil the brief of both physicalism and dualism is a question I shall not address in this thesis.

experience doesn't supervene on the subject's brain state, yet its phenomenal character does (Block (1996) and Burge (1979, 2003) are prominent examples of this view). Most representationalists indeed see the relation between content and phenomenal character as a one-direction relation: the phenomenal character supervenes on the content, but not vice versa. Since externalism about content is widely accepted, while externalism about phenomenal character is not and strikes many philosophers as implausible,⁵³ here I will consider supervenience about phenomenal character alone. This is also because we are interested in supervenience as a motivation for (2#), which is a claim about introspectively indistinguishable experiences, i.e. about experiences that are indistinguishable with respect to their phenomenal character.

In addition, in order to be relevant to (2#), the supervenience claim needs to be local. Global supervenience claims that the total state of any possible world could not have been any different mentally without differing physically. But the situations in perception and hallucination differ physically from one another in that in one case, but not in the other, there is an object. As far as we know, the presence of the object could be a necessary condition for the occurrence of an experience with a certain phenomenal character. Global supervenience would allow for the neurally

⁵³ This is due to the understanding of phenomenal characters that many of these representationalists have in terms of intrinsic properties, or qualia, an understanding of phenomenal character that the relationalist and any other externalist about phenomenal character is committed to reject. Whether such an understanding of phenomenal character is indeed compatible with externalism about content is another issue, which goes beyond the scope of this thesis. See Hellie (2010) for compelling arguments to the effect that externalism about content is incompatible with internalism about phenomenal character.

matching situation to produce a hallucination with a phenomenal character introspectively different from that of perception, or even for the neurally matching situation not to produce any conscious experience at all, or to produce an experience with no phenomenal character at all, due to the context dependence of mental occurrences allowed by global supervenience. Thus, the supervenience thesis needed for the argument from hallucination is a claim of local supervenience about phenomenal character. According to local supervenience, the obtaining of the brain state is sufficient for the occurrence of an experience with a specific phenomenal character.

Local supervenience, together with (a), clearly entails (2#). If it is possible to reproduce the brain state of a genuine perception in the absence of the object perceived, then, in virtue of local supervenience, the brain activation will be sufficient to produce an experience with the same phenomenal character as the genuine perception. The two experiences would then of course be indistinguishable.

However, this form of local supervenience is at least controversial and many have pointed to numerous potential counterexamples: neural replication alone would be enough for sameness of phenomenal characters even in all sorts of cases where we might find it problematic.⁵⁴ Local supervenience would entail that a disembodied brain in a vat whose neurons are stimulated appropriately would have a hallucination indistinguishable from the target perception. But proponents of embodied cognition would

⁵⁴ The Twin Earth thought experiments proposed by Burge (1979) and Putnam (1975) are cases of neurophysiologically indiscernible people whose thoughts have different contents.

reject this. They maintain that the beyond-the-brain body plays a significant causal role in many features of cognition (see Noë and Thompson 2004; Noë and Thompson 2004; O'Regan and Noë 2001; Varela, Thompson, and Rosch 1991; Hurley 1998). Similarly, neural replication in a subject deprived of the cognitive capacities allowing them to recognize objects would be enough for the subject to have a hallucination indistinguishable from the target perception. This would be rejected by those who maintain that perceptual capacities require the deployment of conceptual capacities or other cognitive capacities that must be already in place (see for instance McDowell 1994 and Brewer 1999). In order not to beg the question against these theoretical options, whose acceptance goes beyond the acceptance of relationalism and can be held by representationalists too, one needs to qualify local supervenience with two clauses:

(LS): Neural replication is sufficient for production of experience with phenomenal character in brains of subjects who (1) are in a functioning body, (2) are capable of in the relevant cognitive capacities required for experiencing.

Finally, as Fish notices, empirical data, although not conclusive, speak against the possibility of inducing hallucinations in a person who was born blind. Although he couldn't find any study of direct electrical stimulation of the brain in blind subjects, he found studies that used transcranial magnetic stimulation (TMS) (Gothe et al. 2002). The study resulted in a marked difference between blind and non-blind patients in the ability of TMS to

produce simple visual phosphenes. It must be premised that even the systematic production of phosphenes in blind people would not prove that supervenience should be accepted (as phosphenes are not the kind of experience that is IIIH). However, the study showed that the ability to produce phosphenes decreased as the severity of the blindness and the length of time of being blind increased, and TMS could not induce any phosphenes in the congenitally blind. Despite not being conclusive evidence against local supervenience, this suggests that previous ability to perceive might also be required in order for one to hallucinate, making neural activation not sufficient.

Notice that, even with these qualifications, local supervenience, if true, would be enough for securing (2#). In the thought experiment, in fact, there is no assumption about the subject's being blind, or being a disembodied brain, or being deprived of the relevant cognitive capacities. The thought experiment asks us to imagine that the very same person, that is, a person with a functioning body and normal cognitive capacities, who genuinely saw an apple a moment before, can be induced through brain replication to have an indistinguishable experience of the apple when the apple is removed from the environment.

Most importantly, local supervenience provides much more than is needed to secure (2#) in the traditional argument from hallucination. Accepting local supervenience, even thus qualified, provides one with a direct argument against relationalism. If local supervenience is true, relationalism has to be false: relationalism claims that the phenomenal

character of perception is partly determined by the objects and events perceived, but local supervenience claims that the phenomenal character supervenes on the brain state, so the presence or absence of the objects of perception cannot make any difference in the phenomenal character of one's experience. Therefore, if local supervenience is true, there will be no need to accept IND for relationalism to be in trouble.

According to Broad, local supervenience would provide a 'knock-down argument' against relationalism:

Suppose it could be shown that the occurrence of a certain disturbance in a certain part of a person's brain at a certain time is the immediate sufficient condition of his then having an experience which he would naturally describe as seeing or hearing or feeling a foreign object of a certain kind in a certain place.⁵⁵ Then it would follow at once that the actual presence of such an object in that place at that time cannot be a necessary condition of the occurrence of the experience. From this it would follow at once that the experience cannot be, as it appears to be to the person who has it, a prehension of the object in question. In that case the utmost that could be alleged is that the presence of such an object in the place in question at a

⁵⁵ 'Immediate' here suggests that Broad might have in mind the 'same proximate cause, same immediate effect' principle, rather than the local supervenience principle. However, it is difficult to decide which principle exactly he is suggesting (Fish 2009, for instance, assumes that Broad is talking about supervenience). This is indicative of a widespread tendency in this field. Neither is it surprising, as the two principles have the same result with respect to the argument from hallucination (see Section 5.5 for further discussion).

somewhat earlier date is an indispensable causal ancestor of that disturbance in the brain which is the immediate sufficient condition of the occurrence of the experience. (Broad 1952: 10)

We might not be willing to concede to Broad that the above would be a knock-down argument. Hinton would resist it, claiming that it is possible to disjunctively characterize the event supervening on the brain state. Discussing a related argument that will occupy us in the next section, the argument from the 'same cause, same effect principle', he claims:

The impulse reaches certain specified structure, and then—what? My continuation was,—‘and then, one perceives a flash of light or has the illusion of doing so, as the case may be, according to the nature of the initial stimulus. (Hinton 1973: 75)

We will see in the next chapter how the relationalist could provide a disjunctive account of the event supervening on the mental state. For our current purposes, it is worth noticing that even Broad acknowledges that it would be precipitate to accept ‘the sweeping assertion that the occurrence of a certain kind of disturbance in a certain part of a person's brain is the immediate sufficient condition, as distinct from an immediate necessary condition, of his having an experience’ (Broad 1952: 10). Although tempting, local supervenience is far from granted. Indeed, as Fish points out, ‘many of the reasons given for thinking that local supervenience is true are not as compelling as they are often assumed to be’ (Fish 2009: 42).

Fish thinks that many accept local supervenience as an inference to the best explanation for the fact that ‘hallucinating subjects take themselves to be having veridical perceptual experiences, and this needs explaining’ (Fish 2009: 118). The reasoning would go as follows:

The fact that hallucinating subjects take themselves to be having perceptual experiences suggests that hallucinations are mental events that have a phenomenal character that is the same as, or at least similar to, the phenomenal character of veridical perception. Yet given what we know about the etiology of hallucinations, the only thing that seems to be necessary for a hallucination to occur is brain activity of the right kind. A plausible inference to the best explanation would conclude that this neural activity suffices for the existence of the hallucination, together with the phenomenal character. If so, then the hallucinatory phenomenal character is a property whose instantiation is assured by the neural activity alone. [...] If neural activity alone suffices for a hallucination, to instantiate a phenomenal character of a certain kind, then if that neural activity also occurred in a case of veridical perception, it would likewise suffice for the instantiation of a phenomenal character of the same kind. (Fish 2009: 118)

The main problem with this line of reasoning is that there is no evidence that hallucination involves exactly the same brain activity as

genuine perceptions. On the contrary, the etiology of hallucination varies a lot across types of hallucinations and pathologies or conditions in which they occur and often involves the activation of brain regions that are not activated during genuine perception.⁵⁶ Also, as we have extensively seen in the previous chapter, there is a crucial difference between IIIHs and hallucinations that one simply happens to be unable to tell apart from perception. So this argument trades on mischaracterizing what happens when one hallucinates, both on the neuronal and the phenomenological level. As it stands, the fact that some people hallucinate can at best prove that the phenomenal character of hallucination supervenes on the activation

⁵⁶ In schizophrenic patients, there is evidence that a too high level of dopamine in the limbic system (amygdala, hippocampus, ventral striatum, cingulate) plays an important role in the occurrence of hallucinations and other psychotic symptoms (Laruelle, Kegeles, and Abi-Dargham 2003). The effect of dopaminomimetic agents on the limbic system also seem to be responsible for hallucinogenic episodes in patients with Parkinson's disease (Wolters and Berendse 2001), especially when administered L-dopa, which has a strong dopaminergic effect (Onofrij et al. 2006). LSD and other psychoactive drugs boost serotonin transmission and induce an increased excitability of the visual cortex, which might lead to hallucinations (Oliveri and Calvo 2003). Hallucinations involved in epileptic seizures are connected with stimulation of the amygdala or hippocampus (W. Penfield 1955)(Penfield 1955; Vignal et al. 2007). Studies have found that genetic factors are connected with an enhanced sensitivity of neurotransmitters to dopamine (Malhotra et al. 1998). Moreover, studies have highlighted abnormalities in the volume of brain structures of schizophrenic patients, including a reduction in volume of the superior temporal gyrus and increased volume of ventricles (Weiss and Heckers 1999). Studies using magnetic resonance on schizophrenic patients suggest that, during inner speech, the alterations of white matter fibre tracts responsible for the connectivity between frontal and parietotemporal speech-related areas lead to abnormal coactivation in regions related to the acoustic processing of external stimuli. This abnormal activation may account for the patients' inability to distinguish self-generated thoughts from external stimulation (Hubl et al. 2004). In a study involving Charles Bonnet patients, Ffytche et al. (1998) found brain activation of content-specific areas of the visual cortex (in particular in the ventral occipital lobe) and noticed that the content of the hallucination reflected the functional specialization of the specific region activated. These few examples indicate the variety and complexity of factors (bio-chemical, genetic, of brain structure, and of brain areas activated) involved in the aetiology of different types of hallucinations occurring in different conditions.

of some brain states. It doesn't show that the very same phenomenal character that one could have when genuinely perceiving can occur also in hallucination, where it seems that brain activity is sufficient for the occurrence of an experience with phenomenal character.

But even if it were the case that some hallucinations were introspectively indistinguishable from perception in an impersonal way, the argument would be begging the question against relationalism. Even if the phenomenal character of hallucination supervened on brain activity, it would still be possible for the relationalist to deny that the same is true for the phenomenal character of perception. Nothing about the etiology of hallucination can show that, in the case of perception, the existence of the object experienced is not necessary for the occurrence of the phenomenal character. This is precisely the debate at hand here: the relationalist contends that, even when impersonally indistinguishable through introspection, a perception and hallucination will have different phenomenal character, because the phenomenal character of a genuine perception is partly constituted by the presence of the object. The proponent of local supervenience cannot justify supervenience on the basis of a principled rejection of this possibility.

Since real hallucinations cannot support local supervenience, proponents often turn to imaginary cases of hallucination, where one is in exactly the same brain state as when one has a genuine perception and has an experience that is indistinguishable from perception (or which has the same phenomenal character as perception, proponents of local

supervenience would say). But as will be clear from the discussion so far, we are considering local supervenience as a ground for the possibility of (2#), which in the thought experiment allows for the move from neurally matching situations to IIIHs. Trying to justify local supervenience with such a thought experiment involving neural replication and IIIHs would thus amount to committing to blatantly circular reasoning, as Fish points out:

To get from neural replication to sameness of experience [in the thought experiment...] local supervenience is just assumed. So, as it stands, such a thought experiment does not underwrite local supervenience, but rather takes it for granted. (121)

The fact that many philosophers take the thought experiment to involve cases in which, as Fish puts it, we 'get from neural replication to sameness of experience', is indeed an indicator of the fact that local supervenience underwrites this kind of counterfactual reasoning. Local supervenience delivers more than would be needed in order to secure (2#). Conversely, if the mere possibility of IIIHs brought about by neural replication involved in the thought experiment was motivated by reasons independent of local supervenience (we will consider another candidate for grounding this possibility further below), it would still not be possible to ground the acceptance of local supervenience in such a possibility. For local supervenience would predict that that neural replication would produce matching neural state has the very same phenomenal character as the

corresponding genuine perception: one would simply have that they are indistinguishable.

5.4 Is there Empirical Support for Local Supervenience?

Given that neither reflection on ordinary hallucinations reported in the psychopathological literature nor thought experiments can be used to establish local supervenience (even in its qualified version), lest we beg the question against relationalism and incur a charge of circularity, one might try to turn to empirical data that allegedly provide evidence for the claim that hallucinations that causally match an experience are also impersonally indistinguishable from it. Fish notices that there is widespread conviction that local supervenience is an empirical hypothesis supported by data about brain stimulation experiments (see Fish 2009: 123). Although this strategy often remains implicit, Lowe (1992: 80), for instance, explicitly mentions Penfield's studies on direct brain stimulation to support the supervenience claim. Similarly, Robinson mentions Penfield's studies as classical 'empirical evidence' 'for the sufficiency of brain stimulation for experience' (Robinson 1994: 152 and footnote).⁵⁷

Although dated, Penfield's work remains to date the most extensive corpus of data on brain stimulation generating hallucinatory experiences. This might be due to the fact that direct brain stimulation is practiced for therapeutic purposes, and not in the context of theoretical research about the

⁵⁷ As noted earlier, Robinson wouldn't be happy to accept supervenience, since he rejects physicalism. However, he believes that brain stimulation suffices for experience in virtue of the 'same proximal cause, same immediate effect' principle.

nature of experience or to assess hypotheses about the nature of the relation between neural states and psychological event.⁵⁸ Recent research on the effects of direct brain stimulation are part of an emerging experimental treatment of neuropsychiatric conditions such as depression, anxiety, Parkinson's disease, and chronic pain. There are so few the studies monitoring whether direct brain stimulation induces experiences and a small fraction of these give enough details about the phenomenology of the experiences induced to support any speculation whatsoever about the correlation between mental occurrences and the brain. More recent studies are Diederich, Alesch, and Goetz (2000), which reports the occurrence of a hallucination in one patient with Parkinson's disease upon deep stimulation of the subthalamic nucleus, and Halgren (1982), which reports 'vivid formed dream- or memory-like hallucinations, or intense feelings of familiarity induced by direct electrical stimulation of the limbic system'. Both studies tested only one subject and the first-person descriptions of the reported effects of stimulation are meagre.⁵⁹

Compared to these sparse results, Penfield's final summary of his research on epileptic seizures spanning from the late 1930s to the 1960s (Wilder Penfield and Perot 1963) is impressive. It collects the data of 520 epileptic patients whose cortex of the temporal lobe was electrically

⁵⁸ As noted by Penfield in a discussion with Libet during a study week whose contributions are collected in Eccles (1966) 'experiment is never a primary purpose [of stimulation of the human brain]: the primary purpose is therapeutic (Eccles, ed. 1966:176).

⁵⁹ Taylor, Walsh, and Eimer (2010) also induced phosphenes by applying transcranial magnetic stimulation, but we have seen in the previous chapter that phosphenes are not indistinguishable from genuine perceptions.

stimulated and contains detailed first-person descriptions of the experience undergone. These data were collected while conducting surgeries to remove specific parts of the temporal lobe areas considered responsible for epilepsy. During the operation, Penfield discovered that patients under local anaesthesia could respond to his questions even during surgery.⁶⁰ Among 520 patients, 40 patients (7.7%) reported complex auditory or visual, or even multimodal (auditory and visual) experiences.

None of the experiences described by the patients can possibly count as an IIIH. The direct brain stimulation induced an epileptic seizure and in some cases the seizure was accompanied by hallucinatory episodes. Hallucinations in both spontaneous and induced epileptic seizures have phenomenal traits very different from genuine perception and could in principle be easily distinguished from them, as did the experiences reported by Penfield's patients. Although some patients described their hallucinations as a 'true experience' (Penfield and Perot 1963: 643), most of them described them as being 'like a dream' (id: 680), or like a 'memory'. The experiences described in Penfield's study indeed had traits typical of hallucinations in epileptic seizures, such as a sense of déjà vécu and complex visual hallucination accompanied by a feeling of 'strangeness' (Aleman and Larøi 2008). Penfield describes these experiences as 'recall of past experiences' due to the 'activa[tion] of a neuronal record'.

⁶⁰ The fact that studies of this extent haven't been reproduced can be explained by progress in surgical treatment, which has become less invasive and tentative, and pharmacological treatment, which has replaced surgery for many patients (Shorvon, Perucca, and Jr 2009).

Although Penfield sometimes seems to stress the genuine experiential nature of the episodes undergone by his patients,⁶¹ he never conceives of these experiences as akin to genuine perception. On the contrary he calls them 'a new order' of phenomena (Penfield and Paine 1955, reported by Aleman and Laroi 2008). (Jackson 1931) calls these states 'dreamy states'. This 'dreamy nature' of the experiences is confirmed by (Vignal et al. 2007), who revisited Penfield's studies.

Other phenomenal traits distinguish Penfield's patients' experiences from genuine experience. For instance, one patient explains that '[he] see[s] the people in this world and the people in that world too at the same time' (Penfield 1955: 679). This kind of superposition of the hallucinatory scene over the perceived one, accompanied by an awareness of their distinctive nature, is precisely one of the phenomenological traits that allows one to tell through introspection that one is hallucinating rather than genuinely perceiving. Moreover, Penfield notes that, no matter how vivid and rich the experience was, none of the patients took their experience for a genuine perception, even for a few seconds. Clearly, thus, Penfield's direct brain stimulation cannot be interpreted as inducing IIIHs.

⁶¹ As noted by Hinton (1973), the ordinary notion of 'experience' is very general: 'a transparent, colourless, or empty verb, akin to "undergo" or "do"' (Hinton 1973: 13). It can be found in locutions such as 'the experience of swimming the Bosphorus' (6), where there isn't any specific reference to conscious sensory experience. But even if we understand 'experience' in a narrower way, encompassing 'states of consciousness' (Hinton 1973: 22), 'experience' doesn't only apply to perceptual experience, but to all sorts of conscious mental occurrences such as pain and bodily experience, but also imagining and visualizing, dreaming, recollecting, and so forth. So we shouldn't read too much into Penfield's and his patients' use of the expression 'experience'. Fish, it seems to me, does read too much into it when he sees Penfield as committed to the claim that these experiences are genuine veridical and auditory experiences in the sense that they 'have perceptual-like phenomenal character' (Fish 2009: 130).

Penfield's studies show that neural activation *of some sort* (we have no evidence that the neurons activated in Penfield's studies are exactly the same as those activated in normal perception, and there is some evidence to the contrary) is enough to produce an experience with a phenomenal character of *some sort* (there is evidence that the experiences produced by Penfield's direct brain stimulation are phenomenologically very different from occurrences of genuine perception).⁶² But obviously this falls short of showing that conscious hallucinations brought about by the very same brain state as a perception is possible, and even less that this experience will be impersonally indistinguishable from the corresponding perception.

Given that considerations on hallucinations—either naturally occurring, induced by brain stimulation, or introduced as counterfactual hypotheses in thought experiments—cannot support local supervenience,

⁶² My diagnosis of how we should interpret the results of Penfield's study differ from that offered by Fish. I maintain that Penfield's studies only show that direct brain activation is sufficient to produce an experience with a phenomenal character, although it is not clear that either the same brain area involved in perception are activated in Penfield's studies, or that the phenomenal character produced upon such activation is indistinguishable from that of genuine perception. Fish, instead, claims: 'These results show that, in these subjects, brain activity alone is sufficient for the occurrence of a visual experience. But this is not to accept (yet) that the brain stimulations results support the claim that brain activity is sufficient for experiences with phenomenal character' (Fish 2009: 124). With reservations, Fish maintains that it might be granted that Penfield's studies show that the hallucinations induced by direct brain stimulation can be indistinguishable from a perception, but not that they have a phenomenal character. Fish's divergent interpretation has two explanations. On the one hand, he fails to distinguish between different notions of indistinguishability, and therefore is prone to take experiences that one merely fails to distinguish from perception as indistinguishable in the relevant sense. On the other hand, Fish's claim that the hallucinations resulting from brain stimulation lack a phenomenal character is motivated by his independent commitments about the nature of phenomenal character and the particular version of disjunctivism he proposes, whereby hallucinations lack any phenomenal character whatsoever and only seem to have one in virtue of their cognitive effect and functional role. I will discuss Fish's view on this point and argue for my disagreement in the next chapter.

one might turn to other empirical data that do not involve hallucinations. One type of evidence that Fish considers (and dismisses) is a number of clinical studies of perceptual impairment undergone by patients with lesions on certain areas of the brain. Fish for instance mentions Zeki's (1990, 1991) review articles of studies showing that patients with damage to the ventral surface of the occipital and temporal lobes lose the ability to perceive in colour, while preserving otherwise normal vision. These kinds of data show, unsurprisingly, that neural activity in fairly well-defined areas of the brain is necessary for the occurrence of experience with certain kinds of phenomenal character. But clearly this correlation is not enough to support local supervenience, which is a claim about sufficiency. One might still try to run an argument for local supervenience as an inference to the best explanation, claiming that a certain brain activity is necessary for the occurrence of a certain phenomenal character because it is the product (in the sense that it is directly caused or supervenes on) of activity in a specific area of the brain. But as Fish points out, there is no reason to prefer this explanation to others. An alternative explanation—Fish proposes—would be to consider the brain activity of a certain kind as an enabling condition of the awareness of, say, red, rather than as the sufficient cause for my awareness of red (see Fish 2009: 137).

In conclusion, contrary to what is often maintained—although rarely explicated or argued for—neuroscience doesn't support local supervenience. The truth is that we know very little about the connection between brain states and phenomenal experience to either confirm or disconfirm local

supervenience. Rather than being supported by empirical evidence, local supervenience is a powerful assumption that informs scientific thinking and practice.

5.5 The Causal Argument from Hallucination

An alternative explanation of the link between the mind and the brain assumed in (b) such that it can support (2#) claims that there is a casual link between the two. Accepting the existence of a causal link between the mind and the brain is less demanding than accepting local supervenience and is compatible with non-physicalist views, such as Robison's.

The purported uncontroversial claim we are looking for to characterise the link between neural activity and conscious experience Robison gestures to in (b) could then be some causal link. We have already seen, however, that accepting that there is a suitably general causal law whereby psychological occurrences depend on brain states is not enough to grant that an experience and a hallucination brought about by the same brain state are IIIHs. However, one might think that there is a further principle that applies to causal laws and makes (2#) true.

Many are convinced that such a principle is what is often known as the 'same proximal cause, same immediate effect' principle ('SC/SE' for short). The idea is that the same kind of proximal cause necessarily entails the same kind of effect. Notice that this is significantly stronger than 'experiential naturalism'. The latter claims that if in two situations the effects of an identical cause differ, then some other conditions are in place.

This is because we take causes to have generality, so two causes cannot result in different effects unless other conditions that differ are in place. SC/SE can be seen as hinging on the principle of Newtonian mechanics that there is no 'spooky action at distance' and on what Hume calls the 'locality requirement'—the requirement that causes be spatially and temporally contiguous with their effects. However, SC/SE is committed to more than the refusal of spooky action at a distance. Rejecting the possibility of so-called 'spooky action at distance' predicts that if a distal event has to count as causing an effect, this must be through a chain of causes that are spatially and temporally close to it. This means that, once the distal cause has produced an intermediate cause in the causal line or process, it becomes redundant: the efficacy of the event further up in the chain is pre-empted by the efficacy of the proximal cause, so that if the distal cause were removed, but the proximate cause remained in place, the same effect would be preserved. SC/SE adds something more:

(SC/SE) In the causal chain, the occurrence of the same type of most proximate cause is sufficient for the obtaining of the same type of effect.

This adds the idea that the proximal cause is *sufficient* for the occurrence of the immediate effect and requires the further assumption (which is not contained in the idea of rejecting 'spooky action at a distance') that nothing but the proximal cause may contribute to the effect, and that

nothing may interfere with the proximal cause and prevent it from producing the effect it generally produces.

If we apply this principle to the experiential case, SC/SE would predict that the reproduction of a certain pattern of neural activation will necessarily give rise to the very same experience irrespective of whatever distal cause precedes the neural state and irrespective of whatever further condition applies to the situation. Although accepting the 'same proximal causal, same immediate effect' is less demanding than local supervenience, since it doesn't impose that a certain brain activity be sufficient for a given mental occurrence in all possible worlds, it delivers, for our present purposes, the same result as accepting local supervenience: it establishes that brain activity is sufficient for a mental occurrence, given the laws of nature. Therefore, if we accept the principle, we will be able to prove more than is required to establish (2#): it can directly motivate the acceptance of the claim that the resulting hallucination has the same phenomenal character as perception.⁶³

⁶³ For instance, Foster claims:

Given their exact similarity to what occurred earlier, it seems reasonable to assume that the artificially induced neural signals and the brain responses they elicit will produce a visual experience which subjectively matches the earlier perception, so that it will be with Henry, subjectively, exactly as if he sees a qualitatively identical salmon making a qualitatively identical leap. But, crucially, it also seems reasonable to suppose that, in addition to *subjectively* matching (to being introspectively *indistinguishable*), the two resulting mental episodes will be, at the fundamental level of description, of *exactly the same psychological type—instances of exactly the same psychological state*. (Foster 2000: 24. Italics in original)

Here Foster makes explicit that the 'same proximal cause, same immediate effect' principle entails not only that the experience resulting from neural replication and the target perception introspectively match, but also that they are 'exactly the same psychological type—instances of exactly the same psychological state'.

The principle is often used to mount an independent argument against relationalism that is more direct than the traditional argument from hallucination, because it doesn't need to rely on IND. Indeed, claims (a) and (b) in support of (2#) in the traditional argument from hallucination that I borrowed from Robinson were used by him as the premises of such an argument, which he calls the 'revised causal argument'. In the literature the argument is more often dubbed the 'causal argument from hallucination' and I will adopt this terminology. It goes as follows:

1. It is theoretically possible by activating some brain process which is involved in a particular type of perception to cause an hallucination which exactly resembles that perception in its subjective character.
2. It is necessary to give the same account of both hallucinating and perceptual experience when they have the same neural cause. Thus, it is not, for example, plausible to say that the hallucinatory experience involves a mental image or sense datum, but that the perception does not, if the two have the same proximate—that is, neural—cause.

These two propositions together entail that perceptual processes in the brain produce some object of awareness which cannot be identified with any feature of the external world. (Robinson 1994: 152)

Robinson presents the claim that experiences caused by the same neural activity are subjectively indistinguishable—what I earlier labelled (2#) in the context of the traditional argument—as independent from the claim that experiences caused by the same neural activity have the same phenomenal character⁶⁴ and as not requiring the assumption of the ‘same proximal cause, same immediate effect’ principle. As I explained earlier, the motivations Robinson explicitly provides for the claim that ‘it is theoretically possible by activating some brain process which is involved in a particular type of perception to cause a hallucination which exactly resembles that perception in its subjective character’ (premise (1) in Robinson’s causal argument and premise (2#) in the traditional argument from hallucination) are:

(a) It is possible to ‘freeze’ a neural state by artificially preserving activated all the neurons that were activated when having a perception.

(a) There is a ‘fairly general agreement that our mental life is strongly tied to what happens in the brain’ (Robinson 1994: 152).

Robinson thus seems to believe that the link between mind and brain is a minimal requirement, independent of SC/SE, and that premise (1) of the causal argument (or 2# in the traditional argument) relies on such a link

⁶⁴ Robinson here talks about the causally matching experience deserving the same account as the target perception, not about the sameness of their phenomenal characters. However, I substitute ‘same account’ with ‘same phenomenal character’ in reconstructing Robinson’s view, because the latter is more specific and less prone to ambiguity, and because the target of the current discussion is relationalism, which is a claim about phenomenal characters. I believe this substitution is harmless, because Robinson is committed to the claim that two experiences with the same account have the same phenomenal character and vice versa.

alone. However, as noted before, simply invoking a strong tie between the mind and the brain is too vague, and under a natural understanding of such a link (what Martin calls 'experiential naturalism') insufficient to secure (1). Robinson doesn't provide any further specification of this link. Therefore, it seems to me that Robinson already has SC/SE in mind and that it is this principle that motivates (1) as well as (2). Presenting (1) and (2) as two successive steps in the argument is a rhetorical strategy, reliant on the fact that (1) seems intuitively less contentious than (2).

The following reconstruction of the argument thus better reflects Robinson's reasoning:

1. When one perceives an object O, the presence of O is a distal cause, while the firing of certain neurons is the proximal cause of the effect, E, namely the occurrence of a conscious experience as of O;
2. It is possible to directly stimulate the brain in such a way as to exactly reproduce the neural state one is in when perceiving O, in the absence of O;
3. In causal chain, the occurrence of the same type of most proximate cause is sufficient for the obtaining of the same type of effect (SC/SE);
4. From (1) and (3): the firing of certain neurons is sufficient for the occurrence of the effect E in the case of genuinely perceiving an O;

5. From (2) and (4): in the case where one directly stimulates the brain, and there is no external object O, the firing of certain neurons is sufficient for the occurrence of the effect E, the occurrence of a conscious experience as of O;
6. From (4) and (5): the same effect E occurs irrespectively of the presence or absence of the distal object O;
7. If the same effect E occurs in both the situation when one perceives O and when one has a hallucination of O induced by direct stimulation, perception and hallucination deserve the same account.

Should SC/SE be accepted? One could object that Newtonian mechanics and its commitment to the locality requirement embedded in the principle is surpassed and proven wrong by more recent quantum mechanics. However, it doesn't seem implausible to insist that, even if quantum entanglement proves to violate this principle, this doesn't mean that the principle doesn't apply for macro-physical phenomena that can be explained with classical mechanics. Moreover, relationalists would be in a weak position if their rejection of the causal argument from hallucination was entirely dependent on resolutions about the compatibility of different paradigms in theoretical physics.

But once one has made the suggestion that different theoretical principles might be relevant to different domains of reality, it is legitimate to wonder whether it makes sense to extend the principles of Newtonian mechanics that we adopt to explain the macro-physical world to psychological phenomena. Even if we accept some form of physicalism,

applying the idea of spatial and temporal contiguity to the causal link between neural and mental events may indeed be tricky. First, it is not clear where, if at all, perceptual experiences are spatially located, so it becomes unclear what it means for a cause to be spatially adjacent to the effect. Moreover, even temporally, it is not entirely straightforward to identify the most proximal cause. However, granting these difficulties, it seems intuitive to claim that the firing of neurons happens fractions of second after the causal impact made by the object through light rays on the subject's retina. Similarly, someone who is not bothered by the claim that mental occurrences have a location might maintain that, although it is not clear where an experience is located, wherever it is, it must be closer to whatever happens in the brain of the perceiver than it is to what happens in the environment outside him.

There is a further problem with locality: its intuitive strength relies on the modern atomistic idea that space is made up of discrete positions that can be occupied by objects, so that for any location, there is a next location, such that if two objects are in contiguous locations, they touch each other. However, more recent theories claim that space is not divided into particles; it is rather measured in rational or real numbers, such that for any two locations it is always possible to find intermediate locations.⁶⁵ So it makes no sense to talk of objects touching and contiguous locations: at best, we can talk about arbitrarily small distances between two locations.

⁶⁵ See LePoidevin (2002) for a discussion of this point.

In conclusion, there are several difficulties with SC/SE. However, the bullets that the proponent of the causal argument from hallucination needs to bite in order to preserve the application of SC/SE to experiences will not do too much damage to her dentistry. The most serious problem with SC/SE doesn't lie in commitments to do with the metaphysics of causation and other claims in the philosophy of science. The problem is rather that SC/SE meets all the objections previously addressed to local supervenience: we should allow that factors such as background enabling conditions, the embodiment of the experience, the requirement of conceptual capacities etc. play some role additional to the proximate cause in the resulting effect.

