

Copyright
by
Brian Christopher Cutter
2015

The Dissertation Committee for Brian Christopher Cutter certifies that this is the approved version of the following dissertation:

Sensory Experience and the Sensible Qualities

Committee:

Michael Tye, Supervisor

Adam Pautz, Co-Supervisor

Ian N Proops

Richard M Sainsbury

Robert C Koons

Alex Byrne

Alexander Pruss

Sensory Experience and the Sensible Qualities

by

Brian Christopher Cutter, B.A.

Dissertation

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Doctor of Philosophy

The University of Texas at Austin

May 2015

Acknowledgements

Thanks are owed first to my supervisors, Michael Tye and Adam Pautz, for years of encouragement, conversation, and feedback. Both are extraordinary teachers and writers. Their influence on my thinking has been enormous and should be evident throughout my dissertation to anyone familiar with their work. I would also like to thank my other committee members—Robert Koons, Mark Sainsbury, Ian Proops, Alex Byrne, and Alexander Pruss—for the time and energy each of them put into reading and commenting on my work. Robert Koons has been especially generous with his time over the years, and I have benefitted immensely, both intellectually and personally, from our conversations and from his work. Miriam Schoenfield, Josh Dever, Michelle Montague, Galen Strawson, Hans Kamp, and Jon Litland also deserve thanks for very helpful feedback on various portions of the dissertation. I am grateful for many helpful conversations with my fellow UT graduate students, including especially Richard Davis, Alex Grzankowski, Jonathan Drake, Katherine Piatti, Derek Anderson, Alex Grossman, Enrico Grube, Brian Miller, Jonathan Vanderhoek, Casey Woolwine, and Jeremy Evans. I owe a special debt of gratitude to my undergraduate mentor, Patrick Frierson, for encouraging me to pursue philosophy at the graduate level and, more importantly, for being a model of intellectual and moral virtue. Thanks also to Jesse Hinde, Keefe Piper, and my parents for inspiring in me the love of wisdom. Finally, and most importantly, I would like to thank my wife Rand for her untiring and unconditional love and support and for giving levity to my life.

Sensory Experience and the Sensible Qualities

Brian Christopher Cutter, Ph.D.

The University of Texas at Austin, 2015

Supervisor: Michael Tye

Co-Supervisor: Adam Pautz

My dissertation defends a package of interrelated positions on the metaphysics of the sensible qualities (shape, color, pitch, loudness, flavor, heat, cold, etc.) and sensory experience. It is organized around four questions at the core of philosophical theorizing about the sensible qualities. The first is the question of *reductionism*: are the sensible qualities reducible to either physical properties (i.e. properties definable in the canonical vocabulary of the physical sciences) or response-dependent properties (e.g. Lockean dispositions to affect perceivers in certain ways)? I put forward novel arguments and refined versions of traditional arguments in support of a negative answer to this question. For at least some of the sensible qualities, including many of those traditionally classified as “secondary qualities,” reductionism is untenable.

If I am correct that the sensible qualities are not reducible to physical or response-dependent properties of external objects, the next question arises: do they belong to external objects at all? This is the question of *realism*. Many philosophers have held that a negative answer to the question of reductionism leads—or should lead—to a negative answer to the question of realism. Against these philosophers, I defend an affirmative

answer to the question of realism and respond to arguments from non-reductionism to irrealism.

If I am correct that the sensible qualities really belong to external objects but aren't reducible to any of their physical properties, a third question arises: how are the sensible qualities (especially the so-called "secondary qualities") related to physical reality? This is the question of *integration*, a special case of the more general question of how, in Sellars's terminology, the Manifest Image is related to the Scientific Image. In response to this question, I develop and defend a theory structurally parallel to Russellian monist positions on the mind-body problem. I argue that the Russellian monist framework is actually poorly suited to answer the question it was originally designed to answer—the question of how conscious experience is related to physical reality—but well suited to answer the corresponding question about the sensible (especially secondary) qualities.

Table of Contents

Chapter 1: Introduction	1
1.1 The Manifest Image and the Scientific Image	1
1.2 A Summary of Things to Come	5
Chapter 2: Against Relationism	13
2.1 Introduction	13
2.2 Intrinsicity and Categoricity	14
2.2.1 Revelation?	20
2.2.2 The Look of a Lack or the Lack of a Look?	22
2.2.3 Confusing Color Qualia with Colors?	24
2.3 The Argument from Modal Relations	28
2.3.1 Preliminaries	29
2.3.2 Against Inegalitarianism	30
2.3.3 From Egalitarianism to Non-Relationism	43
2.3.4 Privileged Status	45
2.4 Perceptual Variation	46
2.4.1 The Argument from Perceptual Variation for Relationism	47
2.4.2 The Argument from Perceptual Variation Against Relationism	58
Chapter 3: Against Physicalism	74
3.1 Introduction	74
3.2 The Structure-Mismatch Objection	75
3.3 The Categoricity Objection	76
3.4 Disjunctive Property Objections	77
3.5 The Similarity-Grounding Argument	80
3.5.1 Color Physicalism Leads to Color Disjunctivism	81
3.5.2 Grounding	84
3.5.3 Similarity	87
3.5.4 An Aside on Naturalness	90
3.5.5 Color Shades are Similarity-Grounding Properties	91

3.5.6	Disjunctive Properties do not Ground Similarity	92
3.5.7	Structured Properties and Disjunctive Properties	94
3.5.8	Extrinsic Similarity Grounding?	98
Chapter 4:	Realism	104
4.1	Introduction	104
4.2	The Empirical Justification for Realism	107
4.3	The Perceptual Argument	109
4.4	Premise PA1	111
4.4.1	Internalist Theories of Justification: Dogmatism and Rationalism	113
4.4.2	Externalist Theories of Justification: Process Reliabilism	115
4.4.3	Abductivism/Classical Foundationalism	120
4.5	Objections to (Non-Reductive) Realism: Skeptical Challenges	126
4.5.1	Against S1	131
4.5.2	A Challenge to S3	138
4.5.3	Evolution, Safety, and Luck	139
4.6	Objections to (Non-Reductive) Realism: Causation and Simplicity	144
4.7	Beyond Color	147
4.7.1	Non-Spatial Qualities	148
4.7.2	Bodily Sensations	151
4.8	Shape and Special Relativity	155
4.8.1	Edenic Content and Edenic Shape	156
4.8.2	Chalmers's Case Against RES	158
4.8.3	A Defense of RES	160
4.9	The Argument from Phenomenal Relationism	169
Chapter 5:	Russellian Monism about the Secondary Qualities	174
5.1	The Non-Reductive Realist's Dilemma	174
5.2	Secondary Quality Russellian Monism	179
5.3	The Non-Reductive Realist's Dilemma Revisited	186
5.4	Considerations Favoring PRM Provide More Support for SQRM	189

5.5 The Happy-Synthesis Argument	151
5.5.1 Advantages of Happy-Synthesis Argument for SQRM	194
5.6 The Argument from Qualitativeness	199
5.7 The Argument from Categoricity	201
5.8 Conclusion	205
References	206