In his formulation of the causal argument, Foster allows for such conditions

granted that the relevant neural factors are qualitatively the same on the two occasions, and *assuming that we hold constant all the other relevant factors* pertaining to the subject's current bodily and psychological condition, there seems to be nothing in the causal influences at work which might lead to a difference in psychological outcome. (Foster 2000: 24; my italics)

Foster acknowledges that a series of background conditions and factors might alter the capacity of the proximate cause to produce the relevant effect: the subject must be in the relevant mental states and must have the relevant cognitive capacities; and other physical conditions regarding the subject might be needed too (for instance, the brain in

question must be part of a functioning body). Strictly speaking, then, for Foster the appropriate brain activity is not sufficient for the occurrence of an experience of a certain kind. However, since he thinks that a hallucination brought about by neural replication would be 'exactly the same psychological kind' as the target perception, he clearly doesn't consider the presence or absence of the distal object as one of the '*relevant* factors [...] which might lead to a difference in psychological outcome'. What is the rationale for thinking that the embodiment of a brain is a relevant factor, while the presence of the mind-independent object is not? Foster doesn't explain this and it might be thought that this is arbitrary.

Martin suggests that excluding the presence of the object from the relevant factors consists precisely in having already made assumptions about the nature of experience and its phenomenal character and thus begging the question against relationalism,

for the principle so conceived rules out the possibility that relational states of affairs or events can form part of a causal nexus where relational states of affairs may differ purely in their distal elements. (Martin 2006: 368)

Relationalism claims that the phenomenal character of perception is a relational property whose relatum is the object and therefore constitutively depends on the object one perceives. Relationalists don't need to accept that some spooky action at a distance is possible: they agree that the causal role that the object plays in determining the experience is mediated by a chain of

events, including brain activity of a certain type. However, the distal object also has a constitutive role in determining the phenomenal character of perception. The mere presence of the object cannot, alone, produce the perception: a causal link between such an object and the psychological occurrence must be in place, and this will certainly include brain activity of a certain type. But, although not sufficient for the occurrence of the perceptual experience, for the relationalist the presence of the object is a necessary condition for its occurrence. Thus, the absence of the distal object is sufficient for the failure of the relational occurrence that is perception to obtain.

The fact that the same proximal cause involved in perception can produce an experience indistinguishable from a genuine perception in the absence of the distal object is compatible with perception being different from the hallucination caused by neural replication, in virtue of its constitutive relation with the distal object.⁶⁶ Therefore, the presence of the object should be included among the conditions taming SC/SE. But if we do include this condition, it doesn't follow that the hallucination brought about by the same neural state necessarily has the same phenomenal character. Thus, the causal argument from hallucination fails. This also means that SC/SE doesn't support (2#), because, as far as we know, the presence of the object could make an *intropsectibly noticeable* difference in the phenomenal

⁶⁶ A similar remark is made by Langsam when he claims that 'only intrinsic changes that result from the operation of identical causes must be the same' (Langsam 1997: 190), where 'intrinsic changes' exclude changes in relations obtaining between the object and the subject of experience.

character of perception with respect to that of a hallucination resulting from an identical neural state.

5.6 Causation and Constitution

The relationalist's response to the causal argument lies in the suggestion that we may well reject the possibility of 'spooky action at distance' and agree that the distal object plays a causal role in producing the experience only through the mediation of causes further down in the causal chain. However, the distal object can still make a non-causal contribution to the phenomenal character in virtue of being its constituent. An opponent might challenge this response by arguing that a cause cannot be a component of its effect: if we accept that something plays a causal role in bringing about an effect, we cannot also say that it is a metaphysical condition on the existence of that effect, such that its absence is sufficient for effect's failure to obtain. The relationalist, then, one might be tempted to say, cannot claim that the object is the distal cause of the perceptual experience and also a necessary constituent of it.

Hinton considers this objection in the following passage:

There is the feeling that [...] every *effect* [...] must be what you might call 'narrowly identifiable'; meaning that one can state the what-it-is of it, to a degree of exactitude which satisfies normal human interest in the matter, without having to know what its proximate, let alone remote, cause is. [...] The feeling [...] is based only on the tautologous fact that an effect is always

narrowly identifiable. [...] An objector may [...] wish to assert that an effect can always be described, and is always described, in terms that *do not logically entail the previous or simultaneous occurrence of any cause*, proximate or non-proximate. (Hinton 1973: 80–81; italics in original)

The objector to which Hinton refers argues that, once we start thinking about experience as the outcome of a causal process, we realise that any view committed to the claim that the event brought about by a causal process logically entails the existence of any of the causes involved in the causal process is untenable. This is because it seems to violate the widely accepted idea, which goes back to Hume, that ‘only logically distinct existences can be causally related’ (Strawson 1988: 52).

Hinton characteristically doesn’t say who these opponents are or where concerns about the implications of this Humean Principle for the account of perception can be found in the literature. However, we know that such concerns were at the forefront of the debate on the casual theory of perception, which dominated the debate on perception for more than a decade, starting with the publication of Grice’s ‘The Causal Theory of Perception’ in 1961.

Proponents of the causal theory, championed by Grice (1961, 1988), Pears (1976), and Strawson (1974, 1988), maintain that, as Strawson puts it:

The idea of the presence of the thing as accounting for, or being responsible for, our perceptual awareness of it is implicit in the

pre-theoretical scheme [of commonsense realism] from the very start. (Strawson 1988: 103)

Causal theorists claim that the ordinary notion of perception is a causal concept. They are interested in the analysis of the concept of perception and claim that the idea of a causal link with the object plays a fundamental role. As Snowdon notices, it is not entirely clear what philosophers mean by 'causal concept', and he identifies at least two different ways of understanding it:

causal concept₁: 'C counts as a causal concept just in case it is an a priori truth that a (positive) judgement containing the notion C necessarily requires for its truth, that is to say entails, that a (certain sort of) causal condition obtains' (P. Snowdon 2011: 125).

causal concept₂: 'a concept such that it is a necessary condition for a subject to count as having that concept that she or he holds that the concept applies only if a causal link of some kind obtains' (Snowdon 2011: 125).⁶⁷

Whether Strawson (1979) has in mind causal concept₁ or causal concept₂ is not particularly relevant here. I will suppose he accepts causal

⁶⁷ Snowdon also identifies a weaker and more vague notion, causal concept₃—'the idea that by and large any relatively mature person with the concept takes it that it applies only if a (sort of) causal link obtains' (Snowdon 2011: 125)—but he thinks it is mainly used by developmental psychologists rather than philosophers, and certainly not by those interested in the project of conceptual analysis, such as causal theorists.

concept₂, only because it is less committal than the former. What is interesting for our purposes here is that he sees the Humean Principle as possibly posing a challenge for the causal theory:

The correctness of the description of a perceptual experience as the perception of a certain physical thing *logically* requires the existence of that thing; and the *logical* is thought to exclude the *causal* connection, since only logically distinct existences can be causally related. (Strawson 1979: 136–7. Italic in the original)

Strawson is worried by a problem that is the converse to that highlighted by Hinton's fictional opponent. But we will see that the strategy he proposes for tackling the challenge that the Humean Principle poses to the causal theory also neutralises the challenge for Hinton and the relationalist, at least as long as the Humean Principle is formulated in terms of logical entailment, as is the case in both Hinton and Strawson's formulations.

Strawson might be seen as presenting what looks like an inconsistent triad:

1. Humean Principle: cause and effect involved in a causal relation are logically distinct existences, so it cannot be the case that the effect *logically* requires the existence of the cause to obtain;

2. Causal Theory: it is a necessary condition for the concept of perception to apply that the perception is caused by the object perceived.
3. Relationalist intuition: 'The correctness of the description of a perceptual experience as the perception of a certain physical thing *logically* requires the existence of that thing.'

Strawson argues that the conjunction of (2) and (3) does not violate the Humean Principle. Things and events can be described in many different ways, and perception can correctly be *described* in such a way that it both logically requires the existence of the thing perceived and requires that the very same thing is causally related to it. This situation has many parallels:

Gibbons would not be the historian of the decline and fall of the Roman Empire unless there had occurred some actual sequence of events more or less corresponding to his narrative. But it is not enough, for him to merit that description, that such a sequence of events should have occurred and he should have written the sentences he did write. For him to qualify as the *historian* of these events, there must be a causal chain connecting them with the writing of the sentences. [...] The recently much canvassed 'causal theory of reference' merely calls attention to another instance of the causal link which obtains between thought and independently (and anteriorly) existing thing

when the former is rightly said to have the latter as its object.

(Strawson 1979: 139)

As Steward puts it in her discussion of Strawson's treatment of the Humean Principle:

It may often be logically required, in order that something (Gibbon's writing of certain sentences [...]) qualifies for a *description* of a certain sort, that the thing in question be causally related in a certain way to something else (a series of historical events [...]). But of course, Gibbon's writing of his sentences is a distinct event from those he is writing about.

(Steward 2011: 140–1)

In order for Gibbon to qualify as a historian of Roman Empire, the Roman Empire must have existed, and there must have been a causal chain between the events occurring during the Roman Empire and Gibbon's writing about them. Similarly, it is certainly required, for something to count as a perception of an object, both that the object is present and that the perception is caused by the object, but this doesn't mean that the perception and the presence of the object are not distinct occurrences, in the same way as Gibbon's writing and the events described in his writing are distinct events.

The fact that it is possible to describe a visual perception in such a way that the truth of 'S sees P' entails the truth of 'There is P (in front of S)'

doesn't make it less true that the occurrence of the perceptual experience and the object perceived are two distinct entities. This allows Strawson to claim that perception is a causal notion, a claim that would be in tension with the idea that perception entails the presence of the object if this meant that perception and the object are not distinct existents. But his response can be used to defend relationalism as well. If correctly understood, the Humean Principle is not in contradiction with the correctness of the description of perception as entailing the presence of the object. Believing otherwise would be a category mistake:

Only distinct natural items ('distinct existences') can be *causally* connected; only things capable of truth or falsity (propositions or propositional contents, the contents of assertions, hypotheses, descriptions, etc.) can be *logically* connected. The fact that the truth of a *description* of one item logically requires the truth of the existence of another has no force at all to show that the item in question are not distinct natural existences, capable of being causally linked. The idea that these natural items *themselves* could be logically linked is nonsense, a category howler. (Strawson 1998: 314; italics in original)

A fundamental component of Strawson's response to the challenge is the distinction between what can be *logically* connected (or not) and what can be *metaphysically* connected (or not). It wouldn't make sense to understand the Humean Principle as claiming that the effect cannot *logically*

require the existence of the cause. Indeed, it is possible to describe all sorts of things in such a way that the truth of that description entails the truth of the description of their cause. If the Humean Principle is to have any force, it must be understood in a metaphysical way, as stating that the existence of the effect is metaphysically independent of the existence of its cause. If we come back to the triad, then, Strawson invites us to revise the way the Humean Principle is formulated: it shouldn't read as a claim about logical inferences—'the cause and effect involved in a causal relation are logically distinct existences'—, but rather as a metaphysical claim—'the cause and effect involved in a causal relation are metaphysically distinct existences'.

However, if we understand the Humean principle as a metaphysical principle, the response that Strawson envisages may work as a defence of the causal theory of perception, but it doesn't work as well as a defence of relationalism. For Strawson, what needs to be accommodated with the Humean Principle is the intuition that there is a (natural) way of describing perception such that the truth of this description entails the existence of the object. This description, he says, remains silent with respect to the metaphysics of perception and what is required for one to perceive. But it is precisely its silence with respect to the metaphysics that makes this response unsuitable for the relationalist. Saying that the event brought about by the causal chain can be *described* in relational terms, such that the truth of the description logically entails the presence of the object, leaves it open that other descriptions capture the event and that the correctness of some of

these might not logically require the existence of the thing.⁶⁸ Indeed, the relational description is compatible with the Humean Principle because it doesn't reflect any metaphysical dependency between the perception and its object.⁶⁹

If this is all the relationalists have to offer, one might wonder what they mean when they claim that perception is constituted by the object in front of one. With their talk of constitution, relationalists might be seen to be trading between the platitude that the truth of 'S sees O' requires the existence of an O and a more committal metaphysical claim that the occurrence of the experience involved in a perception of an O metaphysically involves the presence of O. But only the platitude is compatible with the Humean Principle, correctly understood in metaphysical terms. Or so the argument goes.

This worry can be related to a widespread dissatisfaction with the formulation of relationalism in term of constitution. What does it mean to say that perception is partly constituted by the object? Is there anything more to it than that the platitude 'S perceives O' logically entails the existence of O? Such scepticism is expressed once again by Strawson, in his response to Snowdon, whose relational view he writes off as a 'non starter' (Strawson 1998: 314):

⁶⁸ In a similar way, Steward (2011: 141) notices that the disjunctivist is unlikely to be satisfied by the kind of reconciliation proposed by Strawson.

⁶⁹ In fact, it seems to me that the success of the Strawsonian defence of the causal theory is conditional on the assumption that the description of perception in relational terms does *not* reflect the metaphysics of perception at all: perception can be fundamentally a causal concept only if perception doesn't metaphysically require, for its occurrence, the existence of the distal cause.

It is not clear to me how Snowdon conceives of this [...] relation. He seems to suggest that when a subject, for example, sees an object, the visible object is somehow, partly *constitutive* of the experience. Again it is not quite obvious how to understand this. But the plainest gloss on it is the following. If a subject S sees an object O, then S's experience is accurately described in the words 'S sees O'. But the truth that S sees O logically requires the existence and presence of O. It would be logically impossible for *that* experience to occur without the presence of O. So the relation between experience and object is certainly non-accidental: it is logical. (Strawson 1974: 313)

Strawson sees the claim that the object is a constitutive part of perceptual experience as a way of characterizing the relation occurring between object and experience, when the experience is a genuine perception, alternative to the causal relation. The latter, according to Strawson, is a non-starter because it fails to characterize the nature of the relation between the object and the perceptual experience. It fails to do so because, as we have seen, a logical relation holds only between things capable of truth and falsity, and hence not between the object and the perceptual experience themselves, but descriptions of them. But a logical relation between *descriptions* of entities remains silent with respect to the nature of the relation between these entities.

This seems an accurate, or at least a very plausible, way of understanding Snowdon, who in the paper to which Strawson responds,

presents the notion of constitution as 'a way to specify a noncausal relation in which objects can stand to experiences when those experiences amount to perceptions of those objects' (Snowdon 1998): 302). Moreover, in his paper 'Perception, Vision and Causation', where he first introduces disjunctivism as part of his strategy to undermine arguments for the casual theory of perception, Snowdon presents disjunctivism as follows:

The disjunctive picture divides what makes looks ascriptions true into two classes. In cases where there is no sighting they are made true by a state of affairs intrinsically independent of surrounding objects; but in cases of sightings the truth-conferring state of affairs involves the surrounding objects. (Snowdon 1980): 186)

Saying that look ascriptions are made true by either states of affairs intrinsically independent of surrounding objects or by states of affairs involving surrounding objects is compatible with the fact that in both cases there is the very same kind of experience, namely a metaphysically independent state of affairs that doesn't involve the object. The difference is that sometimes, when we make look-statements, we are in fact referring to things that not only are experiences, but are 'sighting[s]', genuine perceptions. That is to say, those experiences are also correctly described as perceptions and what makes this description true is a more complex state of affairs that involves the experience and its surroundings. Although Snowdon claims that a look-statement may be true in virtue of 'two distinct sorts of

states of affairs' that have 'a quite different nature' (Snowdon 1990: 120), it is not clear how these two states of affairs are different, beyond the platitude that, in order for one to successfully perceive O, an O must be present; the state of affairs we have to consider is one that includes the presence of such an object.

It is not clear whether Snowdon is committed to the claim that the only interesting relation between the object and perception is a relation of logical entailment. It might be that he thinks that the logical entailment is all he needs to counter the causal theory, so, in the passages above, he only subscribes to a relation of logical entailment, even though he is inclined to accept a stronger, metaphysical dependency between the object and perception. But what Snowdon exactly commits to is irrelevant to our current discussion. We do not need to take Snowdon to be representative of relationalism: despite his clear inclination towards relationalism, Snowdon has never explicitly engaged in a discussion of relationalism—understood as a view about the phenomenal character of experience—and his interest in disjunctivism is not primarily motivated by a defence of relationalism. So, even if Snowdon defends disjunctivism, it is not clear that he is committed to relationalism as well.⁷⁰

⁷⁰ Martin notices that, for Snowdon's purposes, i.e. undermining a line of argument against the causal theory of perception, disjunctivism may take a form that mirrors Williamson's treatment of knowledge, which is 'consistent with (but does not strictly demand) a view on which the subjective component of one's mental life, all aspects of what it is like for one to be as one is, are comprised of the common states, beliefs and appearances, while the states present in only certain cases, knowledge or perception, while of explanatory importance in other realm, play no role in determining the conscious life' (Martin 1997: 97). This would be inconsistent with relationalism (that Martin calls Naïve Realism), as this seeks to give an account of phenomenal character.

The relationalist is committed to more than the simple claim that there is a logical relation between perception (suitably described) and the object perceived. She is committed to saying that the object is a metaphysical constituent of the phenomenal character of perception. This is not to say that there is an interesting relation that obtains between two entities: the object and the experience. Talking in these terms—as Snowdon (Snowdon 1998: 302) does—suggests that there is an experience that is independent from the object and could in principle exist in its absence. If understood this way, it is not clear what exactly the specification that the relation is a relation of constitution adds: thus understood, the act of perception and the object remain metaphysically distinct. Relationalism claims that perception *is* a relation. Perception amounts to becoming aware of an object, property, or event: it *is* the obtaining of a relation. The fact that the object constitutes the perception is thus an obvious implication of the fact that perception *is* a relation to the object and no relation may exist without its relata, which thus constitute it.

The following passage by Fish is representative of this idea, shared by relationalists, that perception *is* itself a relation with the object rather than being an occurrence that bears some interesting relation to the object:

The distinctive feature of naive realism⁷¹ lies in the claim that, when we see the world, the subject is acquainted with the

⁷¹ As we saw in the first chapter, what I have labelled here ‘relationalism’ corresponds to what many philosophers call ‘naïve realism’. My choice reflects the importance I attribute to stressing that the view in question construes perception *itself* as a relation.

elements of the presentational character—the mind-independent objects and their features—where ‘acquaintance’ names an irreducible mental relation that the subject can only stand in to objects that exist and features that are instantiated in the part of the environment at which the subject is looking. (Fish 2009: 14)

Thus, it is *not* an option for the relationalist to claim that the Humean Principle is compatible with relationalism. What other options are available to the relationalist, then? It might be easier to see the available options if we consider what exactly are the supposed inconsistent claims that the relationalist needs to reconcile:

- 1* Humean Principle (metaphysical version): events that are causally connected are metaphysically distinct existences;
- 2* Causal Intuition: the distal object plays a role in the causal process leading to perception.
- 3* Relationalist commitment: The object partly constitutes the phenomenal character of the experience, so that the occurrence of the perception requires the existence of object.

The relationalist commitment and the Humean Principle are indeed incompatible, unless one denies that the object plays any role in causing the perceptual experience, which seems very difficult to do. A way to deny (2*), however, would be to say, following Davidson (1967), that objects are

not in the business of causing, because only events can be causes. Thus one might say that there is no contradiction because what plays a constitutive role, as relationalism has it, is the object, while what causes the experience is an event involving that object, such as its surface reflecting light in a certain way, which exists independently of the perceptual occurrence. What does the causing (an event) is not a constitutive element (an object).

Irrespective of whatever we take the metaphysics of causation to be and the specific merits of this response, it cannot be the relationalist's final say on the matter. Among things one can see, there are not only objects, but events too. And the relationalist is committed to claiming that, in the case of the perception of events, the event also partly constitutes the phenomenal character of the experience (Soteriou 2010; 2013: 89). But one could hardly deny that the event seen is among the causes of S's seeing the event, as events are indeed in the business of causing. So we may take 'object' here as a placeholder for whatever is the object of perception—be it an event, a property; an object or whatever else—and accept (2*).

But should we accept (1*)? Should we accept that it applies unconditionally across the board and that that perception is not an exception to it? One way of challenging the Humean Principle is offered by Child (1992). His strategy is in some ways similar to Strawson's strategy against the logical version of the Humean Principle, inasmuch as he presents ordinary cases where things are taken to be causes of other things, although they are also constitutive parts of them. Child argues that there are many occurrences that are both related by a causal relation and by a metaphysical

relation of constitution, such as cases of events that are part of a bigger event and also the cause of it. An example of this is the bombing of Pearl Harbour as both the cause of the Pacific War and a part of it (Child 1992: 310). The Pacific War is not metaphysically independent of the bombing of Pearl Harbour: it wouldn't be the event that it is considered to be if the bombing of Pearl Harbour wasn't part of it. However, it remains true that Pearl Harbour caused the Pacific War. This example shows that there are some counterexamples to the Humean Principle. Even if there might be cases where it applies, it cannot be applied across the board. Cases where a portion of an event is also a cause of the event itself are immune to Humean Principle.

Child suggests that we should understand perception along these lines, treating the relation of constitution as a part-whole relation, where the object is a component or part of the perceptual experience. The problem with this analogy is that there is a difference between the case of Pearl Harbour and the Pacific War and the case of a perception and the object perceived. An aspect of our understanding of the bombing of Pearl Harbour as *part of* the Pacific War is that we think of the Pacific War as unfolding over a period of time that starts with the bombing of Pearl Harbour. On the contrary, we do not think that the event of perception begins at the time the object perceived begins to exist. The object one perceives may have ceased to exist hundreds of years before the occurrence of the perception. A famous example of this is perception of stars that have extinguished by the time their light reaches us: we do not think that the perception began when the

star was formed, nor do we think it began at the time the star emitted the light that travels towards us and is eventually seen. This is true even if we consider events, rather than objects. When one sees a comet crossing the sky, we do not think that the seeing begins at the time the comet passes close to sun, heats up, and begins to emit gas—again an event which occurs at least a few minutes before the light from the burning reaches Earth. This indicates that the part–whole model is not particularly helpful here.

Although Child's suggestion that the relation between perception and its object is a part–whole relation may be imperfect and partly misleading, it does indeed show that there are counterexamples to the Humean Principle. Perception might be one of these exceptions, although not exactly for the reasons proposed by Child. It is not difficult to see what the relationalist could say to support the claim that perception represents an exception to which the Humean Principle doesn't apply. The Humean Principle certainly can't apply to relations and relational states. It would be ridiculous to claim that a relation could exist independently of its relata. A relation doesn't obtain without its relata, so it cannot be metaphysically independent of them. Accepting the Humean Principle across the board would be tantamount to denying the possibility, for any relation, of its being caused by one of its relata. But it is not clear why we should deny this. On the contrary, there is no difficulty in accepting that an object causes a perception and that, nonetheless, the perception is a relation to this object and, as such, metaphysically depends on it.

Still, one might think that there is a deeper intuition about the Humean Principle that needs to be true of relations and relational states too. We can grant, one might argue, that an object causes something whose existence constitutively depends upon the object that did the causing. But if we want to make sense of the idea that there is genuine causation between the object and its effect, i.e. the perceptual experience, it must be the case that the relation involving the object that is caused obtains only in virtue of the obtaining of something else, whose existence is independent of the existence of the cause. In other words, it seems that we can grant that the obtaining of a relation to an object (S perceives O) is among the effects of that object only if this relation obtains in virtue of the occurrence of an effect which is, in turn, metaphysically distinct from its cause (and hence could obtain in the absence of the object).

But relationalism doesn't seem to have any difficulty in meeting this constraint. In fact, we have seen in this chapter that relationalism is perfectly compatible with the idea that there is something that is caused by the object of perception and that could exist even in the absence of the object. An opponent of relationalism would see this as evidence that perception, although understood as intrinsically relational, needs to involve a non-relational occurrence, an experience whose existence is metaphysically independent of the object.

The relationalist can retort that it might well be that a weaker version of the metaphysical Humean Principle holds for perception and that we must assume that the relation to an object that perception obtains only in

virtue of the obtaining of another occurrence which is itself metaphysically independent of its causing object. However, there is no need to identify this metaphysically independent occurrence with an experience. Why should one introduce a further mental occurrence, an experience, when there is already an occurrence, in virtue of which the perception occurs, that is metaphysically independent of the object perceived? The alternative candidate is the neural state. We have seen that there doesn't seem to be any reason to reject the possibility that exactly the same neural state that occurs when one perceives an object may occur in the absence of the object when one hallucinates, and indeed relationalists can accept this. Thus the obvious thing to say for the relationalist is that the object causes the neural state, an effect that is metaphysically independent of the object. In virtue of this neural state, one has a perception of the object, which is the obtaining of a relation to the object, and is thus metaphysically dependent on the existence of the latter. The perceptual relation is itself an effect of the object, but it is so only in virtue of the brain state. There is no reason, then, to think that relationalism cannot accommodate the Humean Principle properly understood.

5.7 Where Does This Leave Us?

I have discussed what principle, claim, or intuition might underwrite the acceptance of premise (2#) in the revised argument from hallucination. On the one hand, I have considered local supervenience as a possible motivation for (2#). Such a principle, however, can appeal only to

physicalists, and most importantly, it requires acceptance of a very committal and controversial version of physicalism, that many who are otherwise attracted to physicalism would reject. It is also more committal than (2#), as it begs the question against relationalism, while (2#) could in principle be accommodated by relationalists. Thus local supervenience is far from being the fairly uncontroversial principle on which the acceptance on (2#) may rely. Far from being in a position to motivate (2#), local supervenience is often argued for using arguments that assumes (2#).

On the other hand, one might think a good candidate for motivating (2#) is 'experiential naturalism'—the idea that experience is subject to physical and psychological causes (if the latter are disjoint from the former). 'Experiential naturalism' is widely accepted as a methodological assumption for both theoretical and empirical work and rejecting it would be quite demanding;⁷² however it doesn't entail anything like (2#). I then considered the possibility that experiential naturalism might succeed in delivering (2#) if paired with another assumption, the 'same proximate cause, same immediate effect' principle. I have argued that this principle is not less problematic than local supervenience, and it equally begs the question against relationalism.

Finally, in the previous section I considered a line of argument that is meant to undermine the relationalist's response to the causal argument. The relationalist argues that, even if the distal object causes the perception, it

⁷² Although demanding, it is not impossible to reject 'experiential naturalism'; and Martin mentions the early sense-datum theorists, some phenomenologists, such as Merleau-Ponty (1943) and Valberg (1992) as examples of philosophers who reject it.

may also play a constitutive role in determining its nature. If it is true that nothing that figures in the causal chain of a thing can be one of its metaphysical components, then the causal argument is rehabilitated. I have argued that even this argument fails because, even if it were true for most things involved in causal processes, this cannot be true of relations and relational states of affairs.

Although it doesn't presuppose physicalism, SC/SE has practically the same consequence as local supervenience: they both entail that the neural activation is sufficient for the occurrence of a phenomenal character of a certain kind, all other relevant factors remaining unchanged. The difference is that supervenience secures this conclusion in all possible worlds (or in nearby possible worlds, depending on the strength of the supervenience claim), while the SC/SE principle secures it only given the current psychophysical laws. Therefore, both supervenience and SC/SE are used to mount independent arguments from hallucination against relationalism, which do not make use of IND. However, I have argued that these arguments fail.

The first lesson to be learned from this discussion, then, is that the most promising version of the argument from hallucination against relationalism remains the traditional one, in which IND plays a pivotal role. IND remains central to assessing the tenability of relationalism in the face of hallucinations. Opponents of relationalism ought to provide motivations for IND, and relationalists need to explain how they can accommodate it. This stands in contrast with a recent trend in the debate on the argument

from hallucination, which tends to take the traditional argument to be flawed, while focusing on some version of the causal argument from hallucination, seen as the real threat to relationalism.⁷³ Proponents and opponents of relationalism alike share this attitude. Matthew Nudds, for instance, notices that:

Many people find the common kind assumption compelling because they accept some version of the causal argument [from hallucination].⁷⁴ (Nudds 2013: 274; italics in original)

Thus, Nudds seems to suggest, a defence of relationalism should focus on responding to the causal argument, rather than assessing the key assumption in the traditional argument, i.e. IND. Indeed, Robinson (1994), before presenting the causal argument from hallucination that he deems successful, discusses an argument along the lines of the classical argument from hallucination, but ultimately rejects it as question-begging. He sees the argument as relying on a principle that is a close cognate of IND, and spells it out as follows:

⁷³ Although I have presented the argument from local supervenience as independent from the causal argument, local supervenience is sometimes used to motivate something along the lines of SC/SE or replaces it in an argument with the same structure. This is unsurprising, given the similar implications of the two principles.

⁷⁴ As we saw in Chapter 2, according to the common kind assumption, whatever kind of experience does occur when one perceives, the same kind of experience can be present when one is hallucinating.

Things can be qualitatively similar only if they are of the same general kind and have the same ontological status (Robinson 1994: 88),

only to conclude that it is 'not so obviously true that it can be used to support a controversial conclusion' (Robinson 1994: 88). Similarly, influential discussions of the argument from hallucination incorporate elements of the causal argument or the argument from local supervenience that make the commitment to IND redundant (Smith 2002: 200; Foster 2000; Sollberger 2007, 2012). If my diagnosis of the local supervenience and the SC/SE principle are correct, however, these are not less, but more controversial than IND. Therefore, those more direct ways of arguing against relationalism without relying on IND are unsuccessful.

The second lesson of the present discussion is that the reasons implicitly or explicitly put forward to motivate (2#)—experiential naturalism, local supervenience, and SC/SE—either fall short of delivering it or are themselves largely controversial.

Does this mean that we should reject (2#)? This would be premature. All of the positive motivations considered here fall short of supporting (2#); however it is not clear that one can so easily deny it. So far, we have established that there is no conclusive argument for (2#), so it is possible for one to deny it. But it might turn out that denying it would require one to commit to incredible views. Instead of considering what arguments support (2#), we now need to consider what would it take for one to reject it. One who wants to reject (2#) needs to offer an alternative picture of what

happens to the subject after her brain has been stimulated to induce a brain state that perfectly matches the brain state of a genuine perception in the absence of the relevant distal cause. If the matching neural state doesn't produce an experience that is indistinguishable from the genuine experience, what else happens? Is there any credible story for what follows the reproduction of the matching neural state?

One option would be to say that nothing at all happens. With the rejection of local supervenience and SC/SE, we have rejected the idea that brain activity of a certain kind is sufficient for the occurrence of conscious experience of a certain kind. So it might well be that, without the presence of the suitable distal object, no conscious experience at all occurs. Maybe there is some transmission of (misleading) information, which might eventually have behavioural and cognitive effects on the subject, akin to cases of blindsight (Weiskrantz 1986; Weiskrantz 1996; Weiskrantz 2009; Azzopardi and Cowey 2001).⁷⁵

The discussion of local supervenience presented in this chapter, however, poses serious problems for this line of response. Empirical data

⁷⁵ Notice that this option, which, as far as I know, has not been defended, is significantly different from Fish's (2009). Fish too maintains that hallucinations brought about by neurally matching situations have no phenomenal character whatsoever, for reasons I will discuss in the next chapter. However, in his view, the lack of phenomenal character in such hallucinations is no reason to deny (2#). For Fish, when one has a matching hallucination, it seems to one as if one has a conscious experience, with a naïve realist phenomenal character (a character consisting in putting the subject in relation to mind-independent objects), even though this apparent phenomenology is only an illusion, the by-product of the fact that perception and matching hallucinations share behavioural and cognitive effects. In the view I suggest here, on the contrary, the hallucinating subject might display some behaviour in common with someone veridically blindseeing, such as guessing that a certain type of object is in the blinded visual field. However, like the blindseer, the subject would not take herself to have any conscious experience. This means that a lack of phenomenal consciousness in such cases is identifiable.

such as Penfield's studies show that *some* brain activity—even if it doesn't exactly reproduce the brain activity of a genuine perception—is sufficient for the occurrence of an experience with a phenomenal character of *some* kind—although there is no evidence that it can produce an *impersonally introspectively indistinguishable* conscious experience. If, in general, some brain activity is sufficient for the occurrence of some phenomenal character, it would be very difficult to explain how a brain state that exactly matches the brain state occurring in a genuine perception would fail to produce any phenomenal character at all. It is more natural to suppose that the more the brain state resembles the pattern of neural activation that occurs when genuinely perceiving, the more likely it is that this will generate a conscious experience with phenomenal character.

Even if one accepts that reproducing the mental state is enough to produce a conscious experience with phenomenal character, one can still resist (2#) by denying that the experience thus produced is introspectively indistinguishable from a genuine perception, at least in the impersonal sense of indistinguishability required for the argument to succeed (see Chapter 3). But what would ground this claim? We saw in the previous chapter that hallucinations are introspectively indistinguishable in an impersonal way only if the failure to distinguish them cannot be attributed to any interference having to do with the subject or the conditions of introspection. In the cases of ordinary hallucinations scrutinised in the previous chapter, we were always able to identify some factors that might be responsible for the subject's failure to discriminate: typically, the fact the subject's

cognitive and discriminatory capacity or attentiveness are impaired by a psychoactive substance, brain damage, a medical condition, etc. Moreover, whatever interference occurs, this is likely to correspond to some neurological difference between the hallucination in question and any hypothetical perception.

It was part and parcel of the thought experiment introducing philosopher's hallucinations that nothing is wrong with the subject hallucinating and that she is in exactly the same brain state as she would be when having a perceptual experience: her cognitive capacities are not impaired in any way and she has the same psychological states that accompany and support the experience she has in the case of genuine perception. *Ex hypothesis*, then, none of the factors that could explain why the subject fails to discriminate her current hallucination from the genuine perception of a moment before are present. One would have to indicate some other conditions or factors that interfere with the subject's ability to discriminate two otherwise introspectively distinguishable experiences. However, it is not clear what these factors could possibly be. What makes it the case that the subject fails to discriminate between her current matching hallucination and the previous genuine perception? In the absence of a response to this question, the relationalist is not in a position to challenge (2#) in a convincing way. Although these are not conclusive reasons for (2#), they show that the relationalist would have to answer very pressing questions for which no obvious answer is yet available. Therefore, the relationalist is better off accepting (2#), at least for the sake of argument,

and looking somewhere else to challenge the argument from hallucination. In the next chapter, I will consider a line of response that can be found in Austin's discussion of the argument from illusion.

5.8 Conclusions

In Chapter 3, I argued that IND is motivated by the belief, supported by what I called the argument from superficiality, that relationalists cannot deny IND without also denying the non-negotiable assumption that phenomenal characters are superficial. We have seen that relationalists can accommodate a difference in phenomenal character between perception and an indistinguishable hallucination by appealing to conditions that may interfere with the subject's ability to introspect. Thus, in order to motivate IND, one needs to introduce the notion of impersonal indistinguishability. This means that both the argument from superficiality and the traditional argument from hallucination are successful only if there are hallucinations that are introspectively indistinguishable from perception in an impersonal sense (IIHs).

In the previous chapter, I argued that there is no empirical evidence that can support the actual existence of IIHs. In this chapter I explored potential motivations for claiming that IIHs are metaphysically possible. Although I ultimately accept this possibility, examining the motivations that have been put forward to support the metaphysical possibility of IIHs allowed me to consider the solidity of alternative arguments against relationalism. These arguments, which do not rely on IND, and which

would thus be effective even if the relationalist proved capable of denying IND without violating the superficiality constraint, rely respectively on the principle of local supervenience and the 'same proximate cause, same immediate effect' principle. Since both principles beg the question against relationalism, these alternative arguments fail. Additionally, I argued that, properly understood, the Humean Principle, which may seem to be in tension with relationalist commitments, is indeed compatible with them.

In the absence of alternative arguments from hallucination, the opponent must resort to the traditional one. Hence the debate over relationalism hinges entirely on whether or not we should accept IND. As we have seen, deciding whether IND is to be accepted comes down to deciding whether relationalism can deny it, while preserving the superficiality of phenomenal characters. This will be the task for the next chapter, in which I will argue that relationalists can accommodate the superficiality of phenomenal characters while still maintaining that a perception and an indistinguishable hallucination have different phenomenal characters.

Chapter 6: Introspecting Hallucinations

In Chapter 3 I examined the grounds for accepting the key assumption in the argument from hallucination, namely the idea that when two experiences are introspectively indistinguishable, they have the same phenomenal character (IND). I argued that IND only stands a chance of being true if we understand ‘indistinguishable’ in an impersonal sense. While various explanations appeal to the fact that some conditions can interfere with one’s introspective capacity to distinguish experiences with distinct phenomenal characters when these are *subjectively* indistinguishable, such explanations are not available in the case of hallucinations that are impersonally introspectively indistinguishable from perception (IIIH). Thus the relationalist allegedly can’t account for their having the same phenomenal character as the perception.

As I argued in Chapter 4, there is no evidence that any actual hallucinations is a case in which no interfering condition could prevent one from noticing that the phenomenal character of the hallucination undergone is actually different from that of typical perceptions. However, in Chapter 5 I argued that the relationalist cannot rule out that IIIHs are metaphysically possible. The possibility of such hallucinations arises from the consideration

that it is possible to have hallucinations brought about by the same neural state bringing about a genuine perception.

It remains to see whether relationalism can account for the fact that a hallucination may be impersonally introspectively indistinguishable from a perception while not sharing with it the same phenomenal character, without thereby assuming that the nature of the phenomenal character is hidden to introspection and hence denying that the phenomenal character of hallucination is superficial.

Unless relationalists provide a response to this challenge, we have reason to believe that the best account for the indistinguishability of IIIHs is in terms of the sameness of their phenomenal characters; hence that we have to accept IND as the only credible way to account for the possibility of IIIHs. In this chapter I will explore whether the relationalist can answer this challenge.

Before addressing this question, I will consider a different strategy for blocking the argument from hallucination. One might be tempted to claim that, even if we accept both IND and the possibility of IIIHs, the conclusion of the argument still doesn't follow. I will explore this strategy in the next section, only to argue that this strategy faces a problem—namely what Martin calls the ‘problem of explanatory screening off’ (Martin 2004: 71), originating from a variation of the causal argument from hallucination. As a result of this argument, the only strategy available to the relationalist remains to reject IND. The screening off worry also precludes certain accounts of the phenomenal character of IIIHs and imposes only a negative

account of them. According to the negative account of IIIHs, the phenomenal character of these hallucinations is parasitical upon the phenomenal character of hallucination, for it is exhausted by its being introspectively indistinguishable from that of a genuine perception.

In what follows I will consider some common objections to this account and respond to them. A crucial element in the response to many of these objections is to understand how relationalists should conceive of introspection. With this understanding of the negative account of hallucination and of how introspection works, it will be possible to appreciate how the relationalist can claim that an IIIH doesn't have the same phenomenal character as a genuine perception without committing to the claim that the phenomenal character of hallucination is hidden to introspection and hence forsaking the intuition that phenomenal characters are superficial.

6.1. An Invalid Argument?

The relationalist could argue that, even if we accept IND and the theoretical possibility of IIIHs, the argument from hallucination is a *non-sequitur*. Even if one agrees that perception and an IIIH have the same phenomenal character, it doesn't follow that the phenomenal character of perception cannot be determined by the mind-independent object that we take ourselves to be perceiving.

There are two strategies that relationalists could pursue to reject this conclusion, while accepting all the premises of the argument. The first

option is to claim that perception and hallucination have *one* phenomenal character (however construed) in common, while perception has an *additional* phenomenal character, to be accounted for in relational terms. This proposal is likely to be resisted by most philosophers, who take phenomenal characters to be maximally determinate. The intuition underlying the widespread acceptance of maximal determination of phenomenal characters is the thought that when a perception and an IIIH is indistinguishable from a perception, their shared phenomenal character best explains their indistinguishability, and if their shared phenomenal character explains their indistinguishability, then it's plausible that this shared phenomenal character is sufficient to account for what it is like to experience each. Otherwise the perception would have introspectible features that would distinguish it from the IIIH, but this is *ex hypothesis* to be excluded. I will discuss in greater detail later on, when considering the screening off problem introduced by Martin, why the phenomenal character common to perception and hallucination should be sufficient to account for what it is like to have both perception and a corresponding IIIH.

Alternatively, an opponent of maximal determination might suggest that the additional phenomenal character of hallucination is not an introspectible feature of the perception, and hence not introspectively accessible. But this would amount to denying that the phenomenal character of perception is superficial. On top of violating the superficiality condition, this strategy would hinder relationalism in its effort to account for the way experience strikes us introspectively as being a relation to the object one

seems to perceive, because the relational phenomenal character would not be accessible through introspection, so it couldn't explain why perception strikes as in this way. I will thus grant that phenomenal characters are maximally determinate, leave this strategy aside, and focus on the second one.

The second option is to claim that a perception and an IIIH have the same phenomenal character, although each is constituted by different bases: the former is constituted by the objects we ordinarily take ourselves to perceive and the other is constituted by something else. This option requires one to think that phenomenal characters can be multiply realized. While denying that phenomenal characters are maximally determinate may seem preposterous, the claim that phenomenal characters are multiply realized is in line with the intuitive understanding of phenomenal characters in terms of what it is like to have an experience. It is possible that what it is like to have an experience is identical to what it is like to have another experience, although what determines the phenomenal character of each experience differs.

This is not dissimilar to an intuitive way of understanding observational properties. The lemon and the visually indistinguishable crafted soap may be said to have the same observational properties: yellowness, lemon-shaped-ness, and coarseness of surface. Similarly, under this view, the veridical experiences of the lemon and the lemon-shaped soap have the same phenomenal character. However, according to the relationalist, since phenomenal characters are partly constituted by the

object of perception, they are realised in one case by a relation to the lemon and in the other case by a relation to the piece of soap.

In ‘Particular Thoughts and Singular Thoughts’, Martin discusses how intentional theories of experience can accommodate the particularity of experience. He argues that intentionalists can ‘embrace the conclusion that experiences of identical twins have the same phenomenal character without having to deny that individual experiences have particular objects as part of their phenomenal nature’ (Martin 2002a: 175) by contrasting the ‘unrepeatable aspects of its phenomenology, what we might call its *phenomenal nature*’—the particular entities involved in a perceptual event—‘with that it has in common with qualitatively the same experiential events, what we might call its *phenomenal character*’ (Martin 2002a: 194). The phenomenal nature identifies one specific event (the perception of one rubber duck on a particular occasion, rather than its identical twin), while the phenomenal character ‘abstracts away from the particulars involved in the individual event and just considers the general attributes that the particulars exhibit’ (Martin 2002a: 194).

Similarly, a relationalist could say that two perceptions of two qualitatively identical rubber ducks are distinct events, each involving the occurrence of a relation to a particular. However, the phenomenal character abstracts away from these particulars, so we can say that the two experiences have the same phenomenal characters. Nonetheless, in each case, the identical phenomenal character is constituted by the different particulars involved in the relation. Such a model, it seems, can be applied

to IIIHs. An IIIH and the relevant perception have the same phenomenal character, although each phenomenal character is realised by the obtaining of a relation to a different particular. A theory of this sort is discussed by Price (1932: 40).⁷⁶ ‘The Selective Theory’, as he calls it, claims that, when one genuinely perceives, one is related to mind-independent physical objects, while when one hallucinates, one is related to sense-data.

If one accepts that phenomenal characters can be multiply realised, the Selective Theory is not the only available option. One can maintain that the very same phenomenal character is realised, in the case of genuine perception, by the obtaining of a relation to mind-independent physical objects and, in the case of IIIH, by a representational content, by non-representational intrinsic properties, or anything else. Therefore, as it stands, the argument is a *non-sequitur*: whatever accounts for and constitutes the phenomenal character of an IIIH doesn’t need to account for and constitute the phenomenal character of the corresponding perception. To make the argument valid, one needs to introduce a further premise that rules out the multiple realisability of phenomenal characters and claims that phenomenal characters are maximally determined. A valid version of the argument from hallucination, thus, would be the following:

1. Relationalism claims that the phenomenal character of conscious perceptual experiences is to be accounted for in terms of the

⁷⁶ Price ultimately rejects the Selective Theories on grounds that have little to do with the issues that matter to us in this context, but have rather to do with his attempt to account for the causal efficacy that we attribute to sensible objects (see Baylis 1950: 456).

- obtaining of a psychological relation to the mind-independent objects with which one is perceptually presented;
- 2#. In some cases of hallucination, one has an experience that is introspectively indistinguishable from an experience one may have in a genuine perception;
 3. In the hallucinatory case, the object one seems to perceive doesn't exist;
 4. (from 3) In the hallucinatory case, the phenomenal character of one's experience is not accounted for in terms of the obtaining of a psychological relation with the mind-independent objects with which one is perceptually presented;
 - 5*. If two conscious experiences are *impersonally* introspectively indistinguishable, they have the same phenomenal character (IND);
 6. Phenomenal characters are maximally determinate;
 7. A phenomenal character cannot be determined by different constitutive bases;
 8. (from 4, 5, 6, 7) The phenomenal character of a perception has the same phenomenal character as the corresponding IIIH and this cannot be determined by the obtaining of a relation to the object one seems to be perceiving, as this is not what determines the phenomenal character of the corresponding IIIH.
 9. (from 1 and 8) Relationalism is false.

Although 6 can be granted, it is not clear that 7 should be. The burden of proof is on the proponent of the argument and there doesn't seem to be any immediate reason to deny multiple realisability.

6.2 The Reversed Causal Argument

As it stands, the argument not only fails to get the spreading step in motion; it also fails to impose any constraint on the relationalist's account of the phenomenal character of IIIHs. Unless one rules out the multiple realisability of phenomenal characters, it seems that a relationalist can give any account of the phenomenal characters of IIIHs they deem fit—in the same way that they can give any account they please of the phenomenal characters of hallucinations that are not subjectively introspectively indistinguishable. However, Martin (2004; 2006) argues that an argument related to the causal argument considered in the previous chapter imposes certain constraints on any account that relationalists can provide of the phenomenal character of hallucinations.

Martin calls the argument in question the 'reversed causal argument', as it reverses the order in which the causal argument from hallucination considers genuine perceptions and a certain kind of hallucination that he calls 'causally matching hallucinations'. These are hallucinations that result from the stimulation of the same brain state activated when genuinely perceiving. At the end of the previous chapter, I granted that hallucinations thus produced are IIIHs. Therefore I will use 'causally matching hallucinations' and 'IIIHs' as coextensional terms. I will

use the first one when I want to highlight the fact that they are brought about by the same brain state that brings about a perception (or when reporting Martin's position), and the latter when I need to stress the fact that they are impersonally indistinguishable through introspection from the perception they causally match.

While Robinson's argument states that whatever event happens in the veridical case, it should also happen in the case of a causally matching hallucination, the reversed causal argument argues that whatever happens in the case of hallucination should also be found in the case of perception.

Martin (2004: 54) presents the argument as follows:

(1) When S sees a pine tree at t, call this situation v, there is in S's body some complete causal condition just prior to t which determined the chance of this event of seeing occurring in v, call this condition N;

(2) It is nomologically possible that N should occur in S even if no candidate object of perception is present and conditions necessary for the occurrence of a perception are not met, and an hallucinatory experience instead occurs; call one such situation h;

(3) Where two situations involve the same proximate causal conditions, and do not differ in any non-causal conditions or the

occurrence of some kind of effect, then the chances for the occurrence of such an effect are the same in both situations;

(4) No non-causal condition required for the occurrence of the effects of N is present in h but absent in v;

(5) Whatever kind of experience occurs in h, there is the same chance of such an experience occurring in v;

(6) Hence whatever kind of experience does occur in situations like h, it is possible that such a kind of experience occurs when one is veridically perceiving. (Martin 2004. 53–54)

In this argument, as in Robinson's, Martin accepts the possibility of a hallucination brought about by the same proximal cause (what Martin here calls 'condition N') that brings about a genuine perception (premises 1 and 2). He also accepts something along the 'same proximal cause, same immediate effect' principle in the qualified form expressed by Foster, which allows that the same proximate cause can fail to produce the same immediate effect if there are differences in the non-causal background conditions necessary for the occurrence of that effect (3). The difference is that Foster imposes what Martin sees as an arbitrary limitation on the *relevant* non-causal conditions, where 'relevant' seems to have the specific function of excluding the distal object or event from playing a non-causal constitutive role (see previous chapter). If we do not restrict the relevant

conditions in this way, the presence of the object can play a non-causal role in determining the nature of perception:

Even if those objects are implicated in the causes of experience, they also figure non-causally as essential constituents. (Martin 2004: 288)

This is the reason why it is not possible, *pace* Robinson and Foster, to conclude from the stated premises that everything that can be found in perception is also to be found in hallucination: we cannot exclude without begging the question that the distal object makes a contribution to the effect by being a necessary condition for the occurrence of perception, in virtue of being a constitutive element of its phenomenal character.

However, Martin thinks that, although considerations about the brain state shared by perception and causally matching hallucinations cannot secure any conclusions about the perceptual case, they can impose some constraints on what one can say about the hallucinatory case. The constraint imposed is that whatever accounts for the phenomenal character of an IIIH has to be something that can occur in the perceptual case too (conclusion 6). In order to resist this argument, one would need to show that the *absence* of the relevant object can play a non-causal role in determining the nature of the phenomenal character of hallucination, in the same way that, in genuine perceptual cases, the *presence* of the object determines the nature of the phenomenal character of the perception. This would mean rejecting premise (4).

In order to challenge this premise, one would need to say that the presence of the object in the perceptual case (i.e. the absence of the absence of the object) would prevent the proximate case from having the effect it normally has in the case of hallucination. In turn, this would require one to consider the role that the absence of the object plays in the hallucinatory case as akin to the non-causal constitutive role that the presence of the object plays in perception. That the absence of the object could play this role is incredible. The presence of the object may have a constitutive role to play in the nature of perception and its phenomenal character, because perception is a relation, and as such it requires the existence of its relata to occur. But it doesn't seem plausible to maintain that the experience involved in hallucination metaphysically requires the absence of the object, unless one claims that hallucinating is being acquainted with the absence of that object.

Notice that the implausibility of this claim doesn't follow from the assumption that one cannot be acquainted with absence. It might well be that, in certain cases, one is perceptually acquainted with absence. For instance, Soteriou—who is a proponent of relationalism—claims that perceptual awareness of absence involves a 'positive experiential occurrence' (Soteriou 2013: 128), understood as experience of empty regions of space—in which something could be seen, but nothing is actually seen. Indeed, this awareness of absence is thought to be ubiquitous and a prerequisite to seeing the spatial boundaries of objects. Similarly, being aware of intervals of times within which something could be heard but

nothing is actually heard is crucial for experiencing things like rhythm, which involves the alternation of hearing sound and silence (Soteriou 2013: 116–129).⁷⁷

The problem is rather that the claim that the phenomenal character of a hallucination is determined by the obtaining of a relation to an absence is phenomenologically inadequate and metaphysically implausible. Hallucinating an apple is not like hearing silence or seeing an empty space. Even if it were plausible to claim that the phenomenal character of the latter is partly constituted by the obtaining of a relation to an absence (of sound, or of space-filling objects), it is not clear how the phenomenal character of the experience of seeming to see an apple can be constituted by the obtaining of a relation to an absence. There is a significant phenomenological difference between seeing an empty space and hallucinating an apple that this proposal fails to capture.

Since there doesn't seem to be any obvious and convincing way of rejecting (4), it is difficult to avoid the conclusion of the reversed causal argument. Therefore, we must accept that whatever happens in the case of hallucination is also to be found in the case of genuine perception: a common psychological effect is brought about by the activation of certain brain areas in both perception and causally matching hallucination. Notice,

⁷⁷ This view stands in contrast to the prevailing idea that we see only present objects and scenes, and thus that perception of absence is dependent on having the negative belief that something isn't the case (Dretske 1969; O' Shaughnessy 2000: 330). Soterious' account of experience of absence is dependant on his account of a spatial sensory field of visual perception and a temporal sensory field of auditory perception. These determine the sense-modality-dependent structural feature of the relation of acquaintance involved in perception and contribute to the phenomenal character of perceptual experience.

however, that this conclusion is compatible with relationalism. Saying that there a common psychological effect occurs in cases of veridical perception and causally matching hallucination doesn't prevent one from saying that, on top of the common psychological effect, something else contributes to the phenomenal character of perception. However, this imposes certain constraints on what can be said to constitute the phenomenal character of hallucination: it must be of a kind that can also be found in perception.

Now the question is: are the accounts of the phenomenal character of hallucination proposed above compatible with this constraint? Can one adopt a sense-datum, intentional, or qualia account of hallucination and coherently say that the phenomenal character of hallucination is also present in perception, while still maintaining that the phenomenal character of perception is partly determined by the obtaining of a relation to a physical mind-independent object? Martin (2004, 2006) argues that that one cannot.

He argues that, if one admits that if a given brain state is sufficient to produce an experience with a phenomenal character that is impersonally introspectively indistinguishable from a genuine perception, it is not clear why the obtaining of a psychological event is sufficient to account for the phenomenal character in the hallucinatory case, while it is not in the perceptual case. It seems that the common psychological event screens off the explanatory efficacy of the obtaining of a relation to a physical mind-independent object in accounting for the phenomenal character of perception. The result would be a merely nominal relationalism, where the obtaining of the relation to an object doesn't really contribute to the

phenomenal character of perception, which is fully explained by the common episode occurring both in perception and IIIH.

In formulating the screening off problem, Martin assumes that phenomenal characters are maximally determinate: ‘if we are intending to give the most determinate and specific account of how things are with a subject at a given time’ (Martin 2006: 371), we cannot avoid competition arising between what is solely to be found in perception (the obtaining of a relation to mind-independent mental object) and what it is to be found in both perception and hallucination, whatever that is. As we have seen, this is hardly a contentious assumption.

The screening off problem seems to apply to all the possible accounts of hallucination we have considered above. Suppose one gives an account of the phenomenal character of causally matching hallucinations in terms of the experience’s possession of certain qualitative properties (qualia). If, as per the conclusion of the reversed causal argument, the presence of an experience with certain qualia is also to be found in perception, ‘the naïve realist aspects of perception [the obtaining of a relation to mind-independent objects] [cannot] themselves shape the contours of the subject’s conscious experience’ (Martin 2004: 64). The experience’s possession of qualia exhausts the explanation of what it is like to have it, even in the case of genuine perception.

The same, it seems, holds if one claims that the phenomenal character of hallucination is determined by a representational content or by a relation to mind-dependent sense-data. If being related to mind-independent

entities or representing certain objects and properties is sufficient for having an experience with a certain phenomenal character in the case of causally matching hallucinations, the obtaining of a relation to mind-independent objects is made explanatorily redundant in the case of genuine perception.

6.3 Attempts to Avoid the Screening Off Problem

How does the screening off worry bear on the Selective Theory discussed by Price and the attempts to provide a positive account of the phenomenal character of IIIHs, and which I discussed in section 6.1? Let's first consider an account of the phenomenal character of IIIHs in terms of sense-data.

A proponent of the Selective Theory could insist that the screening off doesn't affect their view, because, by appealing to the multiple realisability of phenomenal characters, they can grant that something happening both in perception and hallucination obtains: in both cases the same phenomenal character is present. Moreover, in both cases the phenomenal character is constituted by the obtaining of a relation to a certain object. However in each case the object to which one is related, which shapes the contours of one's phenomenal character, is a different object: either a lemon or a lemon-shaped sense-datum. Insisting that the relation to a lemon-shaped sense-datum should screen off the relation to a real sense-datum would be like saying that, since it is possible for one to have an experience that is indistinguishable from a lemon, when one only sees a lemon-shaped soap, the obtaining of a perceptual relation to the lemon-shaped soap should screen off the obtaining of a relation to a real

lemon every time I see a lemon and prevent it from contributing to the phenomenal character of the lemon-perception.

This line of response seems particularly appealing if one conceives of sense-data as mind-independent entities. As we saw in the first chapter, early proponents of sense-datum theories, such as Moore, Russell, Price, and Broad, were not committed to any particular view about the nature of sense-data. The notion of sense-data is generally introduced simply to circumscribe the kind of things of which one can be perceptually aware, without further initial commitment with respect to what kind of things one can perceive (see for instance Moore 1926: 54–55). The debate among sense-datum theorists takes precisely the form of working out, through arguments, the nature of sense-data. In general early proponents of sense-data tend to think of them as non-physical entities, but also as mind-independent (see Martin 2003 and Martin 2016 for a discussion). Only later proponents of sense-data, such as Robinson and Foster, frame them as mind-dependent entities. In particular, Price presents the Selective Theory as committed to the mind-independence of sense-data. If sense-data are mind-independent objects that exist around us independently of our perception of them, the analogy between a lemon-shaped sense-datum and a lemon-shaped soap is even stronger.

However, the analogy between being related to a sense-datum in an IIIH and being related to a lemon-shaped soap in an experience of a soap that is indistinguishable from a lemon breaks down in the context of the reversed causal argument. We do not think that the occurrence of a certain

brain state is sufficient for the obtaining of a relation to a lemon-shaped soap. The obtaining of a relation to a lemon-shaped soap is due to the presence of the soap before one and one's being perceptually related to it. In virtue of the reversed causal argument, instead, the occurrence of a certain brain state should be sufficient for the occurrence of what determines the phenomenal character of experience.

If the occurrence of a certain brain state is sufficient for the obtaining of a psychological event, one might well insist that the sense-data to which one is related exist in the world around us and are ready for one to pick up every time one has one's brain stimulated in a certain way. But this is problematic. One is related to these entities only (and every time) one is in a certain brain state. This generates a paradoxical situation in which a psychological episode is accounted for in terms of the obtaining of a relation to a mind-independent object, but whether or not one is related to this object is *simply* determined by whether one is in the relevant brain state. Even if one manages to solve the metaphysical difficulties of this proposal, the screening off is unavoidable. Since it is sufficient to be in a certain brain state to be so related to a sense-datum, an awkward relation to a mind-independent object should obtain every time one is in that brain state, including when a real lemon is in front of one. But if this is the case, the relation to a lemon-shaped sense-datum should screen off the role that the relation to the real lemon is supposed to play when perceiving a real lemon.

Claiming that sense-data are mind-independent doesn't make the screening off worry disappear.⁷⁸

An intentional or a qualia account of IIIHs cannot escape the screening off either: if being in an intentional state with a certain content or having an experience with certain intrinsic sensory qualities is sufficient for the obtaining of a certain phenomenal character in the case of a hallucination caused by a certain brain state, the that same intentional state or intrinsic property should be sufficient to account for the phenomenal character of the perception brought about by the same brain state.

More recently, Johnston presented an account that allegedly doesn't fall prey to the screening off worry, despite giving a positive account of the phenomenal character of causally matching hallucination.⁷⁹ Johnson claims that:

⁷⁸This poses an additional problem: a proponent of mind-independent sense-data needs to make queer metaphysical commitments. She has to claim that the world is replete of these strange entities that are there all the time ready to be seen—as Price says, an 'infinitely multitudinous and infinitely complicated swarm of miscellaneous sensibilia' (Price 1935: 46), which might strike many as 'monstrous and incredible' (Price 1935: 46). It should be noted that Price writes off this objection as not particularly effective. According to Price, this view of sense-data doesn't really fall prey of the objection of unduly multiplying entities. The Selective Theorist only claims that they exist all the time rather than only when sensed, for 'sense-data certainly exist [and are] found in nature', 'for they are certainly sensed' (Price 1935: 46). However, we should resist Price's assumption that sense-data exist in any case, as this relies on the Phenomenal Principle and the use that proponents of sense-data make of them—something that relationalists can and should reject (see Chapter 2).

⁷⁹It is not entirely clear that Johnston aims to defend a relationalist account of genuine perceptions. He never mentions relationalism and his stated claim is to defend direct realism, which, as we saw in the first chapter, is less committal than relationalism and compatible with intentionalism. However Johnston is often considered a relationalist (Conduct 2012; Logue 2014; and Genone 2016 all dub him a 'naïve realist'). This doesn't seem so far-fetched, considering that he explicitly argues against intentionalism and claims that 'seeing is just having things and the visible features of those

Your seeing the scene before your eyes is your being visually aware of a host of spatio-temporal particulars instantiating parts of such a profile or complex of sensible qualities and relations.

(Johnston 2004: 135)

When one hallucinates, one is aware of the uninstantiated properties of the very same sensory profile:

When we hallucinate we are aware merely of the structured qualitative parts of such sensible profiles. Any case of hallucination is thus a case of ‘direct’ visual awareness of less than one would be ‘directly’ aware of in the case of seeing.

(Johnston 2004: 137)

How exactly we understand this proposal largely depends on one’s commitments regarding the possibility of uninstantiated properties. But, however we read this proposal, there is a difficulty that Johnston needs to address. It is not clear how one can be related to universals. A relational view where one is allegedly related to universals would require one to adopt a radically platonic view of universals, where these are really thought to

things disclosed to one’ (Johnston 2004: 172). Moreover, although he uses a different terminology, he seems to be committed to a substantially relationalist account of phenomenal character, whereby the phenomenal character, or ‘sensible profile’, as he calls it, of genuine perception amounts to being aware of particulars. So maybe the reason why his direct realism is, as he calls it, ‘radical’ (Johnston 2004: 114) is that it entails a commitment to relationalism.

exist and can be accessed. Although Johnston mentions this worry, he fails to give a satisfactory answer (Johnston 2004: 141).

Moreover, one may question the phenomenological adequacy of this account. It is part of the phenomenology of the experience that we seem to be presented with particulars. How can awareness of universals give rise to the experience of seemingly being presented with individuals? And how can one explain the indistinguishability of perception and hallucination if the former provides awareness of individuals and the latter only of universals? Perceptual experience presents objects as spatially located in the visual field and as occurring at the time one has the experience. Universal properties are thought to be a-temporal and a-spatial abstracts. It is not clear that one can experience a universal at a certain spatial and temporal location without it being instantiated.

Supposing we can find a solution to these problems, I can see three possible ways of understanding Johnson's proposal. None of these provides a satisfactory strategy to account for hallucination and avoid the screening off worry:

- a) According to one reading of Johnston's proposal, perception and hallucination have the very same phenomenal character, constituted by a relation to universals. In the case of hallucination, there is nothing more to the phenomenal character than this relation to universals. In the genuine perceptual case this phenomenal character is realised by being related to the objects that instantiate those universals that figure in the phenomenal character. He seems to suggest this reading when he claims

that ‘hallucination is just awareness of the visible profiles that they actually happen to instantiate’ in the corresponding veridical case (Johnston 2004: 172) and that in hallucination there is only an irreducible phenomenal character, not further explained by other factors; while in perception one has the very same phenomenal character, which happens to be realised by being related to the object. It should be clear by now that this proposal is vulnerable to the screening off worry: if this is Johnston’s proposal, he hasn’t provided any reason why what accounts for the conscious experience one has in the hallucinatory case (awareness of visible profile) should suffice to account for the conscious experience one has in the perception brought about by the same brain state.

- b) One might alternatively read Johnston as claiming that phenomenal characters are multiply realised. So the very same phenomenal character is either realised by a relation of awareness to instantiated properties (perception), or a relation to universals (IIIH). This reading is supported by Johnston’s explicit suggestion that sensible profiles are types, not tokens (Johnston 2004: 135). This option seems to fall prey to the screening off worry too. If relation to universal is enough to determine the phenomenal character of hallucination, why are relations to universals not enough for perception too?
- c) According to a third reading, experiences can have multiple phenomenal characters. A matching hallucination only has one phenomenal character, constituted by awareness of universals. Perception has the

same phenomenal character, but also has another phenomenal character, which is the prerogative of only genuine perceptions, constituted by a relation to the individual that instantiates those properties. This third reading is also supported by some textual evidence. Johnston suggests that when one hallucinates, by being aware of uninstantiated properties, one is aware of only *one part* of the sensory profile—the other part of the sensible profile being accessible only through perception, when one is aware of the particulars. He also specifies that:

[W]e are instead aware of a *proper part* of what we are aware of in the corresponding case of seeing, a sensible profile that is no more than a certain layout of qualities. (Johnston 2004: 138; emphasis added)

This suggests that hallucination only has *a partial sensible profile*, or phenomenal character; but this violates the idea that phenomenal characters are maximally determined. Perceptions do not have a unique phenomenal character, but one that is in common in hallucination and one that is unique to perception. This doesn't fall prey to the screening off worry because the worry runs on the assumption that phenomenal characters are maximally determined. However, as we have already seen, claiming that perception has more than one phenomenal character commits one to rejecting the superficiality of the phenomenal character of perception, which is unacceptable.

6.4 The Negative Account of IIIHs

It seems that none of the positive accounts of the phenomenal character of IIIHs is safe from the screening off problem. However, Martin maintains that there is one way to avoid the screening off worry. This is to adopt a negative, derivative account of the phenomenal character of IIIHs (or causally matching hallucination). According to him, this is the only way for the relationalist to meet the condition imposed by the reversed causal argument—i.e. that the same psychological event occurring in hallucination also occurs in perception—without running into the screening off problem.

Martin argues that the phenomenal character of hallucination should only be characterised in a way that is derivative of the phenomenal character of perception: hallucination seems and only seems to be constituted by the obtaining of a relation to the mind-independent object that actually constitutes the phenomenal character of the corresponding genuine perception.

The phenomenal character of hallucination is to be accounted for in terms of the obtaining of a negative epistemic condition it must be indistinguishable from the phenomenal character of a genuine perception. The phenomenal consciousness of causally matching hallucinations is provided by their being ‘essentially failure—they purport to relate us to the world while failing to do so’ (Martin 2006, p. 372); and this is the most specific thing one can say about hallucinations.