Chapter 1: Introduction

1.1 The Manifest Image and the Scientific Image

Many of the philosophical questions that have occupied philosophers since the advent of modern science are questions arising from tensions between two images of the world—what Sellars (1963) calls the “Manifest Image” and the “Scientific Image.” Roughly speaking, the Manifest Image is the image or description of the world suggested to us by ordinary experience, whereas the Scientific Image is the image or description of the world suggested to us by modern science. *Prima facie* tensions between these two images abound. To mention a few:

Free will: Our ordinary experience of deliberating and acting suggests that we sometimes act freely—that we sometimes freely choose among a range of mutually incompatible actions, each of which is, at some point just prior to action, within our power to perform. On the other hand, scientific accounts of human behavior from psychology and neuroscience suggest that our behavior is caused by prior conditions over which we have no control. Such descriptions seem to omit any mention of an agent’s powers to do otherwise than he actually did, and seem, at first blush anyway, to be in conflict with the assumption that agents have such powers.

Intentionality: The Manifest Image portrays us as having intentional states—states that are *about* things in the world, states that *predicate* properties of objects, states that are apt to be *true* or *false*. But such properties and relations as aboutness, predication, truth, and falsity seem to be absent from our physical or scientific descriptions of humans

and other animals. The conciliatory position—that each image offers an accurate, but incomplete, portrait of human beings—is tempting here, but it is also problematic for at least two reasons. First, it offends against ontological parsimony. Second, the intentional states of the Manifest Image seem to be in causal competition with the physical states of the Scientific Image. If the physical states attributed to us by the Scientific Image (e.g. by neuroscience) are distinct from the intentional states attributed to us by the Manifest Image, it's hard to see how the latter could have the sort of causal relevance the Manifest Image represents them as having. For the causal roles ascribed to intentional states by the Manifest Image are occupied, according to the Scientific Image, by (ostensibly non-intentional) physical states.

Conscious Experience: When your arm is sharply pinched, you feel a momentary pain—or so introspection (whose testimony partly constitutes the Manifest Image) suggests. But the scientific description of what goes on in such a case—beginning with the stimulation of nociceptors in the arm, followed by complex patterns of activity in somatosensory cortex, insular cortex, and anterior cingulate cortex, leading in turn to a pattern of signals from the motor cortex to muscles in the arm, resulting finally in your withdrawing your arm from the stimulus—seems to leave out entirely the *feeling of pain*. If the feeling of pain really exists, it seems we must accept that it is therefore something different from, something over and above, the processes and events described by the physical sciences. But this conciliatory position is problematic, again because it offends against ontological parsimony and also because the causal roles ascribed to the feeling of

pain by the Manifest Image (e.g. causing withdrawal behavior) are occupied, according to the Scientific Image, by (ostensibly non-experiential) physical processes and events.

The “Secondary Qualities”: Sensory experience (whose testimony partly constitutes the Manifest Image) represents tomatoes as red, strawberries as sweet, ice cubes as cold, and the squeaks produced by my floorboards as high pitched. But such qualities do not seem to show up in our scientific descriptions of the world. Science describes the surface of a tomato as having a certain complex molecular structure, a certain surface reflectance, certain powers to affect human nervous systems, and so forth. But *redness*—that vibrant quality I experience as spread out across the surface of the tomato when I visually behold the tomato—seems to be entirely left out of the scientific description. Once again, the conciliatory position, according to which the tomato really is red but its redness is something over and above its scientific properties, is tempting, but again problematic for reasons of ontological parsimony and causal exclusion, *inter alia*.

Other examples could easily be given from many other areas of philosophy, e.g. the philosophy of time (does time pass? is the present moment somehow more real than the past and future?), metaethics (are there normative facts, or just natural/descriptive facts?), the philosophy of perception (in perception, are we directly acquainted with external objects?), the philosophy of biology (do bodily organs or their activities genuinely have *purposes* or *functions*, or is all biological activity to be explained solely in terms of efficient causes?), and the philosophy of action (are human actions performed in response to reasons or merely produced by arational causes?).

The Grand Question that animates this dissertation is this: is it possible to accept the Manifest Image and the Scientific Image, both in their full vigor, without diluting either? I regard it as obvious that we should take each image seriously, that we should not dismiss either from the outset, and that we should not deny or dilute either without very good reason. We cannot embrace the Scientific Image while wholly dismissing the Manifest Image, for the latter is the epistemic foundation of the former. Nor can we embrace the Manifest Image while disparaging the Scientific Image—say, by adopting some sort of instrumentalist or anti-realist attitude toward science—for to disparage science is to disparage the common sense at the core of the Manifest Image. As T.H. Huxley (1888) wrote,

Science is [...] nothing but trained and organized common sense, differing from the latter only as a veteran may differ from a raw recruit: and its methods differ from those of common sense only as far as the guardsman's cut and thrust differ from the manner in which a savage wields his club. (p. 77)

Of course, the Grand Question above is much too large a question to address adequately in a single work (or perhaps in a single lifetime). But a single work can address this question as it pertains to certain interesting fragments of the Manifest and Scientific Images. This is what I hope to do in this dissertation for those fragments of the Manifest and Scientific Images that pertain to the *sensible qualities*. “The philosopher,” Sellars tells us, “is confronted by two conceptions, equally public, equally non-arbitrary, of man-in-the-world and he cannot shirk the attempt to see how they fall together in one stereoscopic view” (p. 5). This dissertation attempts to present a stereoscopic view of the

relevant fragments of the Manifest and Scientific Images, a view in which neither image is marginalized and in which each contributes its full color.

1.2 A Summary of Things to Come

By the “sensible qualities” I mean those qualities which are presented to us in sensory experience and which we experience as belonging to the objects of sensory experience. At a minimum, they include shape, color, motion, texture, pitch, loudness, timbre, heat, cold, flavor, and scent. I shall focus especially, though not exclusively, on the so-called “secondary qualities.” which minimally include color, pitch, loudness, timbre, heat, cold, flavor, and scent. An important terminological note: here and throughout, I use the term “secondary quality” merely to denote those sensible qualities traditionally classified under the heading “secondary quality.” By calling a quality a secondary quality, I do *not* mean to suggest that it satisfies any of the substantive criteria that have historically been associated with that term, e.g. that it is not a “catholick affection,” that our idea of it does not resemble it, or that it does not inhere in material bodies. The same goes, *mutatis mutandis*, for my use of the term “primary quality.”