This condition is negative and epistemic because it accounts for the phenomenal character of hallucination simply in terms of what *cannot* be

known of it: when one has a causally matching hallucination, one is not in a position to know through introspection that one isn't having the kind of experience that genuinely relates one to a mind-independent object. The phenomenal character of a hallucination, then, is

nothing more than the obtaining of certain negative epistemological conditions and what follows from them: that it is not possible to know through reflection that this is not a situation of veridical perception. (Martin 2006: 373)

This meets the condition imposed by the reversed causal argument, because indiscriminability is reflexive. Thus, surely the phenomenal character of a genuine perception is indiscriminable from itself. When one has a genuine perception, one is not in a position to distinguish its phenomenal character from a phenomenal character constituted by the obtaining of a relation to a mind-independent object. The event occurring in hallucination—characterised in terms of being indistinguishable through introspection from a genuine perception—is something that also occurs in perception.

However, this doesn't generate a screening off problem. There is no risk that what accounts for the phenomenal character of hallucination will override the explanatory role that relationalists attribute to the object in accounting for the phenomenal character of genuine perceptions. This is due to the fact that the phenomenal character of hallucination lacks any positive autonomous account. Any explanation of the phenomenal character of a

causally matching hallucination is parasitic on the account given to genuine perception, which hence preserves its full explanatory role.

If this is correct, Dancy's diagnosis of the debate on disjunctivism fails to grasp an important aspect of the disjunctive account of experiences required by the relationalist. Dancy complains that disjunctivists take the indistinguishability of certain hallucinations from perception to be 'a matter of definition', while it is only 'a contingent matter' (Dancy 1995: 131), only concerning a limited sub-set of hallucinations. On the contrary, he argues, they should provide a positive account of the phenomenal character of hallucination. Moreover, he claims, disjunctivists unduly attribute a primacy to the good disjunct—perception. This primacy, according to Dancy, is simply a result of taking too seriously the way in which the bad disjunct was originally formulated by Hinton, in terms of 'seeming to see', that makes it look derivative of the good disjunct; while in fact 'there need be no suggestion that the worst case be only understood in terms of the better one' (Dancy 1995: 435). Thus, he recommends that the disjunctivist offer a positive account of hallucination and abandon any attribution of priority to perception.⁸⁰ However, *pace* Dancy, adopting those two proposals would leave the disjunctivist susceptible to the screening off problem.

⁸⁰ In his own words: 'in the standard formulation of the account, misleadingly, [...] we characterize [the second disjunct] solely by saying that it is like what it is not. Presumably, however, there may be available a more direct characterization of the second disjunct, and in a totally explicit version of the theory it would indeed be characterized in that better way. [...] [A]lthough we may be persuaded to read the disjunctive conception as holding two states, a better one and a worse one which is indistinguishable from the better one, we could equally well have read things the other way round, saying that there are two states, a worse one and a better one which is indistinguishable from the worse one. Here the apparent primacy state has vanished, leaving only its preferability, which was never in

6.5 Objections and Clarifications

A number of objections have been raised against the negative account of hallucination. Addressing these questions will help us understand what the negative account of hallucination is committed to.

6.5.1 *The Scope of the Negative Account*

Sturgeon (2006: 193; 2008: 125) and Conduct (2010) argue that the fact that in some cases of hallucinations we know that we are not perceiving puts pressure on the negative account. In the light of what we said in Chapters 4 and 5, it should be clear why this objection should not worry the relationalist in the least. Contrary to what many commentators presume (Sturgeon 2006; Siegel 2008; Hellie 2005, 2007; Conduct 2010), the negative account of hallucination is not meant to provide sufficient and necessary conditions for hallucinations in general. It is only meant to provide an account of the phenomenal character of IIIH in order to respond to the difficulties that the reversed causal argument poses for more positive accounts of hallucination available for most cases of hallucination.⁸¹

doubt. Calling the worse state “seeming to see” expressed that primacy, but was inessential to the main purpose of responding to the argument from [hallucination]’ (Dancy 1995: 435).

⁸¹ This confusion on the scope of the negative view is arguably promoted by the fact that Fish—who defends relationalism and proposes a negative view of

Nor is Martin interested in providing, through the negative epistemic condition, an account of what makes a mental occurrence a sensory experience in general. Martin says a few things that might explain why commentators have taken him to be in the business of providing necessary and sufficient conditions for some psychological episode to count as a sensory experience. He wonders how the negative condition of introspective indistinguishability—which, as we have seen, IIIH shares with genuine perceptions—‘is employed in the positive account of the notion of perceptual experience is general’ (Martin 2004: 301). However, discussing how this criterion can feature in an account of sensory experience in general doesn’t mean that it offers necessary and sufficient conditions for it. On the contrary, he specifies that this only ‘gives one sufficient conditions for having visual experience, and not yet any necessary condition’ (Martin 2006: 392). This is compatible with allowing that those hallucinations that are not introspectively indistinguishable from perception still count as conscious sensory experiences. If we understand ‘introspectively indistinguishable’ in an impersonal sense, the negative epistemic condition fails to apply for most, if not all, actual hallucinations (see Chapter 4).

This doesn’t mean that the relationalist account is incomplete until they have found a set of jointly necessary and sufficient conditions for what

hallucination that he sees as a development of and an improvement on Martin’s account—makes no distinction between causally matching hallucination (or IIIHs) and other kinds of hallucinations. Fish explicitly presents the negative criterion as providing necessary and sufficient conditions for hallucination (Fish 2009: 93–101). He also claims: ‘[T]he essence of hallucination—what distinguishes hallucinations as a class from other mental states—lies in their being indistinguishable from veridical perceptions, not in some antecedently identifiable feature of the state’ (Fish 2008: 156).

counts as a sensory experience. As a matter of fact, Martin doesn't think it is possible to find one. Renouncing the task of giving necessary and sufficient conditions for what counts as sensory experience shouldn't be seen as a limitation of relationalism, that is, as something that relationalists ought to provide, but can't.

In Chapter 2, we saw Hinton claiming that 'experience' is philosophical jargon introduced to capture genuine perception and all those mental episodes that seem like perception but are not. Relationalists tell us what perception amounts to: it is the obtaining of a psychological relation with certain mind-independent objects. However, there is no single way in which an episode can seem but fail to be an instance of perception. Hinton notices that the second disjunct in the disjunctive claim ('S sees a flash of light or only seems to see a flash of light') is itself 'implicitly disjunctive: the illusion of a flash of light is the disjunction of cases that are not, but to the subject are like, seeing a flash of light' (Hinton 1967: 218–219). Hallucinations are cases of things going wrong, but as Austin noticed, and as we considered at length in both Chapters 2 and 4, there are many ways in which things can go wrong.

Notice that Hinton indicated that hallucinations 'to the subject are like seeing a flash of light': this is significantly weaker than impersonal introspective indistinguishability. A mental occurrence can be like perception without the subject taking it to be perception. One might object that this is too general, as many non-sensory experiences are like perception—dreams, experiences under hypnosis, vivid imagining. Again,

we should stress that this is no way meant to be a definition of hallucination. A definition of hallucination that covers all hallucinations and excludes all cases of dreams, imagery, illusion, and other experiences is probably not achievable—under a relationalist paradigm or any other—for the reasons seen in Chapter 4.

Another objection to the negative account of hallucination that our previous discussion allows us to discharge easily complains that the negative condition requires the relation of indistinguishability to be asymmetrical, but ‘indiscriminability can also be regarded as symmetrical and some reasons are needed to exclude that’ (Vega-Encabo 2010: 294). But we have seen at the beginning of Chapter 5 that we have reasons to take the relation of indiscriminability between perception and hallucination, even when involving IIIHs, to be asymmetrical.

6.5.2 *The Alleged Need for a Positive Account*

A more serious challenge to the negative view of IIIHs is the following. One may worry that a negative epistemic condition specifying what *cannot* be known through introspection fails to capture the phenomenology of an IIIH. Smith claims that ‘having some specific psychological feature, not entailed by meeting the negative epistemic criterion, is necessary in order for an experience to be a hallucination’ (Smith 2008: 183).

Smith doesn’t expand much on this concern, but it seems to me that what he worries about is cashed out by Siegel (2008: 218) when she notices that the phenomenal character of an experience determines what it’s like to

have that experience, and we can type experiences on the basis of how they strike the subject introspectively. Simply saying what one cannot know about the phenomenal character of one's IIIH doesn't help us classify experiences with respect to whether they are phenomenally similar to the IIIH in question or not. When one introspectively accesses one's conscious experience, one should not only be in a negative epistemic position. One should also be in a position to know that her current experience (say, of a sausage) is not like the experience of something else (say, of a butterfly).

The proponent of the negative view of IIIHs, according to Siegel, aims to account for the subject's introspective knowledge that she is not having an experience of a butterfly derivatively from the subject's not being in a position to know that she is not having a genuine perception of a sausage:

The difference in character between your sausage-hallucination and, say, a hallucination of a butterfly is supposed to be reflected in the fact that each hallucination consists in a different fact about introspective unknowability. The fact that you can't know by introspection that you're not veridically perceiving a sausage is thus supposed to suffice to make it the case that you are not having a butterfly-experience: for example, you are not hallucinating a butterfly, and you are not having an experience with the character of veridically perceiving a butterfly. (Siegel 2008: 219)

However, Siegel argues, ‘unknowability facts’ (negative epistemic facts such as not being in a position to know that one is not perceiving a sausage) cannot ground positive epistemic facts (such as knowing that an experience is not an instance of perceiving a butterfly). For the negative epistemic situation is one I am in every time I cannot know through introspection that I am not genuinely perceiving a sausage. But the truth of this negative fact doesn’t rule out that I could be having all sorts of conscious experience, which for some reason I cannot distinguish from a perception. Notice that Siegel, like many other commentators, fails to distinguish between simple failure to discriminate and a modal notion of indiscriminability—and in particular impersonally indiscriminable. So a disjunctivist could reply that the negative condition of being in a situation that is impersonally indistinguishable from a genuine perception of a sausage is not compatible with having a perception of a butterfly.

Moreover, to respond to this worry, the relationalist may claim that in saying that there is no positive characterization of the phenomenal character of an IIIH she means to deny the idea that there are specific properties of the experience or a specific event that explains what it is like to have an IIIH. However, there is something positive we can introspectively know about the phenomenal character of IIIHs, namely that ‘they purport to relate us to the world’ (Martin 2006: 372). We cannot know through introspection that they actually ‘fail [...] to do so’. But one might contend that the negative view of hallucination fails to capture this idea. This is true, but again it should be noted that the negative epistemic

condition should be not seen as a necessary and sufficient condition for the obtaining of a hallucination (not even if we restrict it to IIIHs).

The negative condition is only meant to provide the most specific available condition for determining the phenomenal character of an IIIH. But this negative condition must go together with a positive aspect, i.e. the fact that one has an experience that just seems like a perceptual relation to an *F*. As Hinton says, a disjunctive claim of the form ‘S sees a flash of light or only seems to see a flash of light’ doesn’t need to be made true by the same occurrence, and not even by the occurrence of two different events that share certain ‘worn-on-the-sleeves’ properties. It might just be made true by the occurrence of an event of seeing a flash of light and something that just seems like that (see Chapter 2 of this thesis).

The problem with this proposal—one might object—is that although the relationalist *wants* to claim that there is positive, yet derivative, characterization of the hallucinatory experience, it is not clear that the mere impossibility of introspectively discriminating one’s situation from a genuine perception stated by Martin’s negative condition *can* provide this. What the relationalist can say in response to this relies on the picture of introspection she proposes. However, I will explain what the relationalist can say about introspection only after I have presented some other frequent objections to the negative view of hallucination; for the way we should understand introspection is crucial to answering those challenges too.

6.5.3 *Very Bad Cases and Unsophisticated Hallucinators*

Another way of insisting on the need for a positive characterization of the phenomenal character of IIIHs argues that, since the criterion provided by Martin is epistemic, it cannot account for the fact that hallucinations are sensory experiences. This objection relies on the second group of counterexamples that I mentioned in section 6.4.1. These counterexamples are meant to show that the negative epistemic criterion is too inclusive, for it is also met by mental occurrences that are not introspectively indistinguishable from perception. According to Smith:

In virtue of meeting the negative epistemic criterion very bad cases are lumped together with hallucinations, and their non-sensory character is not recognised. Alternatively, the sensory character of hallucinations is not recognised, since they are characterised in a way that applies to the non-sensory very bad cases. (Smith 2008: 186).

Smith provides the following three examples of very bad cases: dreams, hypnosis, and what he calls ‘tachiscopic cases’. These are cases in which, in an experimental setting, a subject faces a white screen upon which a tachiscope projects an image for a few milliseconds. Due to the short time of exposure to the image, subjects are inclined to believe that an image has been shown even when it has not and they are only imagining so.

Smith thinks it that it would be ‘unusual’ for these cases to be cases of sensory experiences, but he doesn’t explain why this is so. It is debatable to what extent dreams are cases of conscious experience. Even if we identify

certain criteria that distinguish between conscious sensory experiences and dreams (such as, for instance, that conscious sensory experience requires being awake), as we saw in Chapter 4 the distinction between dreams and hallucination is not clear-cut: there are at least some borderline cases for which it is difficult to decide whether they are cases of hallucinations or dream (I am thinking in particular of hypnagogic and hypnopompic experiences). Yet we may grant that paradigmatic cases of dreams should be distinguished from paradigmatic cases of hallucinations. Similarly, one might wonder whether it is clear that one doesn't hallucinate when being hypnotized, but we can also grant for the sake of argument that one might, when hypnotized, have all the relevant beliefs to put one in a position in which one is unable to tell that one is not perceiving an *F* without having the relevant conscious sensory experience. It is more controversial to claim that in the tachiscopic case one doesn't have any sensory experience, but rather only visualizes or imagines an image on the screen. And similarly to the case of dreams, the distinction between imagery and hallucination, as we saw in Chapter 4, is not so clear-cut either.

But we can concede to Smith, at least for the sake of the argument, that when one falsely reports seeing an image in the tachiscopic case, one doesn't have a conscious sensory experience. Smith's objection is that, although these are not cases of sensory experience, they are cases that are indistinguishable from a perception. Thus, the negative epistemic condition that is supposed to determine the phenomenal character of IIIHs is met by all these non-sensory cases too. And thus, by failing to distinguish IIIHs

from cases where it only epistemically seems to one as if one is perceiving, the relationalist fails to capture the sensory nature of hallucinations.

A first response on the part of the relationalist to these alleged counterexamples is that the notion of indistinguishability in play in the negative epistemic condition is impersonal, while cases of experience under hypnosis, dreams, or any other situation in which one cannot tell that one is not perceiving are not impersonally indiscriminable through introspection. We have seen that in order to establish that an experience is impersonally indiscriminability from perception, one should rule out that interfering conditions prevent one from noticing that one's experience is actually phenomenally different from genuine cases of perception. Being hypnotized and being asleep are precisely cases where conditions that could prevent one from distinguishing one's experience from genuinely sensory experiences are present. It is also reasonable to suppose that the short time span of the stimulus (or the apparent presence of the stimulus) in the tachisopic case interferes with one's capacity to introspectively discriminate. Thus, the negative condition for IIIH doesn't apply to these cases.

At this point, in the current dialectic, opponents of the negative view of hallucination criticize either the notion of impersonal indiscriminability or its use with respect to experiences and knowledge through introspection (Smith 2008; Siegel 2008; Farkas 2006). However, we have already seen, in Chapter 3, that the proponent of the argument from hallucination needs to introduce this notion in the first place in order to secure the spreading step of the argument from hallucination. Only if we understand

indistinguishability in an impersonal way do relationalists face difficulty in accounting for the indistinguishability of certain hallucinations (IIHs) without assuming the sameness of phenomenal characters and yet without violating the superficiality condition. The problem for the relationalist is not that the notion of impersonal indistinguishability is incoherent. If the notion is incoherent, the very argument from hallucination, and the screening off that leads relationalists to give a negative account of IIHs, doesn't even get off the ground. One can just say that perception and hallucination have different phenomenal characters and that whatever account(s) one give to hallucination doesn't affect the relationalist account of perception. The problem for the relationalist is that it is not clear that she can give a coherent account of IIHs that doesn't understand impersonal indistinguishability through introspection in terms of the sameness of phenomenal characters. I have to postpone an answer to this question until the end of this chapter, after having reviewed other common objections to the negative account of hallucination.

However, as Martin (2006) himself recognizes, invoking an impersonal notion of indiscriminability as part of the negative epistemic condition doesn't suffice to respond to the worry presented by Smith. We can better see the extent of the problem faced by a proponent of the negative view of hallucination if we consider a phenomenal zombie duplicate. A phenomenal zombie would have the same cognitive and behavioral response as someone perceiving an apple, without having any sensory experience. In this case, the relationalist cannot invoke the distinction between the

impersonal and personal notions of indistinguishability, as we can imagine a case of IIIH where the brain of the phenomenal zombie is activated in exactly the same way as in the genuine perception of a conscious subject, but the upshot of this brain state is the possession of certain beliefs without any conscious sensory experience. It seems that, in light of the negative epistemic condition, someone having an IIIH would be in the same situation as the phenomenal zombie (this objection is also developed by Siegel 2008 and further discussed by Sturgeon 2008). Thus, the negative account of hallucination fails to explain how IIIHs are genuine sensory experiences.

It seems, thus, that, in the light of the negative epistemic account, hallucinations can lack a phenomenal character, although they seem to have one. Before considering what a relationalist could say to this objection, I want to consider another objection, which helps us to see where the difficulty lies for the proponent of the negative view of hallucination and what she can to respond. As is the case for the objection I considered in the previous sub-section, the key to a satisfactory answer lies in the way relationalists should account for introspection.

The further objection is once again discussed by Siegel (2008: 213). Siegel argues that the negative epistemic condition fails to account for the possibility of cognitively unsophisticated creatures such as dogs having hallucinations. Suppose a dog has a hallucination of a sausage. The dog lacks the cognitive capacities required to make judgments, and in particular she lacks a notion of experience, so she is unable to formulate the introspective judgment that her experience is not an experience of a

butterfly. The dog's hallucination, then, would be indistinguishable from an experience of a butterfly (or a tuna, a submarine, or anything else). Each kind of hallucination had by the dog would count as indistinguishable from any possible genuine perception.

One might try to respond to this challenge by invoking two precisions I made earlier. The first is the idea that there is indeed something the dog is positively aware of through introspection when having an IIIH: she has an experience that seems just like that of a sausage (but she is not aware that it is not). However, since the dog doesn't have the capacity to reflect on her experience, she cannot have any introspective judgment about her experience whatsoever. But this doesn't mean that her experience doesn't have any phenomenology to it: it just seems to her as if she were genuinely perceiving a sausage. The second precision again appeals to impersonal indistinguishability: we can make suppositions regarding the experience the dog undergoes, even if she doesn't have the relevant introspective capacities. The condition of being impersonally introspectively indistinguishable from the perception of a sausage can apply to the dog's hallucination even if she lacks any introspective capacity, in the same way as a patch of colour can count as perceptually indistinguishable from red even if the subject for which the patch of colour is indiscriminable from red is blind.

However, this line of response is not satisfactory. Critics (in particular Siegel 2008) have argued that the appeal to introspective indistinguishability may at best explain what we take, from a third-person

perceptive, the experience of the dog to be like. It doesn't illuminate what it is like for the dog to have an IIIH.

6.6 The Phenomenal Character as a Ghost

Fish (2009) responds to some of these and other worries⁸² by biting the bullet and agreeing that IIIHs (or, for Fish, all hallucinations, since he doesn't distinguish between different notions of indistinguishability and aims to apply the negative account to all hallucinations)⁸³ lack any phenomenal character whatsoever, while granting that this doesn't deprive the hallucinator's state from having a proper phenomenology. Fish revisits the negative condition for IIIHs introduced by Martin in terms of the cognitive and behavioural effects that the hallucination shares with a corresponding perception. A hallucination is indistinguishable from a perception inasmuch as it produces the same sets of beliefs and behaviours that perception produces:

⁸² Among the motivations that lead Fish to revisit Martin's negative account of (causally matching) hallucinations is the dog problem formulated by Siegel (2008: 213). However, what he seems to take as more pressing motivations, and with respect to which his account is allegedly a significant improvement over Martin's, are an objection raised by Sturgeon (2006: 364–365) regarding the difficulty of understanding Martin's clause 'through reflection' (which he uses in the same sense as I have here used 'through introspection') in the formulation of the negative condition, and an objection raised by Farkas (2006) in understanding the notion of impersonal indiscriminability. However, we should not be moved by these two motivations, for we have already seen that (1) it is indeed possible to make sense of the idea of impersonal indiscriminability through introspection and (2) it is the proponent of the argument from hallucination in the first place who needs this notion.

⁸³ Clearly this exposes Fish to the objection raised in section 6.4.1 that didn't apply to Martin's account.

We can interpret the requirements placed on hallucination [by the screening off problem] as insisting (1) that a hallucination must not fail to produce any judgments or beliefs that would have been produced by a veridical perception of the relevant kind and (2) that a hallucination must not produce any beliefs or judgments that would not have been produced by a veridical perception of the relevant kind. With this in mind, we can see that a hallucination will be indiscriminable from a veridical perception of a certain kind if and only if it produces the same beliefs or judgments that a veridical perception of that kind would have produced. (Fish 2009: 94)

We can reconstruct Fish's reasoning as follows. The phenomenal character of perception is determined by the obtaining of a relation to the object one seems to perceive, but in hallucination there is no suitable object to which one can be related. Thus, we must conclude that hallucination doesn't have any phenomenal character whatsoever, for phenomenal characters are the kind of things that require the obtaining of a relation to the object one seems to see.

This proposal unsurprisingly encountered severe resistance and was seen as definitive confirmation of how far relationalists can go in their denial of obvious intuitions about phenomenal characters. It strikes many as simply incredible that (certain) hallucinations don't have any phenomenal character at all, as Knight eloquently voices:

[On this view] we can believe we are having an experience with a specific character when that experience has no character at all! On this view we are not aware of anything when we hallucinate; we are only the state of seeming to have such awareness. Such empty seemings are strange, and many philosophers have found disjunctivism so conceived a bitter if not impossible pill to swallow. (Knight 2013: 3)⁸⁴

Fish would probably respond to this by claiming that we have no compelling reason to assume that hallucinations have phenomenal character. What we have to explain is the fact that they are indistinguishable from perception. But this, he would argue, can be perfectly explained in terms of the cognitive and behavioural effects that the two share. Having the same cognitive effects generates the impression in the subject that she is having an experience with phenomenal character: ‘The phenomenal character of hallucination is simply a ghost generated by my belief that I am seeing something’ (Fish 2009: 98).

How can this possibly be a response to the objection raised by Smith and Siegel? This seems only to aggravate the position of the disjunctivist by making it very clear that the negative account of hallucination doesn’t ensure the presence of a phenomenally conscious occurrence and that the

⁸⁴ It should be noted that Knight mistakenly attributes this view to Martin. This indicates, I believe, a widespread tendency among commentators to assume that the only coherent way to spell out the negative view of hallucination is to adopt Fish’s option and deny that IIIHs have phenomenal character altogether.

costs of this proposal outweigh the benefits of preserving relationalism (see Pautz 2013).

Fish insists that denying that hallucinations have phenomenal characters doesn't prevent them from being phenomenally conscious. What we need to account for, in order to secure the intuition that hallucinations are phenomenally conscious, and in order to explain their indistinguishability from perception, is the fact that they have a 'felt reality', a term that he borrows from Siegel (2008: 213) without much explanation, but which I assume roughly corresponds to what in Chapter 4 I called a 'sense of reality'. And this, Fish maintains, is explained in turn by the fact that he has certain beliefs:

[I]f subjects believe or judge that a state of a certain kind is present, then it will be for the subjects as though they are in such a state, even in situations in which they are not. This explains how my definition of hallucination is able to account for the fact that hallucinations have a felt reality, even though it also insists that the hallucination itself lacks phenomenal character. (Fish 2009: 98)

This should be enough, according to Fish, to silence the objection that the negative view of hallucination makes a subject of hallucination equivalent to a phenomenal zombie. Moreover, he notices that there is something it is like for the subject who has the hallucination in virtue of the fact that there is something it is like to have the range of beliefs one has

(Fish 2009: 99, n 19). It seems, then, that the apparent sensory conscious phenomenology is due to cognitive phenomenology: the subject of hallucination is the victim of a meta-cognitive error, as she takes a mental occurrence with one kind of phenomenology (a set of beliefs) for a mental occurrence with another kind of phenomenology (a sensory perception).⁸⁵

Fish also maintains that his proposal offers a suitable response to the ‘dog problem’: although the dog doesn’t have the capacity to formulate beliefs, her experience produces certain behavioural effects, and these are enough to explain in which sense her hallucination can be indistinguishable from perception. Both when she sees a sausage and when she hallucinates one, she starts drooling, while she doesn’t do so when she sees a butterfly.

Denying that hallucinations have any phenomenal character might also be seen to offer a neat answer to the challenge that the superficiality constraint poses to the relationalist. A relationalist can explain why an IIIH doesn’t have the same phenomenal character as perception without committing to the claim that the phenomenal character of hallucination is hidden to introspection and thus not superficial with respect to it. Recall the argument from superficiality I introduced in Chapter 3:

1. Some hallucinations are IIIH;
2. An IIIH is phenomenally conscious;

⁸⁵ See Dokic and Martin 2012 for a further attempt to develop, drawing on empirical data, an account of hallucination in terms of meta-cognitive mistakes. According to their proposal, however, perception and hallucination share the same affective phenomenology (which, according to them, constitute the felt reality of experiences), rather than the same cognitive effects. Although this account is very plausible in many cases of hallucination, it is less plausible in the case of IIIHs.

3. (from 2) IIIHs have a phenomenal character;
4. Phenomenal characters are superficial.
5. If phenomenal characters are superficial, the phenomenal character of an experience is accessible (at least in principle) through introspection (superficiality constraint);
6. Relationalism claims that the phenomenal character of an IIIH is different from the phenomenal character of a perception;
7. (from 6) For the relationalist, the phenomenal character of an IIIH is different from what it introspectively seems to be: its real nature is hidden to introspection;
8. (from 2 and 7) Relationalism is committed to the claim that the phenomenal character of an IIIH is not superficial;
9. (from 4 and 8) Relationalism is untenable because it denies the superficiality of phenomenal characters.

Fish can block the argument by denying premise (3): he is committed to claiming that the fact that experience is phenomenally conscious doesn't entail that it has a phenomenal character.

Despite this advantage, however, there are many reasons for being unsatisfied with accounting for the indiscriminability of hallucination from perception in terms of the sameness of cognitive and behavioural effects. One might still be persuaded that the phenomenology of a conscious experience can be fully captured by a set of behaviours and beliefs (see Vega-Encabo 2010: 190) and that accounting for how an IIIH strikes us in terms of a meta-cognitive mistake consisting in taking cognitive

phenomenology for sensory phenomenology is phenomenologically accurate.

There is a further problem with this proposal that becomes evident only once we have clear in mind that the negative account of hallucination is supposed to apply only to IIIHs. It seems that if the genuine perception has a sensory phenomenology, while the corresponding IIIH has only a cognitive phenomenology that one takes to be sensory, the two episodes are not impersonally indistinguishable through introspection in the way I discussed in Chapter 3: a meta-cognitive error seems like the kind of error that can be explained in terms of something's interfering with one's introspective capacities. Thus, Fish's proposal is off the mark and fails to properly address the argument from hallucination and the argument from superficiality. It may be a valid account for (some of) the ordinary actual hallucination, but it doesn't do anything to counter arguments that focus on philosopher's hallucinations, i.e. IIIHs.

Moreover, as Pautz (2013) notices, the viability of this proposal depends on how satisfactorily one can counter traditional arguments against Rosenthal's 'higher-order thought theory' of consciousness, from which the idea that consciousness of hallucination is a by-product of cognitive attitudes derives. Counterexamples in which one has a hallucination but no cognitive and behavioural effects (as is the case for very unsophisticated animals) have been proposed (Siegel 2008). Martin (2013) notices that one consequence of Fish's proposal is that it is impossible for me to have the

same hallucination as a dog, since the dog will never be in a position to have certain beliefs.

However, here I want to press another objection that is more relevant to the concerns I am dealing with in this chapter. Martin (2013) argues that Fish is not immune to the screening off problem, contrary to what Fish (2009: 109–111) claims. Fish acknowledges that it might seem that, since higher-order beliefs suffice to furnish an IIIH with felt reality despite the lack of phenomenal character, they should suffice to furnish the corresponding genuine perception with felt reality too. Yet he contends that the property of having certain cognitive effects is not subject to the screening off worry, because it is an ‘inherited’ property, which depends upon the perception’s possession of phenomenal character:

A hallucination has a felt reality—despite the fact that the hallucination itself lacks a phenomenal character—because the hallucinator falsely believes that he is enjoying a veridical perception. But having such a belief—believing that one’s current mental event is a veridical perception of a certain kind—can be sufficient to account for the felt reality of such an event only if there is already a felt reality to the event that the subject judges himself to be in. Hallucinations have a felt reality because, when we hallucinate, we take ourselves to be perceiving, and perceptions themselves have a felt reality. And, of course, the explanation of the felt reality of veridical perception cannot be an explanation of the same kind or we

would be locked in a regress. So, if veridical perceptual experiences did not have phenomenal character in the first place—phenomenal character that makes it the case that there is something it is like to be in such states—then the relational property of being indiscriminable from such an experience would not explain why we think that there is something it is like to be in a state that instantiates that relational property. (Fish 2009: 109–110)

As Martin (2013) notes, if the felt reality of hallucination is the result of the hallucinator's beliefs and behaviours, as Fish (2009: 98) explicitly claims in the passage quoted above (the fact that 'subjects believe or judge that a state of a certain kind is present, [...] explains how my definition of hallucination is able to account for the fact that hallucinations have a felt reality'), then, since the very same beliefs are also present in a perception, these beliefs and behaviours should be enough to give rise to the felt reality of perception too. The only difference between the perceptual and the hallucinatory case is that in one case the beliefs are true and the behaviour appropriate. But this doesn't make the false beliefs occurring in the hallucinatory case derivative of the beliefs one has when perceiving. Therefore, if in both perception and hallucination there is felt reality, and this can be fully explained in the case of hallucination in terms of the subject's possession of certain beliefs, Fish has done nothing to explain why the same beliefs shouldn't suffice to explain the perception's possession of

felt reality—thus screening off the alleged explanatory power of a positive phenomenal character constituted by the relation to certain objects.

However, here Fish insists that the felt reality of perception is not provided by a set of beliefs that are the upshot of experience; rather it is a feature of the conscious experience itself: presumably what accounts for it is the fact that perception has a phenomenal character realised by the obtaining of a relation to certain objects. But the fact remains that hallucination also has a felt reality, and in this it is impersonally indistinguishable from the corresponding genuine perception—and this is entirely explained by the presence of certain beliefs that are also present in perception. But then it still remains the case that the beliefs can explain why one takes oneself to be perceiving and why one's experience can seemingly have a felt reality, which should still screen off any explanatory role from the positive felt reality of perception, however this is accounted for. Fish claims that the felt reality of perception cannot be explained in the same way as the felt reality of hallucination (i.e. in terms of the subject's possession of certain beliefs), or else 'we would be locked in a regress'. But he doesn't offer any reason why it would be a regress to explain the felt reality of perception in terms of the presence of a set of beliefs—the very same beliefs one has in an IIIH, only in this case, true. The fact that in one situation one holds a false belief doesn't make one's situation explanatorily parasitical on a situation in which the belief turns out to be true.

The negative account of hallucination offered by Fish is negative in the sense that it is committed to claiming that hallucination has no

phenomenal character whatsoever. But the seeming presence of a phenomenal character is characterised in terms of the obtaining of positive conditions shared with the perceptual case. This deprives Fish of any right to claim that his account is not susceptible to the screening off problem. For what puts the account offered by Martin in a position to avoid the screening off problem turns on characterising the phenomenal character of hallucination merely in terms of the subject's inability to distinguish through introspection her current situation from a genuine perception. Thus, it seems that the relationalist should provide an account of impersonal indistinguishability through introspection, which is not reduced to the possession of certain beliefs and behaviours.

It remains to be seen whether such an account can face the objections I have introduced earlier and can account for the impersonal indiscriminability of IIIHs without violating the superficiality condition. I have anticipated that a response to the challenges made by commentators such as Smith and Siegel essentially turns on a disagreement about the way in which we should understand introspection and its connection with phenomenal characters. In the next section I will introduce this response and discuss how it can help to tackle those objections. I will then move on to discussing how the negative account of hallucination can help us understand how relationalism can accommodate the intuition that phenomenal characters are superficial.

6.7 Introspecting

Martin (2006) argues that one can properly address the objections raised above, in particular the dog and zombie problems, in the context of a discussion of the self-awareness of sensory experience. He argues that the opponent's claim that the negative epistemic condition can suffice to account for the phenomenal character of IIIHs and that one should instead appeal to positive sensory features

is forced upon us [...] if we accept a certain picture of the relation between phenomenal consciousness on the one hand, and self-awareness of our cognitive states of response to phenomenal consciousness on the other. That is, one may conceive that the facts about phenomenal consciousness are fixed independently of whether a subject has any perspective on his or her own conscious states and is thereby self-aware of them. (Martin 2006: 377)

It is tempting to think of the phenomenal character of one's experience as something one becomes aware of independently of the higher-order introspective perspective one has on it. According to this picture, in making introspective judgments about the phenomenal character of an experience, one acquires knowledge of something we are already aware of independently of those judgments. We have already considered this view multiple times in this thesis. We saw in Chapter 1 how this picture supports the idea that the phenomenal character of experience should be

characterised in terms of qualitative intrinsic properties (qualia). We also saw in Chapter 3 how this view of self-awareness motivates a restricted version of infallibilism: independently of our making introspective judgments about them, we are pre-judgmentally, directly aware of our experiences.

It seems to me that this picture of the relation between phenomenal character and one's awareness of it is motivated by the idea that one cannot be perceptually aware of the object of one's experience if one is not aware of the phenomenal character of that experience. It is the phenomenal character of the experience that reveals what one is perceptually aware of. This assumption forces on us the idea that we must be aware of the phenomenal character of our experiences independently of higher-order introspective judgments. For very often one is not in a position to formulate higher order introspective judgments on the phenomenal character of one's experience, yet one is perceptually aware of things. According to this picture, our perceptual awareness of the things around us depends on our awareness of the phenomenal character of the conscious experience we have. But do we have reasons to embrace such an assumption?

Such an assumption implies that one is aware of the phenomenal character of one's experience prior to and independently of one's perceptual awareness of the objects of perception. But the discussion of the transparency of experience in Chapter 1 suggests the contrary. We have seen that, although we may become aware of features of the conscious experience itself, when we direct our introspective attention towards it, we

are aware of those features only in virtue of being aware of the properties of the object we perceive (or seem to perceive). We do not seem to be in a position to engage in the introspective activity of attending to the phenomenal character of our own perceptual experience without attending to the objects of which we are perceptually aware. As Soteriou (2013) notices, this suggests that the activity of introspectively attending to one's experience depends on the activity of perceptually attending to and looking at the objects with which one is perceptually presented:

In the case of veridical visual perception it appears to be impossible to comply with the following request: introspect those properties that determine what your experience is an experience as of, but make sure you don't look at anything.
(Soteriou 2013: 193)

What distinguishes looking at something from introspectively attending to the experience one is undergoing, Soteriou argues, are the aims of the outcomes of that activity. When introspecting, one aims to acquire knowledge not only about what is in front of one, but also about how things seem to one to be perceptually. And one does acquire beliefs about one's experience and how this strikes one, rather than beliefs about what is there to be seen. When one introspects one looks at what one is perceptually presented with 'with the intention of discovering something about the experience one is undergoing' (Soteriou 2013: 196). It is in this sense that we should understand Martin's claims that 'a subject's perspective on his or

her own experience is not distinct from their perspective on the world' (Martin 2006: 404). Notice that, in order to be introspectively aware of the phenomenal character of one's experience by intentionally looking at the objects of perception in order find out something about the features of experience itself, a certain cognitive sophistication is required: one needs to have at least an implicit theory of mind and a concept of experience in order to switch one's attention from what one is perceptually aware of to one's experience. Whether this introspective activity is judgmental is an open question, but it seems that one can in principle perform this attentional switch without formulating judgments. However, it remains the case that creatures that do not have a concept of experience lack the capacity to introspect their experience.

But what happens when one hallucinates? When one hallucinates one is not looking at anything. However, one behaves as if one is looking at something and introspecting one's experience, and when one introspects one's hallucinatory experience, when one seems to be looking at those objects and features, one seems to see with the intention of knowing something about one's experience, although one is not looking at them, because looking at something requires one to maintain cognitive contact with those objects—which in turn requires that those objects be present. This explains why, as the disjunctivist maintains, when one hallucinates, one is not only wrong about the access the subject seems to have to the surrounding environment. One is also wrong with respect to the access one seems to have to one's own experience. One seems to be in a position in

which one can direct one's attention towards the object with which one seems to be in cognitive contact, with the aim of gaining knowledge about what the experience of these objects is like. This kind of introspective access to the phenomenal character of one's experience seems to be the kind of introspective access that requires one to look at the objects perceived, which is made possible by perceiving those objects. As Soteriou claims:

[I]n the case of successful perception the subject has available to her the kind of introspective access to her experience that she seems to have, because she has available to her the kind of perceptual access to objects of experience that she seems to have. (Soteriou 2013: 205)

By contrast, when one hallucinates, one doesn't have the kind of introspective access one seems to have to one's experience because one doesn't have the perceptual access to one's experience that one seems to have. It is in this sense that we should understand the idea that the phenomenal character of a hallucination is derivative of perception: it is because one seems to have a perception, and thereby a perspective on the world, that one can engage in that activity of introspectively directing one's attention to the objects one seems to perceive with the aim of finding out something about the features of one's own experience itself. But the access one has to one's own experience is misleading because one doesn't have the perceptual access to the world one seems to have. When one introspects one's own hallucinatory experience, one *seems to have* the kind of

introspective access that one has when one genuinely perceives something, when one has a perspective on the world. But what one comes to know, then, about the phenomenal character of one's hallucination, is just that it is indistinguishable from the phenomenal character one can introspectively discover when one actually perceives something.

Martin suggests that the deeper disagreement between disjunctivists and their opponents lies in a disagreement about the relation between phenomenal character and introspective access to it. But, if my diagnosis is correct, there seems to be an even deeper disagreement concerning the direction of the explanation between having a perceptual perspective on the world and having an introspective perspective of one's conscious experience. The opponent of the relationalist thinks that one needs some kind of awareness of the phenomenal character of one's experience in order to have a perceptual perspective on the world. Drawing on considerations about the transparency of experience, the relationalist maintains that the subject's introspective perspective on her own experience depends on her perceptual perspective on the world.

6.8 Response to Critics and How to Preserve Superficiality

How does this picture of introspection bear on the objections that have been raised to the negative view of hallucination that I considered above? And how does it bear on the way relationalists can try to accommodate the superficiality of phenomenal characters?

The first criticism I considered insisted that the negative condition doesn't suffice to capture the sensory conscious aspect of experience. Martin points out that this idea is motivated by a view of the relation between phenomenal character and introspection whereby introspecting amounts to tracking and taking notice of a phenomenal character that exists and is somehow present to us independently of our introspective perspective on it. The idea underwriting this picture is that one cannot count as having a conscious experience if one is not aware, independently of one's higher-order introspective capacities, of what it is like to have one's experience, i.e. of one's phenomenal character. The relationalist is committed to denying both claims. One's access to the phenomenal character of one's experience is not required in order to have a sensory experience. On the contrary, one becomes introspectively aware of one's conscious experience only in virtue of becoming perceptually aware of certain objects. The phenomenal character comes to be present to one's awareness only in virtue of one's being perceptually aware of certain objects. There is no immediate introspective awareness of phenomenal character of the type proposed by Giustina and Kriegel (forthcoming) (see Chapter 3 for a discussion of their proposal).

The proposed account of introspection also helps to shed light on the suggestion I made earlier in response to an objection directed by Smith and Siegel to the negative account of hallucination. I suggested that the epistemic negative condition contains a positive element to it that helps determine the phenomenal character of the IIIH, in such a way that it

doesn't turn out to have the same phenomenal character as all those experiences that one cannot introspectively discriminate from it. This positive, but yet derivative element is the fact that, in an IIIH, one seems to perceive an *F*. We have seen that one could object that the negative epistemic condition in terms of impersonal indiscriminability doesn't deliver this positive element. Now we can see why it does. The negative condition imposes a condition about the possession and awareness of the phenomenal character of an IIIH. But introspective access to the phenomenal character of one's experience depends on the perceptual access that one has or seems to have on the world. So awareness of the phenomenal character of an IIIH (accounted for in terms of being in a situation that is introspectively distinguishable from one's perception) depends on the perspective one seems to have on the world (the fact that one seems to see *F*).

These considerations help us understand what a relationalist should say to the dog objection too. Since the dog presumably lacks a theory of mind and a concept of experience, it is not capable of introspecting on its experience. Therefore, it doesn't have introspective access to the phenomenal character of its experience: it cannot step back from attending the environment and reflect on the features of its own experience. However, 'how we conceive of their experience is such that, were a creature so to experience and be self-conscious, they would thereby be able to articulate judgments just as we in fact do' (Martin 2006: 395). We can attribute the notion of impersonal introspective indistinguishability because, although

such creatures cannot introspect, we conceive of their experience such that it would have a certain introspectible phenomenal character, were the creature to be equipped with introspective capacities.

The objection to this response was that this still doesn't specify what kind of experience the dog has. But this objection relies on the assumption that one needs introspective awareness on the phenomenal character of one's experience in order to have a perceptual perspective on the world. I have argued that this is not the case. On the contrary, one needs a perspective on the world in order to exercise introspective capacities. Moreover, I have argued that having a hallucination amounts to seeming to have a perceptual perspective on the world. Seeming to see to be perceptually related to a sausage is what makes the dog's hallucination impersonally indistinguishable from the perception of a sausage.

Now that I have briefly addressed some of the familiar objections to the negative view of hallucination, it is time to consider how this proposal can help tackle the challenge I set in this thesis. Does the negative view of hallucination equip the relationalist with the means to account for the impersonal introspective indistinguishability of perception without assuming that they have the same phenomenal character and yet without violating the superficiality condition?

If the phenomenal character of an IIIH is only determined derivatively from the phenomenal character of hallucination, one who denies that a perception and an IIIH share the same phenomenal character is not committed to the claim that the real nature of the phenomenal character

of a hallucination is hidden to introspection. That relationalism is committed to the claim that the real nature of the phenomenal character of hallucination is hidden was premise (7) in the argument from superficiality. This was presented as something that follows from (6): the fact that relationalism claims that the phenomenal character of an IIIH is different from the phenomenal character of a perception. (7) seems to follow from (6) because if no difference between phenomenal characters is detectible through introspection—and ex hypothesis no difference from perception is detectable in the case of a hallucination that is *impersonally* indistinguishable through introspection from a perception—but yet the two experiences are claimed to be different with respect to their phenomenal character, it seems that the only explanation is that what differentiates the phenomenal character of the IIIH from that of perception is something that goes beyond introspection.

The negative view of hallucination helps us see that (7) doesn't need to follow from (6). The relationalist can claim that the IIIH doesn't have the same phenomenal character as a perception not because it has some non-introspectible property that distinguishes it from the phenomenal character of perception. The phenomenal character of an IIIH is different from that of a perception because it doesn't amount to the obtaining of a psychological relation to the objects one seem to perceive. However, it doesn't have a hidden nature: all there is to it is the fact the experience cannot be distinguished from a situation in which one is so related to the objects one

seems to see. This brute indistinguishability fact is accessible through introspection.

If the relationalist is not committed to violating the superficiality condition while denying that a perception and a corresponding IIIH have the same phenomenal character, the proponent of the argument from hallucination loses the motivation she had to support its crucial premise: IND. Thus, the argument from hallucination fails to show that relationalism is untenable.

6.9 Conclusions

In this chapter I first considered a strategy to resist the argument from hallucination. The relationalist could appeal to the principle of the multiple realisability of phenomenal characters and claim that the same phenomenal character is present in a perception and a corresponding IIIH, albeit in each case realised on different bases: in one case a relation to the object one seems to perceive, in the other case something else—a relation to sense-data, an intentional content, qualia...

However, following Martin, I argued that, if one embraces this strategy, whatever account of the IIIH one proposes is liable to screen off the explanatory role that the relationalist attributes to the objects of a perception that causally matches the IIIH. Therefore, Martin suggests, the relationalist should seek to account for the phenomenal character of hallucination in purely negative terms: all there is to the phenomenal

character of an IIIH is the fact that it is impersonally indistinguishable through introspection from perception.

I considered some of the most well-known objections to this account of hallucination. Most of these objections seek to argue, in different ways, that the negative epistemic condition doesn't suffice to account for the fact that hallucinations are sensory conscious experiences. I then considered a revision of the negative account proposed by Fish, who seeks to account for the indistinguishability in terms of sameness of cognitive effects in order to counter some of the objections to Martin's negative account. I argued that accounting for the indistinguishability in terms of the sameness of cognitive effects, as Fish advocates, doesn't offer a valid answer to many of these objections, but most importantly is subject to the screening off worry.

I argued that we should understand the disagreement over the negative view of hallucination against the background of a deeper disagreement about (1) the relation between phenomenal character and introspection and (2) the relation between having a perceptual perspective on the world and having an introspective perspective on one's experience. Opponents of the negative view assume that (1) one can be immediately aware of the phenomenal character of one's experience and that the presence and pre-judgmental awareness of this phenomenal character is prior to and independent of the higher-order introspective access one has to them. They are led to think so because they believe that (2) being perceptually aware of something requires one to be (in an immediate way) introspectively aware of the phenomenal character of one's experience. On

the contrary, drawing on the transparency of experience, the relationalist argues that (1) the phenomenal character doesn't exist independently of one's introspective awareness of it and that (2) having introspective access to our experience depends on our perceptual access to the world. I then discussed how the picture of introspection that I recommended can help counter some familiar objections to the negative account of hallucination.

Finally, I argued that the negative view of hallucination can provide a response to the challenge posed to relationalism by the superficiality of phenomenal characters. If the phenomenal character of an IIIH is only determined by its being impersonally indistinguishable through introspection from a perception, the relationalist doesn't commit to claiming that the nature of the phenomenal character of hallucination is hidden to introspection. All there is to the phenomenal character of an IIIH is accessible through introspection: it is just the fact that one cannot introspectively distinguish one's situation from that of seeing something. But this is clearly different from the phenomenal character of perception, which is determined by a relation to certain objects—a phenomenal character that is clearly not present in hallucination.

As we have seen in Chapter 3, the intuition that the phenomenal characters are superficial, together with the suspicion that it is not possible to deny IND without committing to denying the superficiality of the phenomenal character of hallucination, is the only valid motivation for IND, the crucial premise in the argument from hallucination. Thus, by explaining how the relationalist can coherently claim that IIIH doesn't have the same

phenomenal character as the corresponding experience, I complete my defence of relationalism in the face of the argument from hallucination.

The long journey leading through this vindication of relationalism required us to dispel several widespread misconceptions that hinder the debate in relationalism and the argument from hallucination.

First of all, the relationalist commitment that the phenomenal character of perception is partly constituted by the objects one seems to perceive is often regarded with suspicion. It is believed that the phenomenal character of experiences cannot incorporate objects, as these go beyond the scope of what the notion of phenomenal character is designed to capture: the introspectively accessible aspect of experience, which by definition abstracts away from positing the existence of objects around us (or at least this is what some theorists suppose to be the case). I argued in Chapter 1 that nothing in the notion of phenomenal character rules out the position adopted by relationalism: on the contrary, the transparency of experience and its manifest relationality support such an account of perception. In the light of the model of introspection I recommended in this chapter, whereby having an introspective perspective on the phenomenal character of one's experience depends on having a perceptual perspective on the world, it becomes clearer why it is misleading to think that the phenomenal character of experience should make possible to abstract away from the objects one perceives, in a way that precludes a relationalist account.

Another source of confusion in the current debate concerns the nature and structure of the argument that is meant to threaten relationalism.

Although it is widely accepted that the argument from hallucination poses problems to relationalism, it is not always clear how it does so. Recent debates have focused on specific versions of the argument, such as the causal argument from hallucination. This argument originates from the assumption that hallucinations that causally match perceptions are possible. It is believed to be possible to reproduce the neural state one is in when perceiving an object, in the absence of that object. In virtue of the undeniable link between the brain and the mind, it is argued, we should conclude that exactly the same brain activity would produce exactly the same experience one has when one perceives, which hence must not depend upon the existence of the object.

I argued in Chapter 5 that this argument fails, because it relies on the assumption that brain activity of a certain kind is sufficient for the occurrence of a conscious experience of a certain kind. But this assumption can be motivated only by a commitment to either the local supervenience principle, or the ‘same proximate cause, same immediate effect principle’, both of which beg the question against relationalism and face several objections. The most that considerations about the possibility of causally matching hallucinations can do is to impose upon the relationalist the adoption of a negative account of hallucination. However, this is not a problem for relationalism. The negative view of hallucination is often depicted as an unwelcome *ad hoc* solution, a bullet which the proponent of relationalism needs to bite. On the contrary, a negative view of hallucination is precisely what is needed to solve the crucial challenge that hallucinations

pose to relationalism: denying the indistinguishability principle (IND), without denying the superficiality of phenomenal characters.

It is understandable why, despite the ‘same cause, same effect principle’ and local supervenience clearly being controversial claims, recent discussions of relationalism have focused on versions of the argument on hallucination relying on either of these, rather than focusing on the traditional argument from hallucination hinging on IND. As I argued in Chapter 2, this traditional argument is best seen as a refinement of the traditional argument from illusions and conflicting appearances. Despite being widely and often tacitly assumed, opponents of relationalism have so far failed to present any convincing argument to support its acceptance.

Chapter 3 fills this crucial gap and identifies the intuition underwriting IND: the intuition that, in virtue of its superficiality, the nature of a phenomenal character must be accessible through introspection. Framing the debate over relationalism in terms of a deeper disagreement over the superficiality of phenomenal character and, as I further argued in this chapter, its relation to introspection allowed me to identify what is really at stake in the debate over relationalism, what one needs to commit to in order to challenge relationalism and what resources the relationalist has to respond to this challenge.

Although this poses a serious challenge for the relationalist, in the form of what I called the superficiality constraint, this also imposes more constraints upon the proponent of the argument from hallucination, so much so that it becomes far more difficult than what it is generally thought to

support different premises of the argument and make the argument from hallucination work. In particular, the argument from hallucination cannot succeed without a proper discussion of what it takes for a hallucination to count as indistinguishable from perception, nor can it without a better understanding of how different notions of indistinguishability allow for different explanations of the indistinguishability. Ultimately, I argued, the argument from hallucination has a chance to be successful only if its proponent can make space for the notion of impersonal indistinguishability. This reverses the terms of the current debate, where it is often assumed that the negative view of hallucination needs to accept a notion of impersonal indistinguishability, and such a notion is deemed problematic. I argued that there is a viable notion of impersonal indistinguishability through introspection, but even if there weren't, this would be first and foremost a problem for the proponent of the argument from hallucination, rather than for the proponent of the negative view of hallucination.

Another often neglected but important fact that focusing on the superficiality constraint highlights is that hallucinations are not a uniform class of mental occurrences and do not necessarily deserve all the same account. The only thing they have in common is that they bear some resemblance with perception and often are confused with perception. Most importantly, not all hallucinations pose problems to relationalism. Those that pose a challenge to relationalism are only those that are impersonally indistinguishable from perception through introspection. As I argued in

Chapter 4, there is as of yet no empirical ground available to demonstrate that any documented hallucination fits this condition.

This dispels another taxing misconception that hinders the debate on relationalism: relationalism shouldn't be evaluated on its capacity to account for the phenomenology of ordinary hallucinations, as relationalism is a claim about the phenomenal character of genuine perception, and virtually any account of various sorts of ordinary hallucinations is available to relationalism about perception. It would be an interesting research project to see how we should better account for hallucinations. In the light of the discussion I carried in this thesis, it is safe to predict that a one-size-fits-all account of hallucinations will not be available. However, this research project goes beyond the scope of the assessment of relationalism and is independent of it.

Bibliography

- Aggermaes, A. 1972. 'The Experienced Reality of Hallucinations and Other Psychological Phenomena. An Empirical Analysis.' *Acta Psychiatrica Scandinavica* 48 (3): 220–38.
- Aleman, Andre, and Frank Larøi. 2008. *Hallucinations: The Science of Idiosyncratic Perception*. Washington, DC: American Psychological Association.
- Allen, Paul, Andre Aleman, and Philip K. McGuire. 2007. 'Inner Speech Models of Auditory Verbal Hallucinations: Evidence from Behavioural and Neuroimaging Studies.' *International Review of Psychiatry* 19 (4): 407–15.
- Anscombe, G. E. M. 1965. 'The Intentionality of Sensation: A Grammatical Feature.' In *Analytic Philosophy*, edited by Ronald J. Butler, 158–80. Blackwell.
- Armstrong, David M. 1981. 'What Is Consciousness?' In *The Nature of Mind*, edited by John Heil. Cornell University Press.
- Austin, J. L. 1962. *Sense and Sensibilia*. Oxford University Press.
- Ayer, A. J. 1963. *The Concept of a Person and Other Essays*. New York: St. Martin's Press.
- Ayer, Alfred Jules. 1940. *The Foundations of Empirical Knowledge*. Macmillan.
- Azzopardi, Paul, and Alan Cowey. 2001. 'Why Is Blindsight Blind?' In *Out of Mind: Varieties of Unconscious Processes*, edited by Beatrice De Gelder, Edward H. F. De Haan, and Charles A. Heywood, 3–19. Oxford University Press.
- Badcock, Johanna C., and Kenneth Hugdahl. 2014. 'A Synthesis of Evidence on Inhibitory Control and Auditory Hallucinations Based on the Research Domain Criteria (RDoC) Framework.' *Frontiers in Human Neuroscience* 8 (March). doi:10.3389/fnhum.2014.00180.
- Barnes, Winston H. F. 1944. 'The Myth of Sense-Data.' *Proceedings of the Aristotelian Society* 45 (January): 89–117.
- . 1945. 'The Myth of Sense-Data.' *Proceedings of the Aristotelian Society* 45 (1): 89–118.
- Bayne, Tim, and Maja Spener. 2010. 'Introspective Humility.' *Philosophical Issues* 20 (1): 1–22.
- Bem, Daryl J. 1967. 'Self-Perception: An Alternative Interpretation of Cognitive Dissonance Phenomena.' *Psychological Review* 74 (3): 183–200.
- Benke, Thomas. 2006. 'Peduncular Hallucinosis: A Syndrome of Impaired Reality Monitoring.' *Journal of Neurology* 253 (12): 1561–71. doi:10.1007/s00415-0060-0254-4.
- Bergson, Henri. 1988. *Matter and Memory*. Cambridge, Mass: The MIT Press. Translated from French by Paul N. M. and Palmer W. S. (Originally published in 1986 as *Matière et Memoire*)

- Berrios, G. E. 1996. *The History of Mental Symptoms: Descriptive Psychopathology Since the Nineteenth Century*. Cambridge University Press.
- Berrios, G. E., and T. R. Dening. 1996. 'Pseudohallucinations: A Conceptual History.' *Psychological Medicine* 26 (4): 753–63.
- Block, Ned. 1990. 'Inverted Earth.' *Philosophical Perspectives* 4: 53–79.
- . 1996. 'Mental Paint and Mental Latex.' *Philosophical Issues* 7: 19–49.
- . 2003. 'Mental Paint.' In *Reflections and Replies: Essays on the Philosophy of Tyler Burge*, edited by Martin Hahn and B. Ramberg, 165–200. Mit Press.
- . 2010. 'Attention and Mental Paint.' *Philosophical Issues* 20 (1): 23–63.
- Brewer, Bill. 1999. *Perception and Reason*. Oxford University Press.
- . 2008. 'How to Account for Illusion.' In *Disjunctivism: Perception, Action, Knowledge*, edited by Adrian Haddock and Fiona Macpherson, 168–80. Oxford University Press.
- . 2011. *Perception and Its Objects*. Oxford University Press, USA.
- Broad, C. D. 1925. *The Mind and Its Place in Nature*. Routledge and Kegan Paul.
- . 1952. 'Some Elementary Reflexions on Sense-Perception.' *Philosophy* 27 (January): 3–17.
- Brogaard, Berit. 2015. 'Perception Without Representation? On Travis's Argument Against the Representational View of Perception.' *Topoi*, March, 1–14. doi:10.1007/s11245-015-9312-9.
- Burge, Tyler. 1979. 'Individualism and the Mental.' *Midwest Studies in Philosophy* 4 (1): 73–122.
- . 2003. 'Perception.' *International Journal of Psychoanalysis* 84 (1): 157–67.
- Burnyeat, Myles. 1979. 'Conflicting Appearances.' *Proceedings of the British Academy* 65, sec. 69–111.
- Butler, Dr Jesse. 2013. *Rethinking Introspection: A Pluralist Approach to the First-Person Perspective*. Houndmills, Basingstroke, Hampshire: Palgrave Macmillan.
- Byrne, Alex. 2001. 'Don't PANIC: Tye's Intentionalist Theory of Consciousness.' In *A Field Guide to the Philosophy of Mind*.
- Byrne, Alex, and Heather Logue. 2008. 'Either/Or.' In *Disjunctivism: Perception, Action, Knowledge*, edited by Adrian Haddock and Fiona Macpherson, 29:314–19. Oxford University Press.
- Campbell, J. 2002. *Reference and Consciousness*. Oxford University Press.
- Carruthers, Peter, and Bénédicte Veillet. 2011. 'The Case Against Cognitive Phenomenology.' In *Cognitive Phenomenology*, edited by Tim Bayne and Michelle Montague, 35. Oxford University Press.
- Chalmers, David J. 1996. *The Conscious Mind: In Search of a Fundamental Theory*. Oxford University Press.
- . 2004. 'The Representational Character of Experience.' In *The Future for Philosophy*, edited by Brian Leiter, 153–181. Oxford University Press.

- Child, William. 1992. 'Vision and Experience: The Causal Theory and the Disjunctive Conception.' *Philosophical Quarterly* 42 (168): 297–316.
- Churchland, Paul M. 1984. *Matter and Consciousness*. MIT Press.
- Conduct, M. D. 2011. 'Naïve Realism and Extreme Disjunctivism.' *Philosophical Explorations* 13 (3): 201–21.
- Crane, Tim. 2005a. 'The Problem of Perception.' In *Stanford Encyclopedia of Philosophy*, edited by Edward N. Zalta.
- . 2005b. 'What Is the Problem of Perception?' *Synthese Philosophica* 2 (40): 237–64.
- . 2006. 'Is There a Perceptual Relation?' In *Perceptual Experience*, edited by T. Gendler and J. Hawthorne. Oxford University Press.
- David, Anthony S. 2004. 'The Cognitive Neuropsychiatry of Auditory Verbal Hallucinations: An Overview.' *Cognitive Neuropsychiatry* 9 (1–2): 107–23.
- Descartes, René. 1637. *Discours de la méthode*. Flammarion.
- . 1641. *Meditations on First Philosophy: With Selections From the Objections and Replies*. Oxford University Press Uk.
- Deutsch, Max. 2005. 'Intentionalism and Intransitivity.' *Synthese* 144 (1): 1–22. doi:10.1007/s11229-005-1022-1.
- . 2010. 'Intuitions, Counter-Examples, and Experimental Philosophy.' *Review of Philosophy and Psychology* 1 (3): 447–60.
- Diederich, N. J., F. Alesch, and C. G. Goetz. 2000. 'Visual Hallucinations Induced by Deep Brain Stimulation in Parkinson's Disease.' *Clinical Neuropharmacology* 23 (5): 287–89.
- Dokic, Jérôme, and Jean-Rémy Martin. 2012. 'Disjunctivism, Hallucination and Metacognition.' *WIREs Cognitive Science* 3: 533–43.
- Dorsch, Fabian. 2013. 'Transparency and Imagining Seeing.' In *Disjunctivism – Disjunctive Accounts in Epistemology and in the Philosophy of Perception*, edited by Marcus Willaschek, 5–32. Routledge.
- Dretske, Fred. 1969. *Seeing And Knowing*. Vol. 21. Chicago: University Of Chicago Press.
- . 1981. *Knowledge and the Flow of Information*. MIT Press.
- . 1995. *Naturalizing the Mind*. MIT Press.
- Dummett, Michael. 1975. 'Wang's Paradox.' *Synthese* 30 (3–4): 201–32.
- . 1979. 'Common Sense and Physics.' In *Perception and Identity*, edited by G. F. Macdonald, 1–40. Macmillan Education UK.
- Duppen, Zeno Van. 2015. 'The Phenomenology of Hypo- and Hyperreality in Psychopathology.' *Phenomenology and the Cognitive Sciences*, June, 1–19.
- Eccles, John C. 1966. *Brain and Conscious Experience*. Springer.
- Eperjesi, Frank, and Nabila Akbarali. 2004. 'Rehabilitation in Charles Bonnet Syndrome: A Review of Treatment Options.' *Clinical & Experimental Optometry* 87 (3): 149–52.
- Farkas, Katalin. 2003a. 'Review: The Threefold Cord: Mind, Body and World.' *Mind* 112 (448): 786–89.
- . 2003b. 'What Is Externalism?' *Philosophical Studies* 112 (3): 187–208.

- . 2006. 'Indiscriminability and the Sameness of Appearance.' *Proceedings of the Aristotelian Society* 106 (2): 39–59.
- Farrell, B. A. 1950. 'Experience.' *Mind*, no. 59: 170–98.
- ffytche, D H, R J Howard, M J Brammer, A David, P Woodruff, and S Williams. 1998. 'The Anatomy of Conscious Vision: An fMRI Study of Visual Hallucinations.' *Nature Neuroscience* 1 (8): 738–42.
- ffytche, Dominic H. 2013. 'The Hallucinating Brain: Neurobiological Insights into the Nature of Hallucinations.' In *Hallucination*, edited by Fiona Macpherson and Dimitris Platchias, 45–64. The MIT Press.
- Fish, William C. 2009a. 'Disjunctivism.' In *Internet Encyclopedia of Philosophy*.
- . 2009b. *Perception, Hallucination, and Illusion*. Oxford University Press, USA.
- . 2005. 'Disjunctivism and Non-Disjunctivism: Making Sense of the Debate.' *Proceedings of the Aristotelian Society* 105 (1): 119–27.
- . 2008. 'Disjunctivism, Indistinguishability, and the Nature of Hallucination.' In *Disjunctivism: Perception, Action, Knowledge*, edited by Adrian Haddock and Fiona Macpherson. Oxford University Press.
- Foster, John A. 2000. *The Nature of Perception*. New York: Oxford University Press.
- French, Craig, and Lee Walters. forthcoming. 'On the Argument from Illusion.' *American Philosophical Quarterly*.
- Gersztenkorn, David, and Andrew G. Lee. 2015. 'Palinopsia Revamped: A Systematic Review of the Literature.' *Survey of Ophthalmology* 60 (1): 1–35.
- Gertler, Brie. 2012. 'Renewed Acquaintance.' In *Introspection and Consciousness*, edited by Declan Smithies and Daniel Stoljar, 89–123. Oxford University Press.
- Giustina, Anna, and Uriah Kriegel. forthcoming. 'Fact-Introspection, Thing-Introspection, and Inner Awareness.' *Review of Philosophy and Psychology* 7: 1–22.
- Goldman, Alvin I. 2006. *Simulating Minds*. Oxford University Press.
- Goodman, Nelson. 1951. *The Structure of Appearance*. Vol. 3. Harvard University Press.
- Gopnik, Alison. 1993. 'How We Know Our Minds: The Illusion of First-Person Knowledge of Intentionality.' *Behavioral and Brain Sciences* 16 (1): 1.
- Gothe, Janna, Stephan A. Brandt, Kerstin Irlbacher, Simone Röricht, Bernhard A. Sabel, and Bernd-Ulrich Meyer. 2002. 'Changes in Visual Cortex Excitability in Blind Subjects as Demonstrated by Transcranial Magnetic Stimulation.' *Brain: A Journal of Neurology* 125 (Pt 3): 479–90.
- Grice, H. P. 1961. 'The Causal Theory of Perception, Part I.' *Proceedings of the Aristotelian Society* 121: 121–52.
- . 1988. 'The Causal Theory of Perception.' In *Aristotelian Society Supplementary Volume*, edited by Jonathan Dancy, 121–168. Oxford University Press.