This dissertation is organized around three questions at the heart of philosophical theorizing about the sensible qualities. The first, which is addressed in chapters 2 and 3, is the question of *reductionism*: are the sensible qualities reducible to either response-dependent properties or response-independent physical properties? Here I argue that reductionism fails for many classes of sensible qualities, including (probably all, but at least many of) the secondary qualities. The varieties of reductionism about a given class of sensible qualities may be divided into two broad kinds, *relationist* theories and

physicalist theories. Relationist theories identify the relevant qualities with certain response-dependent properties. A response-dependent property may be understood roughly as property whose instantiations constitutively depend on the subjective responses of perceivers. Put another way, response-dependent properties are relational properties constituted by relations to perceivers or their sensory experiences. Paradigms of response-dependent properties include dispositional relational properties like the disposition to produce heat sensations in (or feel hot to) normal perceivers under normal conditions, or the disposition to produce reddish sensations in (or look red to) me in this specific viewing condition, as well as non-dispositional relational properties like the property of occurrently producing bitter sensations in (or tasting bitter to) some or other perceiver, or the property of occurrently producing greenish sensations in (or looking green to) me in this specific viewing condition. In chapter 2, I put forward four arguments against relationist theories of the secondary qualities. The first is the *argument from intrinsicity*, which derives the falsity of relationism from the apparent second-order fact that colors and many other secondary qualities are intrinsic properties. The second is the *argument from categoricity*, which derives the falsity of (not all forms, but) the most popular forms of relationism from the apparent second-order fact that colors and other secondary qualities are categorical properties. The third is the *argument from modal relations*, which concludes that relationism is false on the basis of (*inter alia*) an intuitively plausible principle about the modal relationship between color and spatial extension—viz., it is impossible for something to have a color without being spatially extended. Finally, after criticizing the well-known argument from perceptual variation *for*

relationism, I put forward an argument from perceptual variation *against* relationism, arguing that the empirical facts of perceptual variation, far from supporting relationism, actually constitute powerful evidence against relationism.

Physicalist theories of a given class of sensible qualities, in contrast to relationist theories, identify the qualities in question with response-independent physical properties. A response-independent property is a property that is not response-dependent. “Physical property” here is intended in a fairly narrow sense to mean a property that admits of real definition in terms of the canonical vocabulary of the physical sciences. (Properties that merely supervene on, but are not identical to, such properties do not qualify as “physical” in my sense.) Paradigms of response-independent physical properties include the surface spectral reflectance of a tomato, the fundamental frequency, amplitude, and wave shape of a pressure wave, the temperature and conductivity of a surface, and so forth. In chapter 3, I advance three arguments (or, better, three classes of argument) against physicalist theories of the secondary qualities. The first class of arguments is the class of *structural-mismatch arguments*. Structural-mismatch arguments begin with the empirical premise that there is no (relevant) class of physical properties that has the structural features (e.g. the resemblance structure) of a given class Q of secondary qualities, and from here conclude by Leibniz’s Law that physicalism about Q is false. The second argument is the *argument from categoricity*, which derives the falsity of (not all forms, but) some popular forms of physicalism from the apparent second-order fact that colors and other secondary qualities are categorical properties. The final class of anti-physicalist arguments I discuss are what I call *disjunctive-property arguments*. Applied specifically

to color physicalism, these are arguments with the following form: (i) If color physicalism is true, then color shades are disjunctive (physical) properties. (ii) Color shades have feature F. (iii) No disjunctive property has feature F. (iv) Therefore, color physicalism is false. I consider five promising instances of this argument form, four of which I discuss only briefly, and one of which (the *similarity-grounding argument*) I develop and defend at length.

Physicalism and relationism, though mutually exclusive, do not jointly exhaust the space of possible positions. The portion of logical space left uncovered by (the disjunction of) relationism and physicalism about a given class of sensible qualities corresponds to the view I call *non-reductionism*, according to which the the qualities in question are response-independent, non-physical properties.¹ The upshot of chapters 2 and 3 is that non-reductionism is true of the secondary qualities.

If I am correct that (at least many of) the sensible qualities aren't reducible to either response-dependent properties or response-independent physical properties of material objects, the question arises: are they instantiated in the ("external") material world at all? This is the question of *realism*, the topic of chapter 4. Here I defend the claim that realism is true about every major class of sensible qualities, where realism about a class of sensible qualities is the thesis that qualities in that class are instantiated

¹ Often the term "primitivism" is used to cover roughly this region of logical space. I prefer "non-reductionism" to "primitivism" because the latter suggests that the relevant qualities are unanalyzable or lack non-trivial real definitions. But in some cases this seems to rule out reasonable alternatives to physicalism and relationism. For example, suppose one holds that a certain shade of orange can be analyzed in terms of primitive hue magnitudes, e.g. being reddish to degree x and yellowish to degree y. On such a view, some determinate colors would have non-trivial real definitions and hence they would not be "primitive." Such an alternative to physicalism and relationism therefore seems to fit more comfortably under the heading "non-reductionism" than "primitivism."

by ordinary physical things, more or less as they seem to be. In this part of the dissertation, in addition to responding to a number of objections to realism, I put forward two positive reasons to accept realism, both of which apply to every major class of sensible qualities. The first is that the various sensible qualities *seem* or *appear* to be instantiated in our environment (and in our bodies). Here I rely on the prosaic idea that we ought to trust appearances in the absence of sufficiently strong countervailing reasons (cf. Huemer 2001, Pryor 2000), and I argue at length that there are no such reasons to reject realism. I respond to a range of objections to (non-reductive) realism about the secondary qualities, as well as a recent objection to realism about primary qualities (e.g. shape) from modern physics.

The second positive reason to accept realism comes from the independently motivated thesis of *phenomenal relationism*. According to phenomenal relationism, the phenomenal properties associated with sensory experience (e.g. the property characterizing what it's like for you to look at a lime in good lighting, or to taste black coffee, or to hear floor boards squeak) are relational properties of a certain kind. More specifically, according to phenomenal relationism, what it is for a subject to have a given (sensory) phenomenal property is for the subject to bear a certain relation—which might variously be called “sensory awareness,” “sensory acquaintance,” or “sensory representation”—to a certain sensible quality or complex of sensible qualities (Campbell 1993, Johnston 2004, Chalmers 2006, Pautz 2009, Byrne 2009, Tye 2014). I argue that, given phenomenal relationism, the assumption of realism about the sensible qualities makes the “hard problem” of consciousness much more tractable. The basic idea is this:

consider the phenomenal property you instantiate when you have a visual experience as of something green. Call this property “G.” Given phenomenal relationism, G is identical to (something like) the relational property *sensorily representing greenness*. Now, if color realism is false, then greenness is nowhere instantiated in the universe. In that case, we don’t bear any physical, causal, or informational relations to greenness. But if don’t bear any physical, causal, or informational relations to greenness, it’s hard to see how we could manage to stand in any interesting natural relations to greenness, such as is involved in sensorily representing greenness. On the other hand, if color realism is true, then greenness is instantiated in the world more or less as it seems to be. In that case, we do stand in various interesting physical, causal, and informational relations to greenness. In terms of these we can at least begin to sketch out possible explanations (either reductive or non-reductive) of how we come to sensorily represent greenness. Realism therefore helps us to explain how we come to enjoy phenomenal properties like G, whereas irrealism renders this fact an intractable mystery. Realism is therefore supported on abductive grounds.