- Haddock, Adrian, and Fiona Macpherson. 2008. 'Introduction: Varieties of Disjunctivism.' In *Disjunctivism: Perception, Action, Knowledge*, edited by Adrian Haddock and Fiona Macpherson. Oxford University Press.
- Halgren, E. 1982. 'Mental Phenomena Induced by Stimulation in the Limbic System.' *Human Neurobiology* 1 (4): 251–60.
- Harman, Gilbert. 1990. 'The Intrinsic Quality of Experience.' *Philosophical Perspectives* 4: 31–52.
- Hellie, Benj. 2007. 'Factive Phenomenal Characters.' *Philosophical Perspectives* 21 (1): 259–306.
- . 2010. 'An Extenalist's Guide to Inner Experience.' In *Perceiving the World*, edited by Bence Nanay, 97. Oxford University Press.
- . 2013. 'The Multidisjunctive Conception of Hallucination.' In *Hallucination*, edited by Fiona Mapherson. MIT Press.
- Hicks, G. Dawes. 1912. 'The Nature of Sense-Data.' *Mind* 21 (83): 399–409.
- . 1938. *Critical Realism*. London, Macmillan.
- Hinton, J. M. 1967a. 'Experiences.' *Philosophical Quarterly* 17 (66): 1–13.
- . 1967b. 'Visual Experiences.' *Mind* 76 (April): 217–27.
- . 1973. *Experiences: An Inquiry Into Some Ambiguities*. Oxford: Clarendon Press.
- Hubl, Daniela, Thomas Koenig, Werner Strik, Andrea Federspiel, Roland Kreis, Chris Boesch, Stephan E. Maier, Gerhard Schroth, Karl Lovblad, and Thomas Dierks. 2004. 'Pathways That Make Voices: White Matter Changes in Auditory Hallucinations.' *Archives of General Psychiatry* 61 (7): 658–68. doi:10.1001/archpsyc.61.7.658.
- Huemer, Michael. 2011. 'Sense-Data.' In *The Stanford Encyclopedia of Philosophy*, edited by Edward N. Zalta, Spring 2011. <http://plato.stanford.edu/archives/spr2011/entries/sense-data/>.
- Hume, David. 1748. *An Enquiry Concerning Human Understanding and Other Writings*. Oxford University Press.
- Hurley, Susan L. 1998. *Consciousness in Action*. Harvard University Press.
- Huxley, Aldous. 1954. *The Doors of Perception: And Heaven and Hell*. New Ed edition. London: Vintage Classics.
- Jackson, Frank. 1977. *Perception: A Representative Theory*. Vol. 87. Cambridge University Press.
- Jackson, John Hughlings. 1931. *Selected Writings. Volume 1*. London: Hodder and Stoughton.
- Jaspers, Karl. 1912. 'The Phenomenological Approach in Psychopathology.' *The British Journal of Psychiatry* 114 (516): 1313–23.
- Jimenez, J. M., M. Todman, M. Perez, J. F. Godoy, and D. V. Landon-Jimenez. 1996. 'The Behavioral Treatment of Auditory Hallucinatory Responding of a Schizophrenic Patient.' *Journal of Behavior Therapy and Experimental Psychiatry* 27 (3): 299–310.
- Johnston, Mark. 2004. 'The Obscure Object of Hallucination.' *Philosophical Studies* 120 (1–3): 113–83.
- Jones, M. R., R. Waggoner, and W. F. Hoyt. 1999. 'Cerebral Polyopia with Extrastriate Quadrantanopia: Report of a Case with Magnetic Resonance Documentation of V2/V3 Cortical Infarction.' *Journal of*

- Neuro-Ophthalmology: The Official Journal of the North American Neuro-Ophthalmology Society* 19 (1): 1–6.
- Knight, Gordon. 2013. ‘Disjunctivism Unmotivated.’ *Phenomenology and the Cognitive Sciences*, 1–18.
- Kriegel, Uriah. 2009. *Consciousness: A Self-Representational Theory*. Oxford University Press.
- . 2013. ‘The Phenomenal Intentionality Research Program.’ In *Phenomenal Intentionality*, edited by U. Kriegel. Oxford University Press.
- Langsam, Harold. 1997. ‘The Theory of Appearing Defended.’ *Philosophical Studies* 87 (1): 33–59.
- . 2000. ‘Why Colours Do Look like Dispositions.’ *Philosophical Quarterly* 50 (198): 68–75.
- Laruelle, Marc, Lawrence S. Kegeles, and Anissa Abi-Dargham. 2003. ‘Glutamate, Dopamine, and Schizophrenia: From Pathophysiology to Treatment.’ *Annals of the New York Academy of Sciences* 1003 (November): 138–58.
- LePoidevin, Robin. 2002. ‘Zeno’s Arrow and the Significance of the Present.’ *Royal Institute of Philosophy Supplement* 50: 57–.
- Lewis, David K. 1986. *On the Plurality of Worlds*. Blackwell Publishers.
- Locke, John. 1690. *An Essay Concerning Human Understanding, 1690*. Menston, Scholar Press.
- Logue, Heather. 2012. ‘What Should the Naïve Realist Say About Total Hallucinations?’ *Philosophical Perspectives* 26 (1): 173–99.
- . 2013. ‘Good News for the Disjunctivist about (One of) the Bad Cases.’ *Philosophy and Phenomenological Research* 86 (1): 105–33.
- . 2014. ‘Experiential Content and Naive Realism: A Reconciliation.’ In *Does Perception Have Content?*, edited by Berit Brogaard. OUP.
- Lowe, E. J. 1992. ‘Experience and Its Objects.’ In *The Contents of Experience*, edited by Tim Crane. Cambridge University Press.
- Lycan, William G. 1996. *Consciousness and Experience*. MIT Press.
- Machery, Edouard, Ron Mallon, Shaun Nichols, and Stephen P. Stich. 2004. ‘Semantics, Cross-Cultural Style.’ *Cognition* 92 (3): 1–12.
- Malhotra, A. K., D. Goldman, C. Mazzanti, A. Clifton, A. Breier, and D. Pickar. 1998. ‘A Functional Serotonin Transporter (5-HTT) Polymorphism Is Associated with Psychosis in Neuroleptic-Free Schizophrenics.’ *Molecular Psychiatry* 3 (4): 328–32.
- Martin, Michael G. F. 1997. ‘The Reality of Appearances.’ In *Thought and Ontology*, edited by M. Sainsbury. Franco Angeli.
- . 1998. ‘Setting Things Before The Mind.’ In *Current Issues in Philosophy of Mind*, edited by Anthony O’Hear, 157–179. Oxford University Press.
- . 2000. ‘Beyond Dispute: Sense-Data, Intentionality, and the Mind-Body Problem.’ In *The History of the Mind-Body Problem*, edited by Tim Crane and Sarah A Patterson. Routledge.
- . 2002a. ‘Particular Thoughts and Singular Thought.’ In *Logic, Thought, and Language*, edited by Anthony O’Hear, 51:173–214. Cambridge University Press.

- . 2002b. ‘The Transparency of Experience.’ *Mind and Language* 4 (4): 376–425.
- . 2004. ‘The Limits of Self-Awareness.’ *Philosophical Studies* 120 (1–3): 37–89.
- . 2006. ‘On Being Alienated.’ In *Perceptual Experience*, edited by Tamar S Gendler and John Hawthorne. Oxford University Press.
- . 2010. ‘What’s In a Look?’ In *Perceiving the World*, edited by Bence Nanay. Oxford University Press.
- . 2013. ‘Shibboleth: Some Comments on William Fish’s Perception, Hallucination & Illusion.’ *Philosophical Studies*, 1–12.
- . 2016. ‘Elusive Objects.’ *Topoi*, June, 1–25. doi:10.1007/s11245-016-9389-9.
- McDowell, John. 1982. ‘Criteria, Defeasibility, and Knowledge.’ *Proceedings of the British Academy* 68: 455–79.
- . 1986. ‘Singular Thought and the Extent of ‘Inner Space.’’ In *Subject, Thought, and Context*, edited by John McDowell and Philip Pettit. Clarendon Press.
- . 1995. ‘Knowledge and the Internal.’ *Philosophy and Phenomenological Research* 55 (4): 877–93.
- . 2013. ‘Perceptual Experience: Both Relational and Contentful.’ *European Journal of Philosophy* 21 (1): 144–157.
- McDowell, John Henry. 1994. *Mind and World: With a New Introduction*. Harvard University Press.
- Merleau-Ponty, Maurice. 1943. *La structure du comportement*. 4e édition. Paris: Presses Universitaires de France - PUF.
- Millar, Alan. 2007. ‘What the Disjunctivist Is Right About.’ *Philosophy and Phenomenological Research* 74 (1): 176–99.
- . 2008. ‘Disjunctivism and Skepticism.’ In *The Oxford Handbook of Skepticism*, edited by John Greco. Oxford University Press.
- Millar, Boyd. 2014. ‘The Phenomenological Problem of Perception.’ *Philosophy and Phenomenological Research* 88 (3): 625–654.
- Millikan, Ruth G. 1991. ‘Perceptual Content and Fregean Myth.’ *Mind* 100 (399): 439–459.
- Moore, G. E. 1903. ‘The Refutation of Idealism.’ *Mind* 12 (48): 433–53.
- . 1926. ‘The Nature of Sensible Appearances.’ *Proceedings of the Aristotelian Society* 6: 179–205.
- Moseley, Peter, Charles Fernyhough, and Amanda Ellison. 2013. ‘Auditory Verbal Hallucinations as Atypical Inner Speech Monitoring, and the Potential of Neurostimulation as a Treatment Option.’ *Neuroscience and Biobehavioral Reviews* 37 (10 Pt 2): 2794–2805.
- Nagel, Thomas. 1974. ‘What Is It Like to Be a Bat?’ *Philosophical Review* 83 (October): 435–50.
- Nanay, Bence. 2014. ‘The Representationalism Versus Relationalism Debate: Explanatory Contextualism About Perception.’ *European Journal of Philosophy* 23 (1): 321–336.
- Nichols, Shaun, and Stephen P. Stich. 2003. *Mindreading. An Integrated Account of Pretence, Self-Awareness, and Understanding Other Minds*. Oxford University Press.

- Noë, Alva, and Evan Thompson. 2004. 'Are There Neural Correlates of Consciousness?' *Journal of Consciousness Studies* 11 (1): 3–28.
- Noordhof, Paul. 2002. 'Imagining Objects and Imagining Experiences.' *Mind and Language* 17 (4): 426–55.
- Nudds, Matthew. 2009. 'Recent Work in Perception: Naïve Realism and Its Opponents.' *Analysis* 69 (2): 334–346.
- . 2013. 'Naïve Realism and Hallucinations.' In *Hallucination*, edited by Fiona Macpherson and Dimitris Platchias, 271. MIT Press.
- Oliveri, M., and G. Calvo. 2003. 'Increased Visual Cortical Excitability in Ecstasy Users: A Transcranial Magnetic Stimulation Study.' *Journal of Neurology, Neurosurgery, and Psychiatry* 74 (8): 1136–38.
- Onofrj, M., L. Bonanni, G. Albani, A. Mauro, D. Bulla, and A. Thomas. 2006. 'Visual Hallucinations in Parkinson's Disease: Clues to Separate Origins.' *Journal of the Neurological Sciences* 248 (1–2): 143–50. doi:10.1016/j.jns.2006.05.025.
- O'Regan, J. Kevin, and Alva Noë. 2001. 'A Sensorimotor Account of Vision and Visual Consciousness.' *Behavioral and Brain Sciences* 24 (5): 883–917.
- O'Shaughnessy, Brian. 1984. 'Seeing the Light.' *Proceedings of the Aristotelian Society* 85 (n/a): 193–218.
- . 2000. *Consciousness and the World*. Oxford University Press.
- Paul, G. A. 1951. 'Is There a Problem About Sense-Data?' In *Aristotelian Society Supplementary Volume*, edited by Gilbert Ryle and Antony Flew, 61–77. Blackwell.
- Pautz, Adam. forthcoming. 'Experiences Are Representations: An Empirical Argument (Forthcoming Routledge).' In *Current Controversies in the Philosophy of Perception*, edited by Nanay. Routledge.
- . 2009. 'A Simple View of Consciousness.' In *The Waning of Materialism: New Essays*, edited by Robert C. Koons and George Bealer, 25–66. Oxford University Press.
- . 2010. 'Why Explain Visual Experience in Terms of Content?' In *Perceiving the World*, edited by Bence Nanay, 254–309. Oxford University Press.
- . 2013. 'Do the Benefits of Naïve Realism Outweigh the Costs? Comments on Fish, '\$<Em Class='a-Plus-Plus'\$>\$Perception, Hallucination and Illusion\$</Em\$>\$.' *Philosophical Studies* 163 (1): 25–36.
- Peacocke, Christopher. 1983. *Sense and Content: Experience, Thought, and Their Relations*. Oxford University Press.
- Pears, D. F. 1976. 'The Causal Conditions of Perception.' *Synthese* 33 (1): 25–40.
- Penfield, W. 1955. 'The Twenty-Ninth Maudsley Lecture: The Role of the Temporal Cortex in Certain Psychical Phenomena.' *The Journal of Mental Science* 101 (424): 451–65.
- Penfield, Wilder, and Kenneth Paine. 1955. 'Results of Surgical Therapy for Focal Epileptic Seizures.' *Canadian Medical Association Journal* 73 (7): 515–31.

- Penfield, Wilder, and Phanor Perot. 1963. 'The Brain's Record of Auditory and Visual Experience a Final Summary and Discussion.' *Brain* 86 (4): 595–696. doi:10.1093/brain/86.4.595.
- Phillips, Ian. 2013. 'Afterimages and Sensation.' *Philosophy and Phenomenological Research* 87 (2): 417–453.
- Phillips, Ian. 2014. 'Lack of Imagination: Individual Differences in Mental Imagery and the Significance of Consciousness.' In *New Waves in Philosophy of Mind*, edited by Mark Sprevak. Palgrave Macmillan. <http://www.palgraveconnect.com/doi/10.1057/9781137286734>.
- Pitcher, George. 1971. *A Theory Of Perception*. Princeton: Princeton University Press.
- Price, H. H. 1932. *Perception*. 2nd ed. London: Methuen.
- Prichard, H. A. 1906. 'Appearances and Reality.—I.' *Mind* XV (58): 223–29.
- Putnam, Hilary. 1975. *Mind, Language, and Reality*. Cambridge University Press.
- Robinson, Howard. 1994. *Perception*. Routledge.
- . 2013. 'The Failure of Disjunctivism to Deal with 'Philosophers' Hallucinations.'" In *Hallucination*, edited by Fiona Macpherson and Dimitris Platchias, 313–30. The MIT Press..
- Rosenthal, David. 2005. 'The Higher-Order Model of Consciousness.' In *Consciousness*, edited by Rita Carter. Weidenfeld & Nicolson.
- Russell, Bertrand. 1910. 'Knowledge by Acquaintance and Knowledge by Description.' *Proceedings of the Aristotelian Society* 11 (5): 108–28.
- . 1912. *The Problems of Philosophy*. Barnes & Noble Books.
- . 1921. *The Analysis of Mind*.
- . 1992. *Theory of Knowledge: The 1913 Manuscript*. Routledge.
- Sacks. 2012. *Hallucinations*. London: Picador.
- Sacks, Oliver. 1985. *Migraine*. London: Pan Books.
- Sarbin, Theodore R. 1967. 'The Concept of Hallucination.' *Journal of Personality* 35 (3): 359–80.
- Schellenberg, Susanna. 2013. 'Externalism and the Gappy Content of Hallucination.' In *Hallucination*, edited by Fiona Macpherson and Dimitris Platchias, 291–312. The MIT Press.
- . 2014. 'The Relational and Representational Character of Perceptual Experience.' In *Does Perception Have Content*, edited by B. Brogaard, 199–219. OUP.
- Schott, G. D. 2007. 'Exploring the Visual Hallucinations of Migraine Aura: The Tacit Contribution of Illustration.' *Brain* 130 (6): 1690–1703.
- Schwitzgebel, Eric. 2011. *Perplexities of Consciousness*. Cambridge, Mass: The MIT Press.
- . 2012a. 'Introspection, What?' In *Introspection and Consciousness*, edited by Declan Smithies and Daniel Stoljar, 29--48. Oxford University Press.
- . 2012b. 'Self-Ignorance.' In *Consciousness and the Self*, edited by JeeLoo Liu and John Perry. Cambridge University Press.
- Searle, John R. 1983. *Intentionality: An Essay in the Philosophy of Mind*. Cambridge University Press.

- Shoemaker, Sydney. 1963. *Self-Knowledge and Self-Identity*. Cornell University Press.
- . 1990. 'Qualities and Qualia: What's in the Mind?' *Philosophy and Phenomenological Research* 50 (Supplement): 109–131.
- . 1996. *The First-Person Perspective and Other Essays*. Cambridge University Press.
- . 2001. 'Introspection and Phenomenal Character.' *Philosophical Topics* 28 (2): 247--73.
- Shorvon, Simon D., Emilio Perucca, and Jerome Engel Jr. 2009. *The Treatment of Epilepsy*. John Wiley & Sons.
- Siegel, Susanna. 2004. 'Indiscriminability and the Phenomenal.' *Philosophical Studies* 120 (1–3): 91–112.
- . 2008. 'The Epistemic Conception of Hallucination.' In *Disjunctivism: Perception, Action and Knowledge*, edited by Adrian Haddock and Fiona Macpherson. Oxford University Press.
- . 2011. *The Contents of Visual Experience*. Oxford University Press. <http://www.oxfordscholarship.com/view/10.1093/acprof:oso/9780195305296.001.0001/acprof-9780195305296>.
- Siewert, Charles. 2004. 'Is Experience Transparent?' *Philosophical Studies* 117 (1–2): 15–41.
- Smith, A. D. 2002. *The Problem of Perception*. Harvard University Press.
- . 2008. 'Disjunctivism and Discriminability.' In *Disjunctivism: Perception, Action, Knowledge*, edited by Adrian Haddock and Fiona Macpherson. Oxford University Press.
- Snowdon, Paul F. 1980. 'Perception, Vision, and Causation.' *Proceedings of the Aristotelian Society* 81: 175–92.
- . 1990. 'The Objects of Perceptual Experience.' *Proceedings of the Aristotelian Society* 64: 121–50.
- . 1992a. 'How to Interpret Direct Perception.' In *The Contents of Experience*. New York: Cambridge University Press.
- . 1992b. 'How to Interpret Direct Perception.' In *The Contents of Experience*. New York: Cambridge University Press.
- . 1998. 'Strawson's Agnostic Materialism.' *Philosophy and Phenomenological Research* 63 (2): 455–60.
- . 2005a. 'Some Reflections on an Argument from Hallucination.' *Philosophical Topics* 33 (1): 285–305.
- . 2005b. 'The Formulation of Disjunctivism: A Response to Fish.' *Proceedings of the Aristotelian Society* 105 (1): 129–41.
- . 2008. 'Hinton and the Origins of Disjunctivism.' In *Disjunctivism: Perception, Action, Knowledge*, edited by Adrian Haddock and Fiona Macpherson, 35--56. Oxford University Press.
- . 2009. 'McDowell on Skepticism, Disjunctivism, and Transcendental Arguments.' *Philosophical Topics* 37 (1): 133–52.
- . 2010. 'Peter Frederick Strawson.' In *Stanford Encyclopedia of Philosophy*.
- . 2011. 'Perceptual Concepts as Non-Causal Concepts.' In *Perception, Causation, and Objectivity*, edited by Johannes Roessler, Hemdat Lerman, and Naomi Eilan. OUP Oxford.

- Sollberger, Michael. 2007. 'The Causal Argument against Disjunctivism.' *Facta Philosophica* 9 (1): 245–67.
- . 2012. 'Causation in Perception: A Challenge to Naïve Realism.' *Review of Philosophy and Psychology* 3 (4): 581–95. doi:10.1007/s13164-012-0116-1.
- Sosa, Ernest. 2007. 'Experimental Philosophy and Philosophical Intuition.' *Philosophical Studies* 132 (1): 99–107.
- Soteriou, Matthew. 2010. 'The Disjunctive Theory of Perception.' In *The Stanford Encyclopedia of Philosophy*, edited by Edward N. Zalta, Winter 2010. <http://plato.stanford.edu/archives/win2010/entries/perception-disjunctive/>.
- . 2013. *The Mind's Construction: The Ontology of Mind and Mental Action*. Oup Oxford.
- . 2016. *Disjunctivism*. Routledge.
- Sprigge, Timothy. 1971. 'Final Causes.' *Proceedings of the Aristotelian Society, Supplementary Volumes*, no. 45: 149–70.
- Steward, Helen. 2011. 'Perception and the Ontology of Causation.' In *Perception, Causation, and Objectivity*, edited by Johannes Roessler, Hemdat Lerman, and Naomi Eilan. Oxford University Press.
- Strawson, Peter F. 1974. 'Causation in Perception.' In *Freedom and Resentment*. Methuen.
- . 1979. 'Perception and Its Objects.' In *Perception and Identity*, edited by G. F. Macdonald, 41–60. Reprinted in *Philosophical Writings*, 2011. Oxford University Press. Macmillan Education UK. doi:10.1007/978-1-349-04862-5_2.
- . 1988. 'Perception and Its Objects.' In *Perceptual Knowledge*, edited by Jonathan Dancy. Oxford University Press.
- . 1998. 'Reply to Paul Snowdon.' In *The Philosophy of P.F. Strawson*. Chicago: Open Court.
- Sturgeon, Scott. 1998. 'Visual Experience.' *Proceedings of the Aristotelian Society* 72 (2): 179–200.
- . 2006. 'Reflective Disjunctivism.' *Aristotelian Society Supplementary Volume* 80 (1): 185–216.
- Taylor, Paul C. J., Vincent Walsh, and Martin Eimer. 2010. 'The Neural Signature of Phosphene Perception.' *Human Brain Mapping* 31 (9): 1408–17. doi:10.1002/hbm.20941.
- Thau, Michael. 2002. *Consciousness and Cognition*. Oxford University Press.
- . 2004. 'What Is Disjunctivism?' *Philosophical Studies* 120 (1–3): 193–253.
- Thomas, Nigel J. T. 2009. 'Visual Imagery and Consciousness.' In *Encyclopedia of Consciousness*, edited by William P. Banks.
- Travis, Charles. 1985. 'Vagueness, Observation, and Sorites.' *Mind* 94 (375): 345–66.
- . 2004. 'The Twilight of Empiricism.' *Proceedings of the Aristotelian Society* 104 (3): 245–270.
- Tudor, Helen, Stephen Davis, Chris R. Brewin, and Peter Howell. 2013. 'Recurrent Involuntary Imagery in People Who Stutter and People

- Who Do Not Stutter.’ *Journal of Fluency Disorders* 38 (3): 247–59.
doi:10.1016/j.jfludis.2013.06.003.
- Tye, Michael. 1992. ‘Visual Qualia and Visual Content.’ In *The Contents of Experience*, edited by Tim Crane, 158–176. Cambridge University Press.
- . 2002. ‘Representationalism and the Transparency of Experience.’ *Noûs* 36 (1): 137–51.
- . 2007. ‘Intentionalism and the Argument From No Common Content.’ *Philosophical Perspectives* 21 (1): 589–613.
- Valberg, Jerome J. 1992a. ‘The Puzzle of Experience.’ In *The Contents of Experience: Essays on Perception*, edited by Tim Crane, 45:18–47. Oxford University Press.
- . 1992b. *The Puzzle of Experience*. Oxford University Press.
- van der Zwaard, R., and M. A. Polak. 2001. ‘Pseudohallucinations: A Pseudoconcept? A Review of the Validity of the Concept, Related to Associate Symptomatology.’ *Comprehensive Psychiatry* 42 (1): 42–50.
- Varela, Francisco, Evan Thompson, and Eleanor Rosch. 1991. *The Embodied Mind: Cognitive Science and Human Experience*. Cambridge, Mass: The MIT Press.
- Vega-Encabo, Jesús. 2010. ‘Hallucinations for Disjunctivists.’ *Phenomenology and the Cognitive Sciences* 9 (2): 281–93.
- Vignal, Jean-Pierre, Louis Maillard, Aileen McGonigal, and Patrick Chauvel. 2007. ‘The Dreamy State: Hallucinations of Autobiographic Memory Evoked by Temporal Lobe Stimulations and Seizures.’ *Brain* 130 (1): 88–99.
- Weinberg, Jonathan M., Shaun Nichols, and Stephen Stich. 2001. ‘Normativity and Epistemic Intuitions.’ *Philosophical Topics*, 29 (1–2): 429–460.
- Weiskrantz, Lawrence. 1986. *Blindsight: A Case Study and Implications*. Oxford University Press.
- . 1996. ‘Blindsight Revisited.’ *Current Opinion in Neurobiology* 6: 215–220.
- . 2009. *Blindsight: A Case Study Spanning 35 Years and New Developments*. OUP Oxford.
- Weiss, A. P., and S. Heckers. 1999. ‘Neuroimaging of Hallucinations: A Review of the Literature.’ *Psychiatry Research* 92 (2–3): 61–74.
- Williamson, Timothy. 1990. *Identity and Discrimination*. Vol. 42. Blackwell.
- Wolters, E. C., and H. W. Berendse. 2001. ‘Management of Psychosis in Parkinson’s Disease.’ *Current Opinion in Neurology* 14 (4): 499–504.
- Zeki, S. 1990. ‘A Century of Cerebral Achromatopsia.’ *Brain: A Journal of Neurology* 113 (Pt 6) (December): 1721–77.
- . 1991. ‘Cerebral Akinetopsia (Visual Motion Blindness). A Review.’ *Brain: A Journal of Neurology* 114 (Pt 2) (April): 811–24.

Résumé de Thèse en Français

Le Relationnalisme et Son Corollaire Disjonctiviste

Le relationnalisme affirme que le caractère phénoménal de la perception est en partie constitué par une relation psychologique non-représentationnelle à des objets indépendants de l'esprit. Il fournit ce qui semble être le compte-rendu le plus simple et le plus intuitif de la façon dont l'expérience nous frappe introspectivement.

Certains aspects du relationnalisme ont été longtemps présents dans l'histoire de la philosophie de la perception, le plus souvent incorporés dans le soi-disant «réalisme naïf», une vision attribuée à l'homme ordinaire, utilisée par les philosophes comme cible vis à vis de laquelle ils construisent leurs positions. Cependant, on assiste seulement récemment, au cours des dernières années, à des tentatives de présenter systématiquement les thèses, les mérites et les implications du relationnalisme, principalement grâce au travail de Snowdon (1980;1990), Martin (1997;2002; 2004; 2006), Campbell (2002), Brewer (2011), Fish (2009) et Soteriou (2013). Le débat récent autour du relationnalisme porte principalement sur l'une de ses implications : le disjonctivisme.

Celui-ci affirme que le caractère phénoménal de l'hallucination diffère du caractère phénoménal de la perception. Ceci est un corollaire de la thèse relationnaliste selon laquelle le caractère phénoménal de la perception est en partie constitué par les objets que la perception nous présente. Si l'on conçoit

le caractère phénoménal de la perception de cette manière, il est évident que l'hallucination ne peut pas avoir le même, puisque dans l'hallucination l'objet qu'il nous semble de voir n'existe pas.

Le terme « disjonctivisme » se doit à la discussion de J.M. Hinton concernant un groupe particulier de propositions qu'il appelle «*perception/illusion disjunctions*» (Hinton 1967a; 1967b; 1973). Hinton affirme qu'il n'y a aucune raison de croire que la catégorie générale de l'expérience dans laquelle les philosophes amalgament la perception et l'hallucination capture vraiment un événement se produisant à chaque fois que l'on perçoit ou bien qu'on hallucine. Hinton affirme que les expressions comme « j'ai une expérience comme d'un éclair de lumière » ou « il me semble de voir un éclair de lumière » sont simplement « une façon plus compacte de dire » qu'une proposition disjonctive est vraie : « soit, je vois un éclair de lumière, soit, je possède une illusion d'un éclair de lumière » (Hinton 1967 : 217).

On trouve dans le débat contemporain plusieurs versions de disjonctivisme, en fonction de l'aspect à l'égard duquel on considère que la perception et l'hallucination diffèrent. Donc, il y a une position disjonctive sur la nature fondamentale de la perception et de l'hallucination (compatible avec plusieurs théories de ce qu'on considère respectivement être la nature fondamentale de la perception et de l'hallucination). C'est ce que Byrne et Logue (2008) appellent « disjonctivisme métaphysique ». Cette version de disjonctivisme est attribuée à Snowdon (1980;1990), qui l'introduit dans le contexte d'une critique de la théorie causale de la perception. Une autre

version de disjonctivisme est le disjonctivisme épistémique, pour lequel la perception et l'hallucination mettent le sujet dans des positions épistémologiques différentes. Le partisan le plus connu de cette thèse est McDowell (1986; 1982; 1995). Enfin, il y a une version de disjonctivisme qui distingue la perception et l'hallucination par rapport à leur caractère phénoménal, présentée par Martin (1997; 2002; 2004; 2006).

C'est sur cette dernière version de disjonctivisme que je vais me concentrer, parce que c'est la position disjonctive qui est entraînée par le relationnalisme, dont il est question dans cette thèse. Par conséquent, dans ce qui suit, je vais utiliser « disjonctivisme » pour désigner exclusivement le disjonctivisme qui concerne les caractères phénoménaux, et je vais mettre de côté toutes les autres versions de disjonctivisme. Il y a des questions ouvertes intéressantes à propos des relations entre cette forme de disjonctivisme et les autres. On peut dire que le disjonctivisme des caractères phénoménaux implique à la fois le disjonctivisme métaphysique et épistémologique, alors qu'il n'est pas clair si l'implication s'obtient dans le sens inverse (cf. Thau 2004; Byrne et Logue 2008; Snowdon 2005, 2009 ; pour une discussion). Cependant, je vais laisser ces questions de côté, car elles ne sont pas pertinentes aux fins de cette thèse, dont le but est de défendre le relationnalisme des objections qui dérivent de l'argument de l'hallucination.

Au-delà d'être un corollaire inévitable du relationnalisme, le disjonctivisme joue un rôle crucial dans la stratégie que les rationalistes adoptent pour contraster l'argument de l'hallucination. L'argument repose sur la possibilité que certaines hallucinations sont indiscernables de la perception

et que, comme on vient de le voir, l'hallucination ne peut pas être constituée par l'obtention d'une relation avec l'objet que l'expérience présente. C'est à dire que la manière dont le relationnalisme caractérise le caractère phénoménal de l'expérience ne s'applique pas aux hallucinations. À partir de ces considérations, l'argument généralise l'échec du relationnalisme dans toute expérience perceptive, y compris quand il s'agit d'une véritable perception d'objet.

La Dialectique du Débat sur le Disjonctivisme

Une grande partie du débat récent autour du relationnalisme a été structurée autour de la dialectique suivante. Les détracteurs du relationnalisme présentent l'argument de l'hallucination. Les relationnalistes répondent que l'argument repose sur une hypothèse injustifiée : que le caractère phénoménal de la perception et de l'hallucination sont de la même nature. Mais ceci est exactement ce que le relationnalisme, avec son corollaire disjonctiviste, dénie. Donc, à moins que le promoteur de l'argument fournisse une motivation pour croire que les perceptions et les hallucinations doivent avoir le même caractère phénoménal, l'argument repose sur une pétition de principe injustifié contre le relationnalisme. À ce point, les promoteurs de l'argument offrent comme motivation de leur conviction, que la perception et hallucination ont le même caractère phénoménal, la remarque que la perception et l'hallucination peuvent être indiscernables par l'introspection.

Souvent, on ne sait pas si, dans les intentions des promoteurs de l'argument, la possibilité d'hallucinations indiscernables de la perception, est

destinée à introduire une abduction ou bien un principe indubitable. Parfois, le disjonctivisme est rejeté dès le départ et n'est même pas pris au sérieux. Ceci suggère que l'affirmation selon laquelle l'indiscernabilité des expériences s'explique en termes d'identité de leur caractère phénoménal est considérée comme étant un principe incontestable—ce que je vais appeler 'principe de l'indiscernabilité' (Siewert 1998; Siegel 2011; Kriegel 2013).

Cependant, il y a très peu de discussions sur les intuitions qui rendent ce principe indéniable. Parfois, cela est motivé sur la base de l'acceptation d'une forme d'internalisme (Farkas 2003). Mais motiver ce principe par une théorie internaliste de l'esprit ne peut certes pas convaincre les relationnalistes, qui—avec bien d'autres—rejettent cette idée. Certains relationnalistes ont identifié la source du principe de l'indiscernabilité dans une vision trop optimiste de nos pouvoirs introspectifs (Hinton 1967, 1970; Martin 1997, 2004, 2006). On trouve aussi quelques rares arguments en soutien à l'idée que la seule façon cohérente de concevoir l'indiscernabilité des expériences est en termes d'identité de leur caractères phénoménaux (Farkas 2006), mais ces arguments ne sont pas parvenus à convaincre les relationnalistes. Ce débat donne parfois l'impression d'une dispute qui ne peut pas être réglée sur son propre terrain, mais seulement par la résolution de controverses beaucoup plus larges et profondes, comme celles entre internalistes et externalistes en philosophie de l'esprit, les discussions sur la nature de la conscience et sur l'introspection et sur ce que signifie avoir conscience de ses propres états mentaux.

Cela donne aussi l'impression d'un domaine divisé en deux champs opposés et incommunicables : d'un côté, ceux qui prennent le principe de l'indiscernabilité pour indéniable et évident, de l'autre côté ceux qui protestent que ce principe est tout sauf qu'évident. Le relationnaliste a certainement raison de se lamenter du manque dramatique de considérations explicites en faveur du principe de l'indiscernabilité et d'exiger des arguments meilleurs et explicites. Cependant, cette plainte est susceptible de laisser impassibles ceux qui déjà souscrivent à ce principe d'indiscernabilité.