Chapters 2-4, summarized above, jointly constitute a defense of a *non-reductive realism* (i.e., the conjunction of non-reductionism and realism) about the secondary qualities. A major theme of these chapters is that the reductionist and irrealist alternatives to non-reductive realism are inconsistent with the way things appear and so fail to respect the Manifest Image. More specifically, irrealism is inconsistent with the first-order appearances (apparently true propositions attributing (first-order) sensible qualities to concrete things in our environment), and reductionism is inconsistent with the second-

order appearances (apparently true propositions that attribute second-order properties to the sensible qualities themselves). If we are to respect appearances, we must accept non-reductive realism.

Now, if I am correct that the sensible qualities really belong to material objects but (at least many of them) aren't reducible to any physical properties of material objects, a third question arises: how are the sensible qualities (especially the secondary qualities) related to physical reality? This is the question of *integration*, a special case of the more general question of how the Manifest Image is related to the Scientific Image. I address this question in chapter 5. Here I develop and defend a theory structurally parallel to Russellian Monist positions on the mind-body problem, which I call *Secondary Quality Russellian Monism* (SQRM). At the foundation of Russellian Monism is the observation, due to Russell (1927), Eddington (1928), and others, that physics characterizes matter in terms of its relational structure, including especially its abstract causal and mathematical structure, but does not reveal its intrinsic qualitative nature. To a first approximation, SQRM says that the secondary qualities are grounded in the intrinsic qualities of matter that physics leaves unspecified, perhaps together with the structural properties in terms of which physics characterizes matter, but not in the latter properties alone. As I shall argue in chapter 5, non-reductive realism leads naturally to SQRM; SQRM is therefore indirectly supported by my arguments for non-reductive realism in chapters 2-4. In addition, I offer a handful of more-or-less independent arguments and considerations in favor of SQRM. A central theme of this chapter is that the Russellian Monist framework is actually poorly suited to answer the question it was originally designed to answer—the

question of how conscious experience is related to physical reality—but well suited to answer the corresponding question about the secondary qualities.

Chapter 2: Against Relationism

2.1 Introduction

In this chapter, I advance several arguments against relationist theories of the secondary qualities. It will be useful to focus on a representative example of a secondary quality in what follows; for this reason I focus heavily, though not exclusively, on relationist theories of *color*. However, many of the objections I raise to relationist theories of color apply equally to relationist theories of other secondary qualities. Moreover, it is antecedently plausible that if relationism is false about color, it is also false about the other secondary qualities. I'm not aware of anyone who accepts a non-relationist theory of color together with a relationist view of (say) heat, cold, flavor, or pitch. So if we can show that relationism about color is false, then it's a fair bet that relationist theories of other secondary qualities are false as well.

In §2.2, I briefly discuss two apparent truths about (at least some of) the secondary qualities that are inconsistent with relationism in all of its forms or in its most popular form. In §2.3, I argue that the modal relationship between color and shape—in particular, the fact that nothing can have a color without being spatially extended—makes relationist views of color untenable. In §2.4, I examine the relevance of perceptual variation to relationist theories of the secondary qualities. It is widely held that the phenomenon of perceptual variation supports relationism. On the contrary, I argue that the phenomenon of perceptual variation provides no support for relationism and in fact constitutes powerful evidence *against* relationism.

2.2 Intrinsicity and Categoricity

When, for example, a ripe tomato looks red to me, [...] I do not experience any part of the surface as producing a certain sort of response in me or anyone else. On the contrary, I surely experience redness as intrinsic to the surface, just as I experience the shape of the surface as intrinsic to it.

—Michael Tye (1995, p. 145)

What standard experiences of color do seem to suggest is that redness (e.g.) is intrinsic and categorial. [...] this much is true of lots of properties—roundness for instance.

—Stephen Yablo (1995, p. 490)

Things in our environment appear to have certain sensible qualities. For example, my desk appears to be brown, rectangular, hard, and smooth. These are *first-order* appearances—apparently true propositions that attribute first-order properties to concrete particulars. But not only do concrete particulars appear to have certain first-order properties; these first order properties themselves appear to have certain second-order properties. For example, roundness seems to be an intrinsic, categorial, and non-disjunctive property; it seems to be incompatible with the squareness, and it seems to entail spatial extension. These are *second-order appearances*—apparently true propositions that attribute second-order properties to first-order properties.

There is an interesting question about the manner in which sensible qualities appear to us to have these second-order features. One possibility is that they *perceptually* seem to have these features. On this option, we might say that perceptual experience, in addition to having first-order content which ascribes to concrete objects various first order properties like colors and shapes, also has second-order content which ascribes to these first-order properties various second-order properties like being intrinsic, being

categorical, etc. Some philosophers have repudiated this view of perceptual content. For example, Byrne and Hilbert (2003b) write, “lemons look yellow, but yellowness does not look any way at all” (p. 793, cf. Byrne 2003). Another possibility is that the sensible qualities *intellectually* seem to have certain second-order features upon rational reflection. On this option, the way in which roundness seems to be intrinsic or softness seems to entail spatial extension is similar to the way in which the identity relation seems to be transitive or knowledge seems to be factive. The proponent of this option may nonetheless maintain that the relevant seemings are intimately connected with perceptual experience, even if the sensible qualities do not perceptually seem to have these features. For example, it is plausible that in order for shapes, colors, or flavors to seem thus-and-so upon rational reflection, one must possess certain shape/color/flavor concepts (in a non-deferential manner), which in turn plausibly requires having experienced the relevant qualities (or, in the case of shapes, having experienced certain basic spatial qualities from which the relevant shapes may be constructed). For our purposes, however, it is not necessary to take a position on the manner in which the sensible qualities appear to have certain second-order properties.

It will be a recurrent theme of this dissertation that the alternatives to my non-reductive realist view of the secondary qualities are inconsistent with the way things appear. More specifically, I shall urge that irrealism is inconsistent with the first-order appearances, and reductionism is inconsistent with the second-order appearances. If we are to respect appearances, we must accept non-reductive realism. Of course, there is no general guarantee that things are as they appear to be, so the fact that irrealism and

reductionism are inconsistent with appearances is no guarantee that they are false.

Nonetheless, I take for granted in this dissertation what is sometimes called the Principle of Phenomenal Conservatism: if it seems to one that *p*, then one has at least *prima facie* justification to believe that *p* (Huemer 2001, cf. Pryor 2000). In other words, we ought to trust appearances in the absence of (sufficiently strong) countervailing considerations.

In this section, I discuss two apparent second-order truths about color, one of which is inconsistent with all forms of relationism, and one of which is inconsistent with the most popular form of relationism.

The first apparent second-order truth that bears on relationism is:

Intrinsicality: colors are intrinsic properties.²

As many have observed, we experience colors as intrinsic or local properties of the surfaces of objects (Chalmers 2006, p. 66; McGinn 1996, p. 542; Johnston 1992, p. 223; Yablo 1995, p. 489, Tye 1995, p. 145). Intuitively, an intrinsic or local property is a property whose instantiation by an object does not constitutively depend on the object's relations to things wholly distinct from itself. In this respect, colors seem to be like shapes. Just as the apparent shape of a golf ball seems to be an intrinsic feature of the golf ball, so too does the apparent color of the golf ball. As Tye (2000) remarks,

[O]ur ordinary experiences of color place (many) object colors on the surfaces of objects independently of what is going on elsewhere in the surroundings. In this respect, color seems to me like shape. We experience the redness of a ripe tomato

² I ignore so-called “unsteady” or “relational” colors, such as are found in holograms and glossy materials (Johnston 1992, p. 227)

as not involving anything *away* from the facing surface of the tomato, as being a *local* feature of that surface, just as we do its shape. To take a relational view of color is to repudiate this commonsense fact. (p. 153)

Colors appear to be intrinsic. And as Tye observes, this apparent truth is flatly inconsistent with any relationist theory of color. If we respect appearances, we will reject relationism.

The second apparent truth that bears on relationist theories of color is:

Categoricity: Colors are categorical (non-dispositional, non-hypothetical, manifest) properties.

The notion of a categorical property is, intuitively, the notion of a wholly *non-modal* property, a property the instantiation of which does not constitutively depend on what goes on in other possible worlds. Yablo (1992) helpfully points out a loose parallel between the notion of an intrinsic property and the notion of a categorical property: intrinsic properties stand to ordinary physical space roughly as categorical properties stand to modal space. The intrinsic properties of an object are, very loosely speaking, those of its properties that don't depend on what's going on elsewhere in ordinary physical space. Analogously, the categorical properties of an object are those of its properties that don't depend on what's going on elsewhere in *modal* space.

If we set aside the secondary qualities, perhaps the clearest examples of categorical properties are *shapes*. Contrast shape properties like *being triangular* or *being spherical* with paradigmatic hypothetical properties like *being dangerous*, *being poisonous*, or *being fragile*. If an object is triangular, then its being triangular is, as it

were, a situation wholly contained within actuality. Being triangular is not a matter of how the object *would* be or behave in counterfactual circumstances, but purely a matter of how it actually *is*. By contrast, if something is fragile, its being fragile is, in some intuitive sense, a situation not wholly contained within actuality. Being fragile is at least partly a matter of how the object *would* behave under certain counterfactual circumstances.

Just as colors seem to be like shapes in respect of being *intrinsic*, so too colors seem to be like shapes in respect of being *categorical*. As Dancy and Hookway (1986) remark, “the way in which color appears [...] seems to me at least to be stubbornly non-dispositional” (cf. Yablo 1995, p. 489; McGinn 1996, p. 545; Boghossian and Velleman 1989, p. 86). Kripke (1980) similarly remarks, concerning those who accept a Lockean dispositionalist version of color relationism, “I suspect many have been bothered by the ‘gut feeling’ that yellowness is a manifest property, just as much ‘right out there’ as hardness or spherical shape” (p. 140n). Yablo (1995), defending what he calls the “naive view” of color, writes, “But what is the argument that redness does not look to be what the naive view says it is: an intrinsic non-dispositional *sui generis* color property? This would seem to be *exactly* how it looks” (p. 489). Similar claims apply, I think, to all or nearly all the secondary qualities.

There are, then, good *prima facie* grounds for accepting Categoricality. How does this bear on relationism? Categoricality is not inconsistent with relationism as such, but it is inconsistent with the most popular version of relationism, namely *dispositionalist* relationism (Locke 1689/1996, Reid 1764/1983, McGinn 1983, Johnston 1992, Cohen

2009). According to dispositionalist relationism, the colors we experience objects to have in color perception are dispositions to produce color experiences of a certain kind in (or dispositions to look a certain way to) certain perceivers under certain conditions.³ Hence, the dispositionalist relationist will accept analyses along the lines of:

x is red (for such-and-such perceivers in so-and-so conditions) =: x is disposed to produce reddish sensations (in such-and-such perceivers in so-and-so conditions).

If we respect appearances, we will accept Categoricity and so reject dispositionalist forms of relationism. To accommodate the apparent fact that colors are categorical properties, the relationist must adopt some non-dispositionalist form of relationism. A natural suggestion is to identify colors not with dispositional response-dependent properties like the disposition to produce reddish sensations, but with categorical response-dependent properties like the property of occurrently producing reddish sensations.

In my view, this non-dispositionalist form of relationism, or something in the vicinity, is the most plausible form of relationism. But it is not very popular among relationists. I suspect that the main reason why relationists typically find a dispositionalist account more attractive than a non-dispositionalist account is that the former allows objects to retain their colors when no one is looking. The ripe tomato on my kitchen counter now has the disposition to produce reddish sensations in normal human perceivers under normal viewing conditions, even though it is not now producing any reddish sensations. The dispositionalist relationist can therefore accept the commonsense

³ Dispositionalist relationism is often just called “dispositionalism.” But the former name is preferable because many non-relationists about color also identify colors with dispositions, such as (response-independent) dispositions to alter light in certain ways (Byrne & Hilbert 2003a, Tye 2000, Matthen 1988).

claim that the tomato on my kitchen counter is red right now—that it has now, when no one is looking, the very color quality I experienced it as having this morning. But the non-dispositionalist relationist cannot accommodate this commonsense claim.

2.2.1 Revelation?

I have claimed that colors appear to have certain features and that absent good evidence to the contrary, we should trust these appearances. In §§2.3-4, I will argue that other apparent second-order features of colors raise further difficulties for relationist theories of color, and in §3, I argue that the apparent second-order features of color raise difficulties for physicalist theories of color as well. Some may worry that these claims bring me dangerously close to the much maligned doctrine of “Revelation,” according to which the intrinsic nature of any color is fully revealed by a standard visual experience of it (Johnston 1992, p. 223).

The misguided nature of this worry is fully revealed by a moment’s reflection on it. I suppose that second-order properties like *being intrinsic* and *being categorical* pertain to the intrinsic nature of those qualities that possess them. So I suppose I am committed to the view that the intrinsic natures of colors are at least *partially* revealed to us, either in visual experience or in rational reflection. But I am not committed to (and indeed I would reject) the view that the intrinsic natures of colors are *fully* revealed to us. Nor is there a slippery slope from the former to the latter, as an analogy with the perception of the first-order features of concrete objects makes clear. Perception informs us of some of the intrinsic features of objects in our environment, e.g. their shapes. In this way, the intrinsic character of an object is at least *partially* revealed to us in ordinary

perceptual experience. From this truism, though, there is no pressure to accept the first-order analogue of Revelation: that the intrinsic character of an object is *fully* revealed to us in perception, i.e. that perception informs us of *all* of the object's intrinsic features.

Let us give the name "Weak Revelation" to the thesis that *some* necessary or essential truths about colors are knowable on the basis of ordinary experiences of those colors. Even if I am not committed to the strong thesis of Revelation formulated above, it might be suggested that my commitment to Weak Revelation is somehow problematic. But such a suggestion would be misguided, for Weak Revelation cannot plausibly be denied. Consider the following claims:

R1. Red is more similar to orange than it is to green.