Dialectiquement, l'insistance des relationnalistes pour dire qu'ils ne sont pas persuadés par le principe de l'indiscernabilité n'est pas en mesure de convaincre l'adversaire que leur disjonctivisme est bénin. Pour se mettre en position de dissiper les résistances répandues qu'il entraîne contre le disjonctivisme, le relationnaliste doit d'abord identifier les motivations sous-jacentes au principe d'indiscernabilité et essayer d'y répondre. Alors, la seule façon de faire avancer le débat est d'identifier les hypothèses qui souscrivent au principe d'indiscernabilité, essayer de comprendre ce qui les rend persuasives, et décider si elles doivent être acceptées, et si le relationnaliste a les moyens d'accommoder ces intuitions.

D'autre part, certaines discussions récentes autour du relationnalisme ont tendance à présenter l'indiscernabilité présumée de la perception de certaines hallucinations comme une donnée dont toute théorie de la perception devrait rendre compte et que le relationnalisme n'est pas en mesure d'expliquer suffisamment. Elles présentent donc l'argument de l'hallucination comme faisant fond sur une inférence à la meilleure explication (cf. Sturgeon

1998; 2006; Fish 2005; 2009; B. Millar 2014; Tye 2007), par conséquent peut-être, les plaintes des relationnalistes qui déplorent le manque de motivations contraignantes en faveur du principe d'indiscernabilité.

Cependant, l'ambiguïté entre ces deux approches a toujours été présente dans les formulations de l'argument de l'hallucination et cela demeure encore aujourd'hui. Cela a poussé le débat autour du relationnalisme loin de l'évaluation de la validité de l'argument de l'hallucination et de la solidité de ses présupposés. Le débat s'est de plus en plus focalisé vers une discussion autour de la mesure dans laquelle le relationnalisme, en raison de son disjonctivisme, peut fournir une explication satisfaisante de la possibilité d'hallucinations qui sont indiscernables de la perception (cf. Johnston 2004; Smith 2008; Siegel 2004; 2008; Fish 2008; 2008; Hellie 2007; 2013; Schellenberg 2014;).

Cependant, même encadré en ces termes, le débat semble bloqué dans une impasse. Beaucoup de considérations différentes contribuent à la manière dont on élabore un compte rendu philosophique de l'expérience perceptive. Certaines de ces considérations sont phénoménologiques, d'autres pas. Le relationnaliste peut affirmer que sa position fournit globalement une meilleure théorie de l'expérience perceptive, même si elle ne fournit pas la meilleure explication de la donnée de l'indiscernabilité, alors que les adversaires semblent travailler avec le présupposé que rendre compte de la possibilité d'hallucinations indiscernables de la perception soit la priorité absolue d'une théorie de la perception.

Cela a amené Fish (2005) à suggérer que le débat entre disjonctivistes (et donc relationnalistes qui souscrivent au disjonctivisme) et leurs adversaires soit « fondé sur un désaccord méthodologique sous-jacent » (Fish 2005: 121) créant ainsi une situation « pas très différente de celle dont se trouvent les paradigmes de recherche scientifiques concurrents » (Fish 2005: 125). L'adversaire du relationnalisme croit que, « même s'il y a d'autres différences entre une perception et une hallucination, si on ne les distingue pas d'une perspective en première personne, alors, en tout cas, il y a un noyau ontologique commun—le noyau consistant dans le fait que le sujet est dans un 'état sensoriel' du même type » (Fish 2005: 122; en italique dans l'original). Les relationnalistes, d'autre part, « regardent au-delà des similitudes d'une perspective en première personne [et] trouvent une différence importante entre la perception et l'hallucination—la présence ou l'absence d'un objet vu." (Fish 2005: 122; en italique dans l'original).

Ma Contribution Personnelle au Débat

Aucun de ces diagnostics du débat n'est satisfaisant. Il y a quelque chose de profondément troublant dans l'idée que la viabilité d'une théorie sur le caractère phénoménal de la perception ne peut pas être décidée sur son propre terrain, ou, pire, qu'une évaluation du relationnalisme ne se résume qu'à un choix méthodologique où aucune approche montre des avantages significatifs par rapport à l'autre. Ce travail vise à résoudre cette impasse et à fournir les outils nécessaires pour faire avancer le débat, en se recentrant sur l'argument de l'hallucination, en identifiant ce que l'on peut dire pour rendre

l'argument convaincant, quelles sont les intuitions que l'argument mobilise et si elles sont valides et contraignantes. Cela nous permettra d'apprécier où se trouve le désaccord fondamental entre relationnalistes et leurs adversaires et ce que les relationnalistes peuvent dire pour satisfaire les contraintes imposées par ce sur quoi les deux parties opposées se trouvent en accord. Pour comprendre où se situe le désaccord, il faut donc d'abord déterminer les points sur lesquels les deux parties sont d'accord—à savoir, les contraintes qui sont imposées à toute théorie du caractère phénoménal de l'expérience.

Je ferai valoir qu'au cœur de l'argument de l'hallucination—et de l'acceptation du principe d'indiscernabilité, si crucial pour l'argument— se situe l'intuition que les caractères phénoménaux sont des propriétés superficielles et qu'il n'y a aucun moyen pour les relationnalistes de rendre compte de la superficialité des caractères phénoménaux. Le but de cette recherche est de montrer que les relationnalistes peuvent en effet rendre compte de la superficialité des caractères phénoménaux, bien que cela se révèle plutôt difficile. C'est seulement en montrant que le relationnalisme a en soi les ressources pour accommoder l'idée que le caractère phénoménal est superficiel tout en préservant son disjonctivisme, qu'on peut faire avancer le débat une fois pour toute.

La Superficialité des Caractères Phénoménaux et le Relationnalisme

L'idée que les relationnalistes soient forcés de nier que les caractères phénoménaux sont superficiels est d'abord liée à une image trompeuse de la

superficialité, celle qui voit les propriétés superficielles comme des propriétés pour lesquelles il n'y a pas de distinction entre être et paraître. Contre cette manière de concevoir la notion de superficialité, je montre que la réflexion sur les propriétés de couleur indique que les propriétés superficielles peuvent apparaître différemment de ce qu'elles sont. Toutefois, il y a un sens dans lequel on peut dire que les couleurs sont des propriétés superficielles, dans la mesure où, apprendre qu'une certaine chose n'a pas la couleur qu'on croyait initialement ne peut se faire qu'au moyen d'une source d'information particulière—dans ce cas, la vision.

Je présente alors une manière d'entendre la superficialité dans laquelle, si une propriété superficielle semble différente de ce qu'elle est, cela ne peut pas s'expliquer en vertu du fait que la propriété a une véritable nature qui est inaccessible à travers la source dédiée de l'information—en ce qui concerne les caractères phénoménaux, cette source d'information étant l'introspection. La vraie nature du caractère phénoménal, alors, doit être en principe accessible à travers l'introspection. Cela dit, il n'est pas nécessaire qu'elle le soit à toute condition et à tout moment.

Cela pose un sérieux défi pour les relationnalistes : comment peuvent-ils affirmer que le caractère phénoménal de l'hallucination est différent du caractère phénoménal de la perception, sans soutenir que la nature réelle du caractère phénoménal de l'hallucination est cachée et inaccessible à l'introspection? *Ex hypothesis*, au moins quelques hallucinations sont indiscernables par l'introspection de la perception, qui semble exactement dire qu'il est impossible de distinguer les unes des autres par l'introspection.

Ce défi, cependant, ne se présente que si certaines hallucinations sont impossibles à distinguer de l'hallucination d'une manière impersonnelle, d'une manière qui exclut qu'on puisse expliquer le fait qu'on risque de discriminer une hallucination d'une perception en raison de la présence de facteurs qui interfèrent avec l'introspection et qui empêchent le sujet de remarquer une différence qui serait en principe accessible à l'introspection.

L'adversaire du relationnalisme doit prouver alors, qu'il existe des hallucinations qu'on ne peut pas distinguer par l'introspection de la perception dans le sens impersonnel (par brévit , 'HIII': hallucination indiscernable de la perception via l'introspection de mani re impersonnelle), ou bien, que ce type d'hallucination est m taphysiquement possible. Que l'argument de l'hallucination doive respecter cette contrainte est un fait trop souvent ignor , ce qui r sulte dans certaines  valuations partielles de la force de l'argument.

Au moyen d'une enqu te qui d ploie des outils conceptuels aussi bien que des donn es empiriques sur la psychopathologie des hallucinations, je montre qu'il n'y a pas de support empirique pour affirmer que des HIII existent. En outre, les arguments propos s g n ralement pour soutenir leur possibilit  m taphysique—y compris l'argument s'appuyant sur la survenance locale et l'argument qui s'appuie sur le principe « m me cause imm diate, m me effet imm diat »— chouent. Cependant, en fin de compte, il est difficile pour le relationnaliste —pour des raisons diff rentes de celles propos es dans le d bat courant—de nier cette possibilit . Donc, le d fi qui se pose pour le relationnaliste est bien r el et il se doit de l'affronter.

Ce défi peut être relevé en adoptant une vision négative de l'hallucination du genre proposé par Martin (2004; 2006), de sorte que le caractère phénoménal de l'hallucination est uniquement déterminé en fonction de son être indiscernable du caractère phénoménal de la perception. Le relationnaliste doit adopter cette conception de l'hallucination afin de pouvoir éviter de nier que les caractères phénoménaux sont superficiels, ce qui serait contre-intuitif.

En se concentrant sur l'idée que les caractères phénoménaux sont superficiels et en explorant ce que leur superficialité impose à une théorie de l'expérience perceptive, cette enquête fournit une nouvelle approche à l'argument de l'hallucination et au débat sur le relationnalisme. Elle permet d'imposer un ordre à un débat où, en l'état, il est difficile de comprendre quelles sont au juste les motivations des différentes positions dans le débat ou quels sont les points d'accord et de désaccord. Je soutiens que le promoteur de l'argument de l'hallucination est motivé par une intuition qui doit être prise très au sérieux : les caractères phénoménaux sont superficiels et il n'est pas évident que les relationnalistes puissent fournir une explication de comment une perception et une HIII diffèrent dans leur caractère phénoménal sans nier la superficialité de cette dernière.

La présente enquête montre que le relationnaliste peut relever ce défi et que les différentes versions de l'argument de l'hallucination ne parviennent pas à prouver un relationnalisme intenable. D'autres arguments peuvent réussir ce que l'argument de l'hallucination ne parvient pas à faire (cf. Burge 2005; Pautz 2009; 2010; à paraître; Noordhof 2002; Siegel 2011; Brogaard 2015 pour les

arguments qui ne se rattachent pas à l'hallucination). Considérer ces arguments ultérieurs sort du champ de l'étude présente. Cependant, je suis convaincue que l'examen de la notion du caractère phénoménal, de la nature des propriétés superficielles, du fonctionnement de l'introspection, et de la notion d'indiscernabilité affronté dans ce travail se révélera fructueux pour une future évaluation d'ultérieurs arguments contre, mais aussi en faveur, du relationnalisme.

De suite je vais détailler la structure de chaque chapitre.

Premier Chapitre

Dans le premier chapitre, je présente le relationnalisme, la théorie selon laquelle le caractère phénoménal de la perception est caractérisé en termes de l'obtention d'une relation entre le sujet et l'objet. J'identifie les thèses fondamentales de cette position et je vais expliquer comment le relationnalisme caractérise au mieux la manière avec laquelle le caractère phénoménal de l'expérience perceptive nous apparaît par une réflexion introspective.

Tout d'abord, je vais discuter de la façon dont on doit entendre la notion de caractère phénoménal et ce à quoi on doit s'attendre d'une théorie qui vise à l'expliquer.

La notion de caractère phénoménal est introduite pour saisir l'idée que cela fait un certain effet d'avoir une expérience. Le *locus classicus* dans la philosophie analytique contemporaine de l'idée qu'il y ait un effet que ça fait d'avoir une expérience consciente est le célèbre article de Nagel « Quel effet

cela fait, d'être une chauve-souris ? ». Cependant, l'idée que, pour chaque expérience perceptive consciente il y ait l'effet que cela fait d'avoir cette perception, et que l'effet que cela fait d'avoir une perception le distingue d'autres états mentaux précède largement la parution de l'article de Nagel. À Nagel, on doit la diffusion et le succès de l'expression « l'effet que cela fait » ('what it is like'). Nagel affirme:

« [...] le fait même qu'un organisme possède une expérience consciente montre que cela fait un certain effet *d'être* cet organisme. Cela peut impliquer d'autres choses [...] mais fondamentalement un organisme a des états mentaux conscients si cela lui fait un certain effet *d'être* cet organisme—un certain effet *pour* cet organisme » (Nagel 1974 : 436. Italics dans l'original).⁸⁶

Ce que Nagel appelle « caractère subjectif », d'autres l'appellent « caractère phénoménal », « caractère conscient » ou « caractère sensible ». Ici, je vais adopter l'expression « caractères phénoménaux », l'expression plus répandue dans les débats récents sur la perception.

Je prends en considération le scepticisme de Snowdon quant à l'utilité d'utiliser la notion de caractère phénoménal et l'expression apparentée « what-it-is-like » (« l'effet que cela fait »). J'explique pourquoi je ne partage pas le scepticisme de Snowdon et pourquoi je pense que la notion de caractère

86 Traduction Française de P. Engel, « Quel effet cela fait-il d'être une chauve-souris? » dans *Questions mortelles*, PUF, 1983.

phénoménal a bien une place dans le discours philosophique. Cela dit, l'usage de la notion de 'caractère phénoménal' dit se limiter à celle d'une variable libre : une expression que l'on utilise pour identifier certains aspects de l'expérience perceptive qui font appel à une meilleure compréhension. Le caractère phénoménal indique l'effet que cela fait d'avoir une expérience perceptive. C'est une question ouverte et controversée de savoir *quel* effet cela fait. La nature du caractère phénoménal d'une expérience n'est donc pas évidente pour nous : contrairement à ce que beaucoup de participants au débat courant assument (tout en étant en désaccord les uns avec les autres quant à ce qu'est le caractère phénoménal de l'expérience), ce n'est pas évident de savoir comment l'expérience perceptive nous affecte, quel effet cela nous fait d'avoir une expérience.

En particulier, je contrecarre l'idée que comprendre le caractère phénoménal des expériences en termes de propriétés intrinsèques que l'expérience possède n'est pas un avec la notion même de caractère phénoménal. Cela laisse ouverte la possibilité que le caractère phénoménal de la perception soit (en partie) constitué par l'obtention d'un rapport à des objets dans l'environnement, selon ce que le relationnaliste maintient.

Dans ce chapitre, j'identifie les deux caractéristiques de l'expérience dont on a accès par l'introspection. La première caractéristique est ce qu'on appelle souvent la transparence de l'expérience. La thèse de la transparence contient un volet positif et un négatif.

Le volet négatif de la thèse de la transparence affirme que, quand je dirige une attention introspective à mon expérience, je prends conscience des

objets qui semblent indépendants de l'esprit et de leurs propriétés— par exemple le buisson de lavande en bas de la rue et sa couleur. À savoir, ce que je rencontre par mon introspection sont les mêmes objets et les propriétés que je rencontre quand je dirige mon attention vers le monde. Il y a donc une évidence introspective à faveur de la thèse que les propriétés des objets que je perçois font partie de ce que dont je prends conscience quand je dirige mon attention introspective vers le caractère conscient de mon expérience.

Le volet négatif de la thèse de la transparence affirme que, quand je dirige mon attention introspective vers mon expérience, il ne me semble pas que je prenne conscience des propriétés de mon expérience de manière indépendante de la prise de conscience de l'objet de mon expérience et de ses propriétés. Ceci est compatible avec l'idée qu'il y ait des propriétés de l'expérience indépendantes des objets que je perçois. Cependant, la thèse affirme qu'on ne parvient pas à en avoir conscience de manière indépendante de la conscience des propriétés et des entités qui font l'objet de la perception. Cela indique un manque d'évidence introspective pour la thèse dont il faut expliquer la nature du caractère phénoménal de l'expérience en termes de propriétés intrinsèques de la perception, telles que les *qualia*.

L'autre caractéristique que l'expérience semble avoir est sa relationnalité : il nous semble d'être liés aux objets que nous percevons d'une manière différente de la manière dont nous sommes liés à eux quand on y pense. Il semble il y avoir une relation psychologique particulière avec les objets de notre perception.

Le relationnalisme—avec sa thèse centrale selon laquelle le caractère phénoménal de l'expérience est constituée par les objets que l'on perçoit et par leurs propriétés (cf. Campbell 2002: 116 ; Martin 2006: 357 ; Brewer 2011: 94 ; Soteriou 2013: 23)—semble mieux rendre compte de la relationnalité et de la transparence de la perception par rapport aux théories alternatives du caractère phénoménal de l'expérience, à savoir la théorie des *sense-data*, le représentationalisme et la théorie des *qualia*.

Cependant, le relationnalisme a rencontré et rencontre toujours une forte résistance, malgré ses mérites explicatifs manifestes. La raison de cette résistance se doit à un argument qui prend la forme d'une *reductio ad absurdum* du relationnalisme. L'argument établit d'abord que, dans certains cas spéciaux d'expérience perceptive (typiquement illusions et/ou hallucinations), le caractère phénoménal n'est pas constitué par les objets que l'on a l'impression de percevoir : le buisson de lavande au bout de la rue et l'ordinateur portable sur mon bureau. La gamme des expériences envisagées inclut les cas d'apparences contradictoires, d'illusions et d'hallucinations, souvent traités de façon très similaire. Puis, l'argument généralise cette conclusion à toute expérience, y compris les cas dans lesquels l'objet qu'il nous paraît de voir est bel et bien présent.

Valberg présente le problème posé par ce genre d'argument dérivé de la réflexion sur les illusions et les hallucinations comme une antinomie (1992a; 1992b). L'antinomie surgit de l'opposition de deux façons parfaitement adéquates de réfléchir sur notre expérience, qui cependant parviennent à des conclusions opposées.

D'une part, on peut suivre le raisonnement problématique détaillé ci-dessus impliquant l'argument des illusions et des hallucinations. D'autre part, on peut être «ouvert à notre expérience» : on peut considérer la façon dont l'expérience nous frappe introspectivement et envisager la meilleure façon de rendre compte de la phénoménologie qui nous est révélée par l'introspection. Si on est ouvert à la façon dont l'expérience nous frappe, on est amené à penser à l'expérience comme à une relation aux objets qui nous entourent : le buisson de lavande dans la rue, l'ordinateur portable sur le bureau, et ainsi de suite. L'antinomie se pose parce qu'on ne peut qu'être ouvert à la façon dont l'expérience nous frappe : on ne peut que savoir que cet objet-ci, cet objet que je vois, est ce à quoi ma perception me met en relation. Mais en même temps, l'argument semble correct : il nous interdit donc de croire ce qu'on connaît déjà de notre expérience. Ceci nous plonge dans une situation paradoxale : il nous faut accepter, mais on ne peut pas accepter que ce que l'on voit est l'objet que l'on voit devant soi.

Valberg remarque qu'il y a deux positions possibles que l'on peut adopter vis-à-vis de l'antinomie, une fois qu'on y est confronté sa portée nous est claire. On peut essayer de la résoudre et « démontrer que le conflit dans lequel il est constitué n'est qu'apparent. » (Valberg 1992b: 42). Cela demande d'exposer une erreur dans l'argument de l'illusion et de l'hallucination. Alternativement, on peut essayer d'expliquer l'antinomie, c'est-à-dire d'expliquer pourquoi « il y a vraiment un conflit tel que nous l'avons décrit » (Valberg 1992b: 42).

On peut voir le relationnalisme comme une position qui essaye de résoudre l'antinomie, en montrant qu'il y a une erreur dans l'argument de l'illusion et de l'hallucination. Les thèses concurrentes, par contre, en raison de leur acceptation de ces arguments, se trouvent en position de devoir expliquer pourquoi et comment le paradoxe a lieu.

Deuxième Chapitre

Ce chapitre analyse la structure de l'argument de l'illusion et de l'hallucination. Tout d'abord nous avons besoin de prouver ses prémisses, pour comprendre s'il y a bien un argument convaincant contre le relationnalisme. Tout d'abord, il faut comprendre si, pour un certain groupe d'expériences, l'argument parvient à montrer que le relationnalisme ne peut pas s'appliquer. Puis, il faut comprendre s'il est possible de généraliser la fausseté du relationnalisme à toutes les expériences, y compris les perceptions authentiques.

Dans sa forme traditionnelle, l'argument utilise des phénomènes très différents, allant des cas de changements de perspective et de réfraction — interprétés comme des cas d'apparitions contradictoires — à des hallucinations complexes et il tend à les traiter tous indistinctement comme des cas d'illusions.

Contre l'argument qui veut que le relationnalisme ne s'applique pas aux cas d'apparences contradictoires, le relationnaliste peut protester que le partisan de l'argument décrit d'une manière inadéquate les phénomènes discutés, en ignorant le bien connu phénomène de la constance de couleur,

forme, dimension et ainsi de suite. En effet, les cas de variations de couleurs sous des illuminations différentes, de variation de forme et dimension selon la position relative du sujet par rapport à l'objet vu, etc, ne sont pas des cas où l'objet ne paraît pas avoir du tout la propriété qu'il possède en effet. Je présente aussi la manière, phénoménologiquement plus adéquate, dont les relationnalistes décrivent et expliquent ces phénomènes.

Pour ce qui concerne les illusions d'optiques, le relationnaliste ne peut pas invoquer la constance perceptive, ni dire qu'il ne s'agit pas de cas où quelque chose nous apparaît avoir des propriétés qu'il n'a pas. La première est l'idée qu'être conscient d'un objet de perception exclut qu'on soit conscient—dans un même acte de perception, au même endroit etc—d'un autre objet. (cf. Snowdon 1992: 74 et French et Walters (à paraître). Je montre que ce principe est clairement faux. D'autre part, un principe qui pourrait le remplacer n'est pas évidemment faux, mais il n'est pas clair non plus qu'il soit vrai.

L'argument de l'illusion s'appuie sur un autre principe crucial, celui phénoménal : l'idée que, à chaque fois qu'à un sujet il paraît (perceptivement) que quelque chose ait une certaine propriété, le sujet est conscient d'un objet qui possède cette propriété. (cf. Robinson 1994: 32). Ce principe est présenté par les partisans des sense-data comme phénoménologiquement motivé. Cependant, on ne peut pas présenter ce principe comme la meilleure manière de rendre compte de ce que c'est, pour le sujet, d'avoir une illusion, car accepter le principe phénoménal nous obligerait à rejeter un aspect très important de la façon dont l'expérience nous paraît introspectivement : c'est à dire, le fait que l'objet qui m'apparaît comme jaune (même s'il ne l'est pas)

me semble bien être cet objet-ci, que je regarde, que je touche et j'indique ; non pas un autre objet (un sense-datum) qui le remplacerait.

J'en conclus qu'on ne peut pas établir que le relationnalisme ne s'applique pas aux cas d'illusions d'optiques et d'apparitions contradictoires. Cependant, il est vrai que le compte rendu que les relationnalistes donnent de la perception ne peut effectivement pas s'appliquer aux hallucinations : aucun argument n'est demandé pour établir cette simple affirmation. Cependant, il n'est pas clair qu'on puisse étendre la fausseté du relationnalisme établie pour le cas des hallucinations aux cas de perception véritable.

Les partisans de l'argument motivent cette généralisation en vertu d'une similitude phénoménologique entre la perception et l'hallucination ou bien, dans certaines versions, de leur indiscernabilité. Toutefois, il est curieux que quelqu'un soit prêt d'accepter une inférence de l'indiscernabilité entre deux expériences et leur identité : normalement, on sait bien qu'on ne peut pas inférer une proposition concernant l'identité métaphysique de deux choses à partir du fait épistémique que quelqu'un n'est pas en mesure de les discerner. Dans ce chapitre, je montre, en m'appuyant sur Hinton, que ce qui motive vraiment cette généralisation est plutôt l'idée que l'indiscernabilité entre la perception et l'hallucination demande une explication et que la meilleure explication de ce fait est de dire que les deux expériences ont des propriétés en commun, dans ce cas, leur caractère phénoménal. L'argument repose donc sur l'acceptation d'un principe, que j'appelle, d'après Martin (2004 ; 2006), « Principe de l'Indiscernabilité », ou pour abrégé, IND :

(IND) : si deux expériences sont introspectivement indiscernables, elles ont le même caractère phénoménal.

Les relationnalistes peuvent rétorquer qu'on ne peut pas invoquer la présence du même caractère phénoménal pour la perception et pour l'hallucination comme la meilleure explication de la phénoménologie de ces expériences, parce que cela nous engagerait à adopter une théorie révisionniste de la perception véritable. Les partisans de l'argument de l'hallucination doivent par conséquent viser à une thèse plus radicale qu'une inférence à la meilleure explication. Ils doivent démontrer que la *seule* manière cohérente d'expliquer l'indiscernabilité entre perception et hallucination, est d'invoquer l'identité de leurs caractères phénoménaux.

Cela nous permet de reconstruire l'argument de l'hallucination comme un argument valide et dont les prémisses ont une certaine plausibilité relative :

1. Le relationnalisme affirme qu'on doit rendre compte du caractère phénoménal des perceptions en termes de l'obtention d'une relation psychologique avec les objets qu'on a l'impression de percevoir.

2. Dans certaines hallucinations, on a une expérience qui est introspectivement indiscernable d'une perception.

3. Dans l'hallucination, l'objet qu'il me semble de percevoir n'existe pas.

4. (de 1 et 3) Dans l'hallucination, le caractère phénoménal de la perception ne peut pas dépendre de l'obtention d'une relation avec l'objet qu'il me semble de percevoir.

5. Si deux expériences conscientes sont introspectivement indiscernables, elles ont le même caractère phénoménal IND.

6. (de 4 et 5) Le relationnalisme ne peut pas être vrai même dans le cas véridique.

Troisième Chapitre

Dans le troisième chapitre j'explore les raisons qui pourraient amener à l'acceptation de l'IND. Il est curieux et surprenant que quelqu'un puisse croire que le fait épistémologique, que certaines hallucinations sont indiscernables de la perception, doit entraîner une thèse métaphysique, telle que l'identité du caractère phénoménal des deux expériences. Si beaucoup de philosophes acceptent l'IND, cela doit être parce qu'ils pensent qu'il y a quelque chose de spécial concernant les expériences et leur caractère phénoménal qui fait que, dans ce cas, l'inférence allant d'une thèse épistémologique à une thèse métaphysique soit valide.

Je considère l'hypothèse que les partisans de l'IND soient motivés par l'idée que l'introspection soit infaillible. Mais celle-ci est une thèse que très peu de philosophes aujourd'hui seraient prêts à accepter. Y a-t-il donc une autre motivation pour accepter l'IND ? Je fais valoir que l'acceptation de l'IND repose sur l'intuition que les caractères phénoménaux sont superficiels. Il y a une certaine conception de la notion de superficialité que, si acceptée, rendrait l'IND nécessairement vraie. Selon cette conception, les propriétés superficielles sont des propriétés pour lesquelles il n'y a pas de distinction entre être et apparaître. Je soutiens que, malgré son succès, cette manière de

comprendre la superficialité est problématique et qu'il existe une autre manière, plus adéquate, de concevoir la superficialité. Je présente donc une théorie alternative à la superficialité, selon laquelle si une propriété est superficielle par rapport à une certaine source d'information, sa nature doit être accessible, au moins sous certaines conditions, à travers la source d'information pertinente.

Cette façon de comprendre la superficialité permet qu'une propriété superficielle apparaisse différemment de ce qu'elle est, à condition qu'il y ait des facteurs d'interférence qui empêchent la source d'information pertinente de fonctionner correctement. Bien que cette façon de comprendre la superficialité laisse ouverte la possibilité qu'un caractère phénoménal soit différent de ce qu'il peut sembler au sujet lorsque celui-ci introspecte son expérience, elle pose des contraintes pour toute théorie du caractère phénoménal. La notion de superficialité impose que, si un caractère phénoménal est superficiel, il doit être accessible à travers la source d'information pertinente (dans ce cas, l'introspection). S'il y a une différence entre le caractère phénoménal de la perception et le caractère phénoménal d'une hallucination, cette différence doit être accessible (au moins en principe et dans des conditions optimales), par l'introspection. Mais il n'est pas clair comment le relationnaliste puisse affirmer que la différence qu'il postule entre le caractère phénoménal d'une perception et celui d'une hallucination soit détectable par l'introspection si ces mêmes expériences sont indiscernables par l'introspection.

Le relationnaliste doit donc expliquer comment il peut soutenir qu'une perception et une hallucination indiscernables ont un caractère phénoménal

différent sans pourtant devoir nier que le caractère phénoménal de l'hallucination est superficiel. Il semble que le relationnaliste ne peut qu'affirmer que l'hallucination a un caractère phénoménal dont la vraie nature est inaccessible à l'introspection. Si le relationnalisme amène vraiment à cette conclusion, il semble que nous devrions rejeter le relationnalisme, car dire que le caractère phénoménal de l'hallucination est inaccessible à l'introspection viole nos intuitions les plus profondes sur les caractères phénoménaux, à savoir qu'ils sont superficiels. Cela pose le problème auquel le reste de cette thèse devra répondre.

Toutefois, afin de prouver que le relationnaliste se trouve à devoir nier la superficialité du caractère phénoménal de l'hallucination, l'adversaire doit montrer que certaines hallucinations sont impossibles à distinguer de la perception d'une manière impersonnelle. J'appelle ce type d'hallucinations 'hallucinations qui sont indiscernables de la perception par l'introspection d'une manière impersonnelle', ou HIII pour abrégé. Ces considérations nous amènent à formuler l'argument suivant contre le relationnalisme — un argument que j'appelle 'argument de la superficialité'. On peut le concevoir à la fois comme un argument direct contre le relationnalisme et comme un argument pour supporter l'IND :

1. Il y a des HIII.
2. Une HIII est une expérience phénoménalement consciente.
3. (de 1) Une HIII a un caractère phénoménal.

4. Si les caractères phénoménaux sont superficiels, le caractère phénoménal d'une expérience est accessible (du moins en principe) par l'introspection (contrainte de superficialité).

5. Le relationnalisme affirme que le caractère phénoménal d'une HIII est différent du caractère phénoménal d'une perception.

6. (de 4) Le caractère phénoménal d'une HIII est différent de ce qu'il semble être : il ressemble au caractère phénoménal d'une perception, mais il ne l'est pas.

7. (2 et 6) le relationnalisme doit affirmer que le caractère phénoménal d'une HIII n'est pas accessible par l'introspection.

8. (de 3 et 6) Le relationnalisme doit nier que les caractères phénoménaux des HIII sont superficiels.

9. Les caractères phénoménaux sont superficiels.

10. (de 8 et 9) Le relationnalisme est intenable.

Ainsi, le relationnaliste devra résister à la conclusion de cet argument. À cette fin, il a besoin de résister à l'une des prémisses de l'argument. On discutera dans les chapitres suivants si le relationnalisme peut résister à l'une ou à l'autre prémisses de cet argument.

Si cet argument est correcte, le relationnalisme faillit parce qu'il n'est pas en mesure d'expliquer le fait qu'une perception et une hallucination peuvent être indiscernables. Cela signifie aussi que l'IND, dans l'argument classique de l'hallucination tel que je l'ai présenté dans le deuxième chapitre doit être accepté. Cependant, cette discussion montre aussi qu'une contrainte s'ajoute à l'argument de l'hallucination classique : l'IND est vraie, seulement

si on interprète la notion d'indiscernabilité de manière impersonnelle. Par conséquent, l'argument est valide, seulement s'il est possible d'établir qu'il y a non seulement des hallucinations qui sont indiscernables des perceptions, mais qu'elles le sont de manière impersonnelle, c'est à dire qu'il y a des HIII.

Chapitre 4

Dans ce chapitre, j'explore si les relationnalistes peuvent résister à la première prémisse de l'argument présenté dans le chapitre précédent. Il est important de remarquer que, si cette prémisse est fausse, l'argument de l'hallucination tombe tout aussi bien que l'argument présenté dans le dernier chapitre. Car l'IND est seulement si on considère la notion d'indiscernabilité d'une manière impersonnelle. Ceci signifie que l'IND peut s'appliquer seulement à des HIII. L'argument de l'hallucination doit donc être révisé comme de suite, avec une modification à la première prémisse :

1. Le relationnalisme affirme qu'on doit rendre compte du caractère phénoménal des perceptions en termes de l'obtention d'une relation psychologique avec les objets qu'on a l'impression de percevoir.

2*. Dans certaines hallucinations, on a une expérience qui est introspectivement indiscernable (de manière impersonnelle) d'une perception. En d'autres termes, il y a des HIII.

3. Dans l'hallucination, l'objet qu'il me semble de percevoir n'existe pas.

4. (de 1 et 3) Dans l'hallucination, le caractère phénoménal de la perception ne peut pas dépendre de l'obtention d'une relation avec l'objet qu'il me semble de percevoir.