R2. Yellow is not identical to blue.

R3. If something is purple, then it is spatially extended (or at least spatially located).

R4. Nothing can be both yellow (all over) and blue (all over).

R5. Scarlet is not a shade of green.

R6. Red is a unitary color, and orange is a binary color.

It is plausible (to say the least) that at least some of R1-R6 are necessary or essential truths about color and can be known on the basis of ordinary color experience. If so, then Weak Revelation is true.

It may be tempting to deny that the self-evidence of R1-R6 is really a matter of these truths being revealed to us in ordinary experiences of the colors, maintaining instead that R1-R6 merely register facts about how we have collectively decided to use color words. For example, it might be suggested that the self-evidence of R2 and R4 only

reflects the fact that the way we use color words ensures that no pair of properties deserve the names “yellow” and “blue” unless they are distinct and mutually exclusive properties, just as the way physicists use their theoretical vocabulary ensures that no pair of properties deserve the names “unit positive charge” and “unit negative charge” unless they are distinct and mutually exclusive properties. But this suggestion is mistaken, as is evident when we observe that R1-R6 express facts which can be known on the basis of ordinary experience without making use of English color terms; they can be known, for example, under perceptual demonstrative concepts, as when we know, when looking at a yellow thing next to a blue thing, that *this* quality is not *that* quality (R2), or that *this* quality excludes *that* quality (R4), or when looking at a purple object, that *this* quality could not be possessed by a spatially unextended object (R3).

2.2.2 The Look of a Lack or the Lack of a Look?

Sometimes it is denied that colors look, or in any other way seem to be, intrinsic or categorical, as I have claimed, and that our temptation to think they do stems from a kind of scope confusion. In particular, from the fact that colors do not look relational or dispositional, we mistakenly come to think that colors look non-relational and non-dispositional (cf. Smart 1975, Tye 2009, p. 142, Armstrong 1968b, pp. 48-9). In other words, according to the present objection, we confuse the lack of a look for the look of a lack.⁴ More precisely, our mistake is owed to a fallacious shift in scope from

A1. It is not the case that colors seem to be relational (dispositional, disjunctive, . . .), wherein the negation takes wide scope, to

⁴ I owe this clever turn of phrase to Mark Sainsbury (personal communication).

A2. Colors seem not to be relational (dispositional, disjunctive, . . .),
wherein the negation takes narrow scope.

There are two serious problems with this scope-shift objection. The first is that the alleged fallacy is so flagrant that it is scarcely believable that one could commit it. An analogy: I happen to endorse the following *de re* modal claim:

B1. Necessarily, Socrates is not a cow ($\Box\neg\text{Cow}(s)$)

Perhaps I am mistaken, though. Perhaps Socrates has a thinner, less restrictive essence than I think. But if I am mistaken, it is not because I am unconsciously sliding from the evident fact that

B2. It is not the case that, necessarily, Socrates is a cow ($\neg\Box\text{Cow}(s)$),

wherein the negation takes wide scope, to B1, wherein the negation takes narrow scope. Like any philosopher, I am liable to make the occasional philosophical mistake. But I flatter myself that I do not make such egregious mistakes as this. The problem with the scope-shift objection is that it attributes to advocates of claims like A2 a precisely analogous mistake.

Perhaps it will be said that while we are not apt to make this kind of fallacious shift across some operators, e.g. “necessarily,” we are apt to make these fallacious shifts across operators like “seems” or “appears.” But we aren’t. It is not the case that the water in the glass before me visually appears to be 72 degrees Fahrenheit, or to be composed of exactly 10^{18} molecules. But there is no temptation to think that it visually appears not to be. The same holds for second-order appearances. When I attend to the shade of blue on

the curtains before me, it is not the case that this color seems to have the property of being Jones's favorite color. But there is no temptation to think that it seems not to be favored by Jones.

The second problem with the scope-shift objection is that it makes a mystery of why we make the specific "mistakes" we do. The proponent of the scope-shift objection maintains that

C1. It is not the case that colors seem to be intrinsic (categorical)

One wonders, then, why we would make the fallacious shift from A1 to A2 rather than making the parallel shift from C1 to

C2. Colors seem not to be intrinsic (categorical)

The scope-shift objection leaves it mysterious why we are disposed to believe A2 rather than C2. The most straightforward explanation for this fact is that, *pace* the scope-shift objection, A2 is true but C2 is not.

2.2.3 Confusing Color Qualia with Colors?

Some reductionists about color accept the following package of commitments, which I shall call the Qualia View:

Color experiences have certain intrinsic, introspectively available qualities called "color qualia." The color qualia exemplified by a color experience determine or constitute the phenomenal character of the experience—i.e. what it's like to undergo the experience. Color qualia, so understood, are distinct from colors. Colors are the properties *represented* by color experiences—the properties *attributed to* or *predicated of*—ordinary material objects by color experience. Colors are, in other

words, the properties we visually experience as belonging to such things as the surfaces of tomatoes. On the other hand, color qualia are intrinsic properties *of* color experience, which are *not* attributed to ordinary material objects by color experience (Block 1990, Peacocke 1983). Although we *experience*, or are *aware of*, color qualia, we do not experience color qualia *as* belonging to such things as the surfaces of tomatoes.

The Qualia View opens up a natural line of response to Intrinsicity and Categoricity objections above. Specifically, it might be suggested that when we judge that colors are intrinsic and categorical, we are confusing the color qualia *exemplified by* our color experiences with the colors *represented by* our color experiences. On this suggestion, when I look at a ripe tomato and judge that *that color*—the color that I experience as belonging to the surface of the tomato—is intrinsic and categorical, I am registering certain genuine features of the reddish quale exemplified by my experience, but I am mistakenly projecting these onto the external (represented) color.

This sort of response is commonly used against objections to reductionist theories of color, relationist and physicalist alike. For example, a common objection to color physicalism is that the apparent resemblance structure of color space is not shared by any collection of physical properties which might reasonably be identified with the colors (§3.2). In response, many color physicalists say that when we make judgments ostensibly about the resemblance relations among colors, e.g. that red is (intrinsically) more similar to orange than it is to green, we are really just registering facts about the resemblance relations among color qualia or color experiences, e.g. the fact that the reddish quale is

(intrinsically) more similar to the orangish quale than it is to the greenish quale (cf. Lewis 1997, p. 330; McLaughlin 2003, p. 115; Cohen 2003, p. 88).

The problem with this line of response is that the Qualia View is false. There are no introspectively accessible qualities associated with color experience distinct from the colors (levels of illumination, etc.) that appear to belong to the surfaces of objects in our environment. I urge you to place a ripe tomato (or some acceptable surrogate) in the middle of your visual field. All sides agree that there is a certain consciously accessible quality somehow associated with the middle of your visual field which is in some way constitutive of what it's like to undergo your current visual experience. Call this quality *Q*. I am quite sure that *Q* and similar qualities are what proponents of the Qualia View have in mind when they talk about "color qualia." But I am also quite sure that they have mischaracterized such qualities. For our purposes, the most important observation to make about *Q* is this: *you experience Q as belonging to the surface of the tomato*. In other words, *Q* appears (looks, seems) to you to belong to the surface of the tomato. (Doesn't it?)⁵ And that means that *Q* does not satisfy the characterization of color qualia given by the Qualia View. Color qualia, unlike colors, are supposed to be intrinsic qualities of experience, and *not* qualities we experience as belonging to external objects. So *Q* is not a color quale, as defined above. But if *Q* is not a color quale, then nothing is a color quale. So there are no color qualia.

⁵ This observation is closely related to the oft-made observation that color experience is "diaphanous" or "transparent" (Harman 1990, Tye 1992).

The Qualia View might be revised to accommodate the fact that we experience Q and other such qualities as belonging to external objects. The Revised Qualia View agrees with the original Qualia View in holding that color experiences, in addition to representing color properties, have certain introspectively accessible intrinsic qualities (distinct from the colors they represent), and that these are properly called “color qualia.” Where the Revised Qualia View diverges from the original Qualia View is that the former allows that we *do* (at least sometimes) experience these qualia as belonging to external material objects like the surfaces of tomatoes. In such cases, however, we are victims of a projective illusion, seeing what is in fact a quality of our experience as belonging to the object of our experience. This Revised Qualia View also allows us to explain away judgments like Categoricity and Intrinsicity in the manner described above.

But given that Q and similar qualities are experienced as belonging to ordinary material objects like the surfaces of tomatoes, surely we can recognize a legitimate sense of the word “color” on which Q and similar properties qualify as colors. Importantly, when I claim that non-reductive realism is true about colors, I mean these theses to apply to Q and other such qualities—the qualities which are directly presented to us in color experience, and which we experience as belonging to the objects of color experience. Concerning colors in my sense, the Revised Qualia View is just a form of irrealism, not reductive realism; it is a view on which our color experience of Q and other such qualities as features of external objects is a grand illusion. I shall defend realism against such views in §4. But crucially, the Revised Qualia View cannot be used to respond to my objections to relationism above, because on the question at issue—whether colors are

response-dependent properties of ordinary external objects—the Revised Qualia View agrees with me: they are not.

2.3 The Argument from Modal Relations

It is common for relationists to adopt an *inegalitarian* view of shape and color, according to which shape is “objective” or response-independent and color is “subjective” or response-dependent. That is, *inegalitarianism* combines a relationist view of color with a non-relationist view of shape. The next argument against relationism, which I call the argument from modal relations, has two stages. In the first stage, I argue that the modal relations between color and shape make the *inegalitarian* view untenable. We must embrace some form of *egalitarianism*, according to which shape and color are either both response-independent (objective egalitarianism) or both response-dependent (subjective egalitarianism). The argument against *inegalitarianism* begins with the following principle about the modal relations between shape and color: it is impossible for something to have a color without having a shape, i.e. without being spatially extended in some way. I then proceed to argue that, given reasonable assumptions, *inegalitarianism* contradicts this principle. Given the immense plausibility of the latter, I conclude that we should reject *inegalitarianism* in favor of some form of *egalitarianism*. In the second stage of the argument, I argue that, given *egalitarianism*, we ought to accept objective *egalitarianism* in particular. The upshot is that (not only shapes, but also) colors are response-independent properties, so relationism is false.

2.3.1 Preliminaries

For simplicity, in what follows I'll assume the color relationist accepts a simple dispositionalist form of relationism according to which redness (for example) is identified with the disposition to produce experiences with a reddish phenomenal character ("reddish sensations") in normal perceivers under normal conditions. Likewise, I'll assume the relationist about shape accepts a similar dispositionalist form of relationism according to which roundness (for example) is identified with the disposition to produce experiences with a roundish phenomenal character ("roundish sensations") in normal perceivers under normal conditions. However, the arguments given in this section are largely independent of this assumption and are adaptable to a wide range of possible relationist views (including, *inter alia*, non-dispositionalist relationist views like those described in §2.2).

I will use expressions like "subjective redness" (or "subjective roundness") to denote the response-dependent property that the relationist about color (shape) identifies with redness (roundness). *Subjective colors* include subjective redness, subjective whiteness, and their ilk. *Subjective shapes* include subjective roundness, subjective squareness, and their ilk.

There is a minor technical complication that is worth mentioning here, if only to justify ignoring it in what follows. It is common for relationists to relativize color properties to kinds of perceivers or to particular perceivers and viewing conditions. On such views, there is no such thing as *the* property of being red. Rather, there are many kind-relative, or perceiver-and-viewing-condition-relative, properties of being red, whose

analyses take the form: “x is red for K =: x is disposed to produce experiences reddish sensations in normal members of K in circumstances which are normal for members of K,” or “x is red for S in C =: x is disposed to produce reddish sensations in subject S in circumstance C.” The commitment to the relativity of color in standard forms of relationism means that there are more color properties than we might initially have suspected. But for our purposes, what is important to recognize is that, given relationism, this proliferation of relativized color properties just gives us *more color properties*. So if colors are relativized to kinds, or to perceivers and circumstances, then whatever is necessarily true of color properties as such must be true of these relativized color properties. For example, suppose that (as I will claim in §2.3.2) it is necessarily true of color properties as such that they are only instantiated by things which have a shape, i.e. things which are spatially extended in some way. Then by Leibniz’s Law, if color relationism is true, this must be necessarily true of the kind-relative response-dependent properties with which the relationist identifies color properties. Because my arguments depend only on claims about what is necessarily true about color properties as such, I shall hereafter (harmlessly) ignore this technical complication about color relativity in typical relationist views.

2.3.2 Against Inegalitarianism

“Come then, let us try to tell you what shape is.” says Socrates to Meno. “Let us say that shape is that which alone of existing things always follows color.” (75b). Whatever the shortcomings of this as an *analysis* of shape, Socrates’s suggestion certainly seems to be

correct in one respect: shape, as a matter of necessity, always follows color, in that whatever has a color must also have a shape. Equivalently, color “modally depends” on shape, in the following sense:

Modal Dependence (MD): It is (metaphysically) impossible for something to have a color without having a shape.⁶

MD has been widely accepted, and indeed regarded as a truism, throughout the history of philosophy. Locke (1689/1996) writes, “It is true, solidity cannot exist without extension, neither can scarlet color exist without extension” (Bk II, ch. 13). In Berkeley’s (1713/1979) *Three Dialogues*, Hylas tells us (and Philonous evidently agrees) that “color can’t exist without extension,” treating this claim as an obvious truth that any philosophical theory must respect (First Dialogue; cf. Berkeley 1710/1982). Later philosophers, taking MD as self-evident, have treated it as a paradigm of the synthetic *a priori* (Husserl 1900/2001, Bonjour 1998, Hospers 1997).⁷ And in recent philosophy, MD (or something equivalent) is regularly used as a clear-cut, illustrative example of a

⁶ Two points: first, in saying that color “modally depends” on shape, I mean nothing more than is contained in the above formulation. I do not mean to suggest that shape properties are in any interesting sense “ontologically prior” to color properties, or that the former “metaphysically ground” the latter. Second, one might object to MD by claiming that it’s possible for a point-sized object to have a color, but that such an object is not correctly described as having a shape. However, my arguments will not be affected if we interpret “having a shape” broadly, so that point-sized objects count as having a shape.