5. Si deux expériences conscientes sont introspectivement indiscernables, elles ont le même caractère phénoménal IND.

6. (de 4 et 5) Le relationnalisme ne peut pas être vrai même dans le cas véridique.

Pour démontrer que la prémisse (1) dans l'argument de la superficialité et la prémisse (2*) dans l'argument de l'hallucination révisé sont vraies, le relationnaliste pourrait d'abord essayer de montrer que la notion des HIII est incohérente. Tout d'abord, donc, je considère si la notion des HIII est cohérente. Pour cela, je propose différentes manières dont on peut concevoir cette notion et je propose celle que je crois être la meilleure interprétation. Enfin, j'analyse la structure que prend la relation d'indiscernabilité quand elle concerne la relation entre les HIII et la perception.

Je soutiens qu'affirmer qu'une hallucination est une HIII signifie que l'on est incapable de reconnaître par l'introspection que l'hallucination ne fait pas partie de la classe des perceptions. La relation établit donc qu'une hallucination particulière est indiscernable de toutes les perceptions : elle semble appartenir à la classe des perceptions. La structure de cette relation est donc une relation entre un particulier (l'hallucination) et une classe (la classe des perceptions). L'affirmation selon laquelle une hallucination est indiscernable de la perception revient à l'incapacité de classer correctement

l'expérience hallucinatoire que l'on vit comme une hallucination plutôt que comme une perception. Je conclus que la notion des HIII est bien cohérente. Il reste à savoir si ce genre d'hallucinations existe vraiment.

Pour démontrer cela, il faudrait démontrer qu'il existe des cas dans lesquels, si les conditions avaient été plus favorables (par exemple, si le sujet était dans un état «altéré» d'esprit, s'il était plus attentif, s'il n'avait pas une déficience cognitive, etc), le sujet n'aurait pas été en mesure de distinguer son hallucination de la perception. En m'appuyant sur les évidences que l'on trouve dans la littérature psychopathologique, je montre qu'il n'y a pas de preuve claire de l'existence des HIII. J'examine les descriptions d'hallucinations que l'on trouve dans la littérature clinique et je conclus qu'il n'y a pas d'évidence empirique pour soutenir qu'il existe des cas de HIII, c'est-à-dire des cas dans lesquels rien n'interfère avec la capacité du sujet à distinguer son hallucination actuelle d'une véritable perception. Il n'y a donc aucune preuve empirique pour accepter (2*) dans l'argument de l'hallucination ou (1) dans l'argument de la superficialité.

La discussion que je mène dans ce chapitre concernant les différentes hallucinations rapportées dans les annales de psychopathologie nous délivre une autre leçon importante, que les philosophes feraient bien d'écouter : les hallucinations sont des phénomènes bien plus variés (quant à leur phénoménologie, étiologie, population concernée et neurologie sous-entendue) que les philosophes ont tendance à assumer. Ceci suggère qu'il est probablement futile d'essayer de parvenir à une définition ou une explication qui vaille pour tous les cas d'hallucination : du phosphène au syndrome de

Carles Bonnet, de la vision double aux hallucinations chez les schizophrènes. Cette remarque est pertinente lorsqu'il s'agit d'évaluer un argument alternatif contre le relationnalisme, du type phénoménologique qu'un détracteur pourrait être tenté de proposer (cf. Millar et Crane pour des exemples sur ce type d'argument) même s'il se trouve qu'il n'y a pas de HIII. L'argument serait que l'hallucination et la perception sont phénoménologiquement très semblables, et pour cette raison nous devons expliquer leur caractère phénoménal de la même manière.

Mais la simple similarité entre perception et hallucination, nous l'avons déjà largement vue, ne peut pas justifier l'hypothèse que les deux aient le même caractère phénoménal. Millar (et bien d'autres philosophes, de manière plus ou moins explicite) présente l'hypothèse que la perception et l'hallucination ont le même caractère phénoménal comme une hypothèse supportée phénoménologiquement. Mais aucune donnée phénoménologique peut, à elle seule, supporter cette hypothèse.

Si l'on regarde la littérature psychopathologique, on voit bien qu'il n'est même pas possible de parvenir à une définition unifiée des hallucinations, et encore moins des perceptions, des illusions et des hallucinations. Il n'y a pas de caractéristiques phénoménologiques partagées par toutes les hallucinations de manière univoque. Et il existe encore moins, donc, des caractéristiques phénoménologiques partagées par toutes les hallucinations avec les perceptions, caractéristiques qui requièrent une théorie unitaire.

En outre, même si l'on admet que seulement quelques hallucinations ont certaines caractéristiques phénoménologiques en commun avec la perception, il n'y a aucune raison de postuler qu'on doive expliquer leur caractère phénoménal de la même façon. Le fait que *certaines* similitudes existent entre perception et hallucination n'est certainement pas quelque chose que le relationnaliste ne voit pas, comme nous l'avons déjà vu dans le deuxième chapitre. Toutefois, les relationnalistes rejettent l'idée qu'il faille expliquer cette ressemblance de la même manière, à savoir, qu'ils rejettent l'IND. Mais, comme on l'a vu, l'IND a des chances d'être plausible seulement si l'on utilise une notion impersonnelle d'indiscernabilité. Par conséquent, l'IND ne peut s'appliquer qu'aux hallucinations si celles-ci sont des HIII. Mais, comme nous l'avons vu dans ce chapitre, aucune preuve phénoménologique ne peut étayer que certaines hallucinations sont des HIII. Par conséquent, il n'y a aucun argument phénoménologique de l'hallucination contre le relationnalisme.

Il y a une tendance à évaluer le relationnalisme sur la base de sa capacité à fournir une théorie viable et crédible du caractère phénoménal des hallucinations. Et on pense souvent que les relationnalistes ne sont pas en mesure de fournir une telle théorie. Cependant, si l'argument de l'hallucination ne parvient pas à son but (et jusque là, nous ne détenons pas les éléments pour affirmer le succès de cet argument), les relationnalistes peuvent dire ce qu'ils veulent sur ce qui détermine le caractère phénoménal de l'hallucination. Ils peuvent dire que le caractère phénoménal des hallucinations est constitué par des sense-data, des qualia, des propriétés intentionnelles, des

relations à des zones de l'espace environnant, des relations à des universaux, ou n'importe quoi. Il se peut aussi qu'on explique le caractère phénoménal de certaines hallucinations d'une certaine manière, et le caractère phénoménal d'autres hallucinations d'une autre, étant donné la différence entre plusieurs hallucinations au niveau de leur phénoménologie et de leur étiologie. Bien sûr, pour toute théorie des hallucinations que le relationnaliste propose, il reste à voir si elle est en soi satisfaisante. Pourtant chaque théorie de l'hallucination devra être évaluée indépendamment pour ses propres mérites, non pas pour les conséquences qu'elle pourrait avoir pour une théorie de la perception véridique. On peut déplorer que les relationnalistes n'aient pas encore assez dit sur la façon dont nous devrions tenir compte des hallucinations. Mais cela équivaudrait simplement à exprimer le dégoût pour le sujet de recherche choisi par le relationnaliste : se plaindre que le relationnaliste s'intéresse au caractère phénoménal de la véritable perception, plutôt que de celui de l'hallucination.

En outre, la discussion dans ce chapitre montre que, même s'il se trouve que nous devons accepter l'existence ou la possibilité métaphysique des HIII, il est toujours possible d'expliquer le caractère phénoménal de toutes les hallucinations qui ne sont pas des HIII comme mieux il nous semble. Le succès de l'argument de l'hallucination imposerait des contraintes uniquement sur ce qu'on peut dire des HIII. Le défaut de réaliser cela est largement dû à l'absence dans le débat actuel, d'une discussion sérieuse sur la notion d'indiscernabilité et de ce qui est requis pour l'IND pour être plausible, entre les partisans et les adversaires du relationnalisme.

Cinquième Chapitre

Bien qu'il n'y ait pas d'évidence empirique pour supporter l'hypothèse qu'il existe des HIII, le promoteur de l'argument de l'hallucination peut dire qu'elles sont tout de même métaphysiquement possibles.

Par conséquent, on peut remplacer la prémisse (2*) dans l'argument avec la suivante proposition modale :

(2#). Il est possible d'avoir une hallucination qui est impersonnelle introspectivement indiscernable d'une véritable perception.

Dans ce chapitre, j'explore si (2 #) devrait être accepté.

Bien que (2#) puisse sembler évident, il n'est pas clair ce qui peut supporter une telle affirmation. (2#) repose sur l'hypothèse selon laquelle il est possible d'activer exactement les mêmes zones du cerveau qui sont impliquées dans la perception et de reproduire l'état du cerveau qu'on a quand on perçoit réellement. Cependant, cette seule prémisse ne peut pas soutenir (2#). Quel principe doit accompagner cette thèse pour faire en sorte que (2 #) soit vrai?

On peut essayer de supporter (2 #) à l'aide de l'un des principes suivants :

(1) *Le principe du naturalisme expérientiel*. Il affirme que si l'on a deux situations dans lesquelles la même cause produit un certain type d'effet dans une situation et un autre type d'effet dans autre , les conditions dans lesquelles la cause intervient doivent être différentes. Par exemple, le grattage d'une allumette génère généralement le feu; mais cet effet ne se produit pas si l'allumette est rayée dans une atmosphère privée d'oxygène, ou si l'allumette

est mouillée. Dans le cas d'une hallucination produite par la même activité cérébrale, le principe prévoit que l'effet qui s'en suit de l'activité cérébrale dans l'hallucination ne peut différer de l'effet qui s'en suit de l'activité cérébrale dans le cas de la perception que si des différentes conditions contextuelles sont en place. Mais il est évident qu'il y a des différences entre le cas perceptif et hallucinatoire : la présence ou l'absence de l'objet externe ou d'un événement. Étant donné que la même activité cérébrale se produit dans des conditions différentes, il est possible que les effets qui en résultent diffèrent dans toutes sortes de façons, y compris le caractère phénoménal de l'expérience qui en découle. L'acceptation de ce principe indéniable, donc, ne garantit pas (2#).

(2) *Le principe de la survenance locale.* Selon ce principe, un certain type d'activité cérébrale est suffisant pour la réalisation d'un état mental conscient. Il faut remarquer que l'acceptation de la survenance locale aurait une conséquence bien plus grave pour le relationnaliste que de supporter (2#). Si ce principe était vrai, on devrait accepter qu'une hallucination causée par le même état cérébral d'une perception ait le même caractère phénoménal. Cela rendrait l'argument de l'hallucination superflu. Mais, la survenance locale est une thèse controversée, pour laquelle, comme je montre dans ce chapitre, il n'existe aucun soutien empirique. De plus, l'acceptation de la survenance locale présuppose (2#) : souvent ses partisans citent (2#) comme une prémisse dans un argument à faveur de la survenance locale. Il faut donc chercher ailleurs pour motiver (2#).

(2) Le principe de la «même cause immédiate, même effet immédiat». Celui-ci peut être considéré comme un développement du principe de la mécanique newtonienne où il n'y a pas d'« action à distance ». Cependant, le principe de la «même cause immédiate, même effet immédiat » va plus loin que cela. Le rejet de la possibilité de la soi-disant action à distance prédit que, si un événement distal doit compter comme cause d'un effet, cela doit se faire par un enchaînement de causes qui soient spatialement et temporellement contiguës. Cela signifie qu'une fois que la cause distale a produit une cause intermédiaire dans un processus causal, celui-ci devient redondant : l'efficacité de l'événement placé en haut dans la chaîne est préempté par l'efficacité de la cause proximale, de sorte que si la cause distale était enlevée, et la cause immédiate maintenue en place, le même effet serait préservé. Le principe de la «même cause immédiate, même effet immédiat» ajoute quelque chose en plus. Elle indique que la cause immédiate est suffisante pour l'apparition de l'effet. Cela nécessite l'acceptation de la thèse supplémentaire (qui ne figure pas dans l'idée de rejeter l'action à distance) que rien sauf la cause proximale ne peut contribuer à l'effet et que rien ne peut interférer avec la cause proximale et l'empêcher de produire l'effet qu'elle produit en général.

Si nous voulions appliquer ce principe au cas expérimental, le principe de la «même cause proximale, même effet immédiat », on aurait que la reproduction d'une certaine activité cérébrale donne lieu nécessairement à la même expérience, indépendamment de toute autre condition applicable à la situation. Bien que l'acceptation du principe de la «même causalité proximale, même effet immédiat» soit moins exigeante que le principe de la survenance

locale, car elle ne prévoit pas qu'une certaine activité cérébrale est suffisante pour une occurrence mentale donnée dans tous les mondes possibles, cela donnerait, aux fins de notre discussion, le même résultat que l'acceptation survenance locale: cela établirait que l'activité cérébrale est suffisante pour l'occurrence d'un phénomène mental (étant données les lois de la nature courantes). Par conséquent, ce principe, comme celui de la survenance locale, permettrait non seulement de justifier (2#), mais offrirait la base pour un argument direct contre le relationnalisme, un argument qui ne repose pas sur l'IND. C'est l'argument causal de l'hallucination :

(1) Lorsque l'on perçoit un objet O, la présence d'O est une cause distale, tandis que l'activation de certains neurones est la cause proximale, de l'effet E, à savoir l'apparition d'une expérience d'O.

(2) Il est possible de stimuler directement le cerveau de manière à reproduire exactement l'activation neuronale qui advient lorsque l'on perçoit O, en l'absence d'O.

(3) Principe de la même cause immédiate, même principe d'effet immédiat.

(4) De (1) et (3) : L'activation de certains neurones est suffisante pour l'apparition de l'effet E dans le cas d'une véritable perception O.

(5) De (2) et (4) : dans le cas où l'on stimule directement le cerveau, et il n'y a pas d'objet distal O, l'activation de certains neurones est suffisante pour l'apparition de l'effet E, à savoir la survenue d'une expérience de O.

(6) De (4) et (5) : le même effet E se produit indépendamment de la présence ou de l'absence de l'objet distal O.

(7) Si le même effet E se produit dans la situation où l'on perçoit un O et quand on a une hallucination d'un O induite par la stimulation neuronale directe, la perception et l'hallucination correspondantes sont de la même nature.

Cependant, le principe de la même cause immédiate, même principe d'effet immédiat n'est pas acceptable. Le relationnaliste pourrait toujours argumenter que l'objet distal détermine de manière non-causale la nature du caractère phénoménal de la perception. Ce principe doit être rejeté et ne peut donc pas motiver (2#).

La première leçon à tirer de cette discussion, est, alors, que la version la plus prometteuse de l'argument de l'hallucination contre le relationnalisme reste l'argument traditionnel, dans lequel l'IND joue un rôle central. Il demeure donc crucial d'établir si l'IND est justifiée pour décider si le relationnalisme est tenable. Les opposants au relationnalisme devraient fournir des motivations pour l'IND, et les relationnalistes doivent expliquer comment ils pourraient l'accueillir. Ceci est en contraste avec une tendance récente dans le débat sur l'argument de l'hallucination, qui tend à prendre l'argument traditionnel vicié, tout en se concentrant sur une version de l'argument causal de l'hallucination, considérée comme la véritable menace pour le relationnalisme. Si mon diagnostic de la survenance locale et du principe de la «même cause immédiate, même effet immédiat» est correct, ces principes sont cependant plus controversés et moins bien motivés que l'IND. Par

conséquent, les stratégies alternatives qui argumentent contre le relationnalisme sans compter sur l'IND sont infructueuses.

La deuxième leçon à tirer de la discussion menée dans ce chapitre est que les raisons implicitement ou explicitement mises en avant pour motiver (2#) – le naturalisme expérientiel, la survenance locale et le principe de la «même cause immédiate, même effet immédiat» – est que soit elles ne peuvent pas justifier (2#) soit qu'elles -mêmes ne sont pas justifiées. Toutefois, cela ne signifie pas que nous devrions rejeter (2 #). Afin de le rejeter, il faut, soit affirmer que, quand on reproduit la même activité cérébrale que l'on a dans une perception en l'absence de l'objet vu, on n'a pas d'expérience consciente. Mais cela serait incroyable. On pourrait alors affirmer que l'on dispose d'une expérience consciente, mais qui n'est pas impersonnelle et indiscernable de la perception causalement correspondante (c'est à dire, on n'a pas une HIII).

Mais que peut-on dire pour soutenir cette affirmation ? On a vu dans le chapitre précédent que les hallucinations sont HIII si l'incapacité de distinguer l'hallucination par la perception ne peut pas être attribuée à une interférence, ayant à faire avec le sujet ou avec les conditions d'introspection. Dans les cas d'hallucinations ordinaires considérées dans le chapitre précédent, nous avons toujours été en mesure d'identifier certains facteurs qui pourraient être responsables de l'échec du sujet à la discrimination : une capacité ou une attention cognitive et discriminatoire affaiblie par une substance psychoactive, une injure au cerveau, une condition médicale, etc. De plus, quelle que soit l'interférence qui se produit, cela est susceptible de correspondre à une certaine différence neurologique entre l'hallucination en question et toute autre

perception hypothétique. Mais lorsque nous introduisons la possibilité d'hallucination qui correspond causalement à une perception, on exclut à priori toute interférence de ce type, qui pourrait expliquer pourquoi le sujet ne parvient pas à distinguer son hallucination d'une perception. Il faudrait indiquer d'autres conditions ou facteurs qui interfèrent avec la capacité du sujet à discriminer deux expériences autrement introspectivement distinctes. Cependant, on ne sait pas ce que ces facteurs pourraient éventuellement être.

Par conséquent, le relationnaliste ferait mieux d'accepter (2#) et de chercher ailleurs des raisons pour bloquer l'argument de l'hallucination. Dans le chapitre suivant, je vais envisager une ligne de réponse qui peut être trouvée dans la discussion d'Austin sur l'argument de l'illusion.

Sixième Chapitre

On a vu dans le Cinquième chapitre que, après tout, la possibilité métaphysique des HIII doit être acceptée. On a vu dans le Troisième chapitre que les explications mobilisées par le relationnaliste pour rendre compte du fait que la perception et l'hallucination ont des caractères phénoménaux différents, lorsqu'il s'agit d'hallucinations qui ne sont pas indiscernables de manière impersonnelle, ne sont pas disponibles dans les cas de HIII. Il semblerait donc que le relationnaliste doit affirmer que la nature du caractère phénoménaux d'une HIII est inaccessible à l'introspection. Fait qui est inacceptable, car dire cela, violerait la superficialité du caractère phénoménaux.

Dans ce chapitre j'explore si le relationnaliste peut relever le défi que la superficialité des caractères phénoménaux lui pose : peut-il expliquer les

HIII sans accepter l'IND, tout en préservant l'idée que leur caractère phénoménal est superficiel? Si le relationnaliste ne répond pas à ce défi, on aura raison d'accepter que la seule manière cohérente d'expliquer comment une hallucination peut être impersonnellement indiscernable par l'introspection est d'accepter l'IND et dire qu'elle doit avoir le même caractère phénoménal qu'une perception correspondante.

Avant d'aborder directement cette question, je considère si une stratégie différente pour bloquer l'argument de l'hallucination. On pourrait affirmer que, même si nous acceptons les IND et la possibilité des HIII, la conclusion de l'argument ne suit pas toujours. On peut accepter que le même caractère phénoménal peut être réalisé par des différentes bases. On peut alors accepter que la perception et les HIII ont le même caractère phénoménal, mais cela ne signifie pas que, dans le cas de la perception, ce caractère phénoménal soit déterminé par les objets que l'on voit.

Cependant, cette proposition fait face à un problème que Martin (2004) appelle « screening off problem », ou problème de l'exclusion explicative. Ce problème dérive d'une variation de l'argument causal de l'hallucination considérée dans le chapitre précédent. Le problème du screening off est la suivante : ce qui suffit à déterminer le caractère phénoménal dans le cas hallucinatoire devrait être suffisant pour le déterminer dans le cas de la perception aussi. Cela rend toute référence à l'objet redondant, dès qu'il s'agit d'expliquer la nature du caractère phénoménal de la perception.

Une conséquence du problème du screening off est que la seule stratégie qui reste à disposition du relationnaliste est celle de rejeter l'IND.

Une autre conséquence est que le relationnaliste ne peut pas expliquer le caractère phénoménal de l'hallucination en indiquant des propriétés positives. Le relationnaliste doit adopter une théorie négative de l'IIIH : le caractère phénoménal d'une IIIH n'est déterminé que par le fait qu'il est impersonnellement indiscernable de la perception par l'introspection.

Non seulement la thèse négative de l'hallucination bloque le problème du screening off mais il fournit également une réponse à l'argument de la superficialité. Le relationnaliste peut expliquer la différence de caractère phénoménal d'une perception et d'une HIII sans nier la superficialité des caractères phénoménaux, car la nature du caractère phénoménal d'une HIII semble être déterminé par quelque chose qui n'est pas accessible à l'introspection. Le caractère phénoménal d'une HIII est différent de celui d'une perception, car il ne dépend pas de l'obtention d'une relation psychique aux objets qu'on perçoit. Cependant, il n'a pas une nature cachée : tout ce qui est du caractère phénoménal est le fait que l'expérience ne peut être distinguée d'une situation dans laquelle on perçoit pour de vrai.

Si le relationnaliste ne doit pas affirmer que le caractère phénoménal d'une HIII est superficiel, le promoteur de l'argument de l'hallucination perd la seule bonne motivation qu'il avait pour soutenir l'IND. Ainsi, l'argument de l'hallucination faillit et le relationnalisme est préservé.

Cependant, de nombreuses objections ont été adressées à la thèse négative de l'hallucination. Je les passe en revue et fais valoir qu'un élément crucial dans la réponse à bon nombre de ces objections est de comprendre comment le relationnaliste doit concevoir l'introspection. Le relationnaliste

doit rejeter une conception de l'introspection selon laquelle le caractère phénoménal de son expérience est quelque chose dont on est conscient de façon indépendante de la perspective introspective d'ordre supérieur qu'on peut avoir sur elle. Selon cette image, quant on formule un jugement introspectif sur le caractère phénoménal, on acquiert une perspective nouvelle sur quelque chose dont on est déjà conscient indépendamment de ce jugement.

Une telle hypothèse implique que l'on est conscient du caractère phénoménal d'une expérience avant et indépendamment de sa conscience perceptive des objets de perception. Mais la discussion sur la transparence de l'expérience menée dans le chapitre 1 suggère le contraire. Nous avons vu que, bien que nous puissions prendre conscience des caractéristiques de l'expérience consciente elle-même, quand nous dirigeons notre attention introspective vers elle, nous sommes conscients de ces caractéristiques qu'en vertu d'être au courant des propriétés de l'objet que nous percevons (ou qu'il nous semble de percevoir). Il ne nous semble pas d'être en mesure de s'engager dans une activité introspective dirigée vers le caractère phénoménal de sa propre expérience perceptive sans porter notre attention sur les objets dont on est perceptuellement conscient. Comme Soteriou (2013) le remarque, cela suggère que l'activité d'introspecter son expérience dépend de l'activité de porter son attention perceptive sur les objets que l'on voit.

Ce qui distingue entre regarder quelque chose et introspecter une l'expérience, sont les objectifs et les résultats de cette activité. Lorsqu'on introspecte, on vise à acquérir des connaissances non seulement sur ce qu'il y a

devant nous, mais sur la façon dont les choses nous semblent perceptivement. Ce qu'on acquière, ce sont donc des croyances concernant l'effet que cela fait de voir ce que l'on a devant, plutôt que des croyances sur ce qu'il y a devant.

Quand on hallucine il n'y a aucun objet de perception sur lequel on peut porter notre attention. Cependant, on se comporte comme si l'on portait notre attention perceptive sur quelque chose. Donc, même si un objet n'existe pas, on se comporte comme s'il existait, et l'activité d'introspection que l'on dirige vers une hallucination est parasitaire de l'activité de se comporter comme si l'on portait l'attention sur un objet de perception. Ceci explique pourquoi, comme soutient le disjonctiviste, quand on hallucine, on se trompe par rapport à l'accès perceptif le sujet semble avoir à l'environnement, mais on se trompe aussi par rapport à l'accès qu'il nous semble avoir à sa propre expérience. Le type d'accès introspectif au caractère phénoménal de l'hallucination semble être le genre d'accès introspectif qui requiert que l'on regarde les objets perçus, ce qui est vrai seulement dans le cas d'une perception véridique. C'est en ce sens qu'il faut comprendre l'idée que le caractère phénoménal d'une hallucination est dérivé de la perception : c'est parce qu'il semble qu'on ait une perception, et donc un point de vue sur le monde, que l'on peut se livrer à l'activité de porter son attention introspective sur l'expérience elle-même.

Il faut comprendre les critiques à la thèse négative de l'hallucination sur le fond d'un désaccord au sujet de (1) la relation entre le caractère phénoménal et l'introspection et (2) la relation entre avoir une perspective perceptive sur le monde et avoir une perspective introspective sur une

expérience. Les opposants à la thèse négative supposent que (1) l'on peut être immédiatement conscient du caractère phénoménal d'une expérience d'une manière indépendante de l'accès introspectif d'ordre supérieur que l'on a sur ce caractère phénoménal. Ils sont amenés à penser ainsi parce qu'ils croient que (2) être perceptivement conscient de quelque chose exige que l'on soit (d'une façon immédiate) introspectivement conscient du caractère phénoménal de sa propre expérience. Au contraire, le relationnaliste fait valoir que (1) le caractère phénoménal n'existe pas indépendamment de la conscience introspective de celui-ci et que (2) avoir un accès introspectif au caractère phénoménal d'une expérience exige que l'on ait un accès perceptuel à son objet.

Si l'on adopte la conception de l'introspection que je développe dans ce chapitre, le relationnaliste peut (1) contrer les objections adressées à la thèse négative de l'hallucination et (2) expliquer comment la perception et les HIII peuvent avoir des caractères phénoménaux sans supposer que le caractère phénoménal de l'hallucination ne soit pas superficiel, et donc (3) rejeter l'IND et l'argument de l'hallucination. Cela complète ma défense du relationnalisme.

Conclusions

Le travail de recherche que j'ai mené, permet de dissiper plusieurs idées fausses qui sont couramment répandues dans le débat sur le relationnalisme et sur l'argument de l'hallucination.

Tout d'abord, la thèse relationnaliste qui soutient que le caractère phénoménal de la perception est en partie constitué par les objets qu'il nous

semble percevoir est souvent considérée avec suspicion. On croit que le caractère phénoménal de l'expérience ne peut pas incorporer les objets, car ceux-ci vont au-delà de la portée de ce que la notion de caractère phénoménal a été conçue pour capturer : l'aspect de l'expérience accessible à l'introspection, qui par définition fait abstraction de l'existence d'objets devant nous. Dans le chapitre 1 j'ai soutenu que dans la notion de caractère phénoménal rien n'exclut la position adoptée par le relationnalisme. Au contraire, la transparence de l'expérience et sa relationalité (du moins apparente) soutiennent le relationnalisme. En faisant fond sur le modèle d'introspection recommandé dans le chapitre 6, selon lequel avoir un point de vue introspectif sur le caractère phénoménal de l'expérience dépend de la présence d'un point de vue perceptif sur le monde, il est plus clair pourquoi il est trompeur de penser que le caractère phénoménal de l'expérience devrait faire abstraction des objets que l'on perçoit, d'une manière qui empêche un compte relationnaliste.

Une autre source de confusion dans le débat actuel porte sur la nature et la structure de l'argument qui vise à menacer le relationnalisme. Les discussions récentes se sont concentrées sur la version causale de l'argument de l'hallucination, que j'ai exploré dans le chapitre 5. Cet argument découle de l'hypothèse que les hallucinations qui ont une cause neuronale identique à une perception sont possibles : à savoir, il est possible de reproduire l'état neuronal dans lequel on se trouve en percevant un objet, en l'absence de l'objet. En vertu du lien entre le cerveau et l'esprit, cet argument conclut que l'activité du

cerveau produit exactement la même expérience que l'on a lorsque l'on perçoit, qui ne peut donc pas dépendre de l'existence de l'objet.

J'ai soutenu que cet argument ne tient pas, car il repose sur l'hypothèse que l'activité cérébrale d'un certain type est suffisante pour l'apparition d'une expérience consciente d'un certain type. Mais cette hypothèse ne peut être motivée que par le principe de survenance local, ou bien par le principe « même cause immédiate, même effet immédiat », qu'on ne peut pas accepter.

Au maximum, ce genre de considérations sur la possibilité d'avoir des hallucinations produites par le même état cérébral qu'une perception, peuvent motiver l'adoption d'une thèse négative de l'hallucination. Toutefois, ainsi qu'on l'a vu dans le chapitre 6, cela ne pose aucun problème pour le relationnalisme. La thèse négative de l'hallucination est souvent décrite comme une solution *ad hoc* importune. Au contraire, une vision négative de l'hallucination est précisément ce qui est nécessaire pour résoudre le problème crucial que les hallucinations représentent pour le relationnalisme : nier le principe d'indiscernabilité (IND), sans nier la superficialité des caractères phénoménaux.

On comprend pourquoi, en dépit du fait que les principes de la « même cause, même principe de l'effet » et de la survenance locale soient clairement controversés, les récentes discussions sur le relationnalisme ont mis l'accent sur ceux-ci, plutôt que de se concentrer sur l'argument traditionnel d'une hallucination reposant sur l'IND. Comme je l'ai soutenu dans le chapitre 2, cet argument traditionnel est un raffinement de celui sur les illusions et sur les apparences contradictoires. En dépit d'être largement et souvent tacitement

supposé, les adversaires de relationnalisme ont jusqu'à présent failli de présenter aucun argument convaincant pour soutenir son acceptation.

Le chapitre 3 comble cette lacune cruciale et identifie l'intuition sur laquelle l'IND repose : l'intuition que, en vertu de sa superficialité, la nature d'un caractère phénoménal doit être accessible par l'introspection. Cadrer le débat sur le relationnalisme en termes d'un désaccord sur la superficialité de caractère phénoménal et sa relation à l'introspection m'a permis d'identifier ce qui est vraiment en jeu dans le débat sur le relationnalisme, quel est vraiment le défi que les hallucinations posent au relationnaliste et quelles sont les ressources que le relationnaliste peut déployer pour faire face à ce défi.

Concevoir l'argument de l'hallucination en ces termes, fait en sorte qu'il est beaucoup plus difficile de ce qu'on considère généralement de formuler l'argument de manière satisfaisante, et de supporter toutes ses prémisses. En particulier, l'argument de l'hallucination ne peut réussir sans une bonne discussion sur ce qui fait en sorte qu'une hallucination compte comme indiscernable de l'hallucination. Il faut considérer comment différentes notions d'indiscernabilité permettent de différentes explications de l'indiscernabilité. En fin de compte, je l'ai soutenu, l'argument de l'hallucination n'a de chance de succès que si son promoteur fait place à la notion d'indiscernabilité impersonnelle. Cela inverse les termes du débat actuel, où il est souvent supposé que la thèse négative de l'hallucination doit accepter une notion d'indiscernabilité impersonnelle, et une telle notion est considérée comme problématique. J'ai soutenu qu'il y a une notion viable d'indiscernabilité impersonnelle par l'introspection, mais même s'il n'y en avait pas, ce serait

d'abord et avant tout un problème pour le promoteur de l'argument de l'hallucination, plutôt que pour le promoteur de la thèse négative de l'hallucination.

Un autre fait important souvent négligé et que j'ai exposé est que les hallucinations ne sont pas une classe uniforme d'occurrences mentales et elles n'ont pas nécessairement le même type de caractère phénoménal. La seule chose qu'elles ont en commun est qu'elles ont une certaine ressemblance avec la perception et qu'elles sont souvent confondues avec cette dernière. Plus important encore, toutes les hallucinations ne posent pas de problème pour le relationnalisme. Celles qui posent un problème pour le relationnalisme sont seulement celles qui sont impersonnelles et indiscernables de la perception par l'introspection (HIII). Comme je l'ai soutenu dans le chapitre 4, il n'y a pas encore de support empiriques disponibles pour démontrer qu'aucune hallucination documentée correspond à cette condition.

Cela dissipe une autre idée fautive qui entrave le débat sur le relationnalisme : le relationnalisme ne doit pas être évalué sur la base de sa capacité à rendre compte de la phénoménologie des hallucinations ordinaires, car le relationnalisme est une thèse sur le caractère phénoménal de la perception véridique, et n'importe quelle théorie des hallucinations ordinaires est compatible avec cela. Ce serait un projet de recherche intéressant de voir comment on devrait rendre compte de ces hallucinations ordinaires. Compte tenu de la discussion que j'amène dans cette thèse, on peut prédire avec confiance qu'il n'y aura pas d'explication « passe-partout » disponible pour toute hallucination. Mais ce projet de recherche va bien au-delà de la portée de

cette thèse, à savoir l'évaluation du relationnalisme, et cela est indépendant de celui-ci.