⁷ Some of those cited here mention instead the non-modal proposition that all colored things are extended rather than the modal proposition MD as a example of the synthetic *a priori*, but they each take for granted that the former is a necessary truth, and so take for granted the truth of MD.

necessary truth in service of more general philosophical discussions about metaphysical modality.⁸

MD is the key background assumption in my argument against inegalitarianism. I regard MD as a basic intuition about the nature of color and shape. It seems to be far more plausible than any specific theory about the nature of color and shape. Theoretical claims of the latter sort are to be tested in part by their ability to accommodate truisms like MD. At the very least, if a philosophical theory conflicts with MD, this should be regarded as a serious, if not decisive, cost for the theory.

From here, the argument against inegalitarianism appeals to one further premise:

P. If inegalitarianism is true, then possibly, something has a subjective color without having a shape.

MD together with P entail the falsity of inegalitarianism. For assume MD and P are true, and now suppose for reductio that inegalitarianism is true. Inegalitarianism and P jointly entail: (1) possibly, something has a subjective color without having a shape. But for the inegalitarian, subjective colors *are* colors. So inegalitarianism and (1) jointly entail: (2) possibly, something has a color without having a shape. But (2) contradicts MD. So (given MD and P) inegalitarianism is false.

All that remains, then, is to establish the truth of P. To do so, it will suffice to describe a scenario *w* such that (i) in *w*, something has a subjective color without having

⁸ A typical example from the opening sentences of Kroedel (2012), a paper addressing general issues about the epistemology of metaphysical modality: “We seem to have plenty of modal knowledge, that is, knowledge that such-and-such is metaphysically necessary and that so-and-so is metaphysically possible. For instance, we know that necessarily everything that is colored is spatially extended, and we know that there might have been flying pigs” (p. 1). Cf. Sider (2003, p. 184).

a shape, and (ii) if inequality is true, then *w* is a genuine (metaphysical) possibility. As I shall argue, these conditions are satisfied by the following scenario, which I call *W*.

W is an entirely non-spatial world inhabited by a vast host of non-spatial objects —objects that are neither spatially extended (though they possess other, non-spatial intrinsic properties), nor spatially related to one another (though they stand in various non-spatial relations, including causal relations, to one another). Among the inhabitants of *W* are minds—that is, subjects of mental states. We might suppose that these minds are mereologically simple, as Descartes held actual minds to be. Or we might instead suppose that these minds are mereologically complex, as materialists hold actual minds to be, built up from simpler parts which causally interact with one another in complex ways that structurally parallel the causal interactions among the parts of an actual human brain. The minds of *W*, like ours, undergo a wide range of experiences, whose subjective characters co-vary in systematic ways with certain intrinsic features of the extra-mental objects which cause them, at least in normal circumstances. Some such extra-mental objects are disposed to produce reddish sensations in normal perceivers under normal conditions. Others are so disposed to produce greenish sensations, others blueish sensations, and so forth.

In other words, some objects in *W* possess *subjective-colors*. But because everything in *W* is entirely non-spatial, these objects are devoid of any shape. So *W* satisfies condition (i) above. What about condition (ii)? Is *W* a genuine metaphysical possibility? *W* seems to be a conceivable scenario. Conceiving *P*, in the relevant sense—a sense involving the *appearance of possibility* (cf. Yablo 1993, Chalmers 2002)—

plausibly provides *prima facie* justification to believe that P is possible.⁹ In other words, given the conceivability of W, we ought to believe that W is possible unless there are sufficiently strong countervailing considerations. It's also worth noting that similar scenarios familiar from discussions of external-world skepticism are widely regarded as possible. For example, the scenario Descartes imagines in *Meditations I-II*, in which his sensory impressions, including his color sensations, are produced in him by a non-spatial being (an immaterial demon) and in which “body, shape, extension, movement, and place are all chimeras” (p. 63), strikes most readers as possible. And indeed, some philosophers, such as David Chalmers (2006, 2012), hold that something like W actually obtains.¹⁰ Nonetheless, I do not positively claim that W is possible. For reasons to be discussed shortly, there may be grounds for denying its possibility. What I do claim, however, is that whatever grounds there may be for denying the possibility of W are not available to the inegalitarian. For, as we'll see, these grounds rely on assumptions about

⁹ This principle of modal epistemology is entailed by, but does not entail, the more general principle of Phenomenal Conservatism mentioned in §2.2. The latter principle, or something like it, is likely to be attractive to color relationists. Color relationists are generally color realists, and their rejection of the irrealist alternative is often justified on something like Phenomenal Conservative grounds. (See, e.g., Cohen 2009, p. 65.)

¹⁰ More precisely, Chalmers maintains that “Edenic” shapes—roughly speaking, the shape properties directly presented to us in experience, and others under the same determinable—are not instantiated in the actual world. If we restrict our attention to Edenic spatial properties, Chalmers would take my description of W to be true of the actual world. However, he allows that “ordinary” or “imperfect” shapes, picked out as those properties, whatever they are, that serve as the normal causal basis of our experience as of Edenic shape, are instantiated in the actual world.

the nature of shape and the nature of experience that cannot be comfortably combined with the commitments of egalitarianism.¹¹

There is one ground for denying the possibility of W which is available to an egalitarian or a subjectivist stripe. W is stipulated to be a world in which some things have subjective colors. But it's somewhat plausible that there is a modal relation between subjective color and subjective shape analogous to the modal relation between color and shape. In other words, it's somewhat plausible that the following modal-dependence principle is true:

Subjective Modal Dependence (SMD): It is impossible for something to have a subjective color without having a subjective shape.

This claim might be supported by arguing that nothing can produce the subjective response-type associated with any color without producing a subjective response-type associated with some shape. Perhaps it's a "law of appearance" that one cannot have (say) an experience as of something red without having an experience as of something spatially extended in some way. If SMD is correct, then given relationism about shape, W

¹¹ Some may find the idea of an entirely non-spatial world, or of non-spatial objects with causal powers, so bizarre that they are hesitant even to accept that there is a *prima facie* case for the possibility of W. For philosophers in this group, it's worth noting that there is a somewhat less bizarre alternative scenario that would serve our purposes nearly as well as W, which I'll call W* (This scenario was suggested to me by Adam Pautz). W* is just like W except that the entities that produce color experiences are point particles rather than non-spatial entities. If we ran the argument with W* instead of W, we would have to give a somewhat stronger interpretation to MD than that given in fn. 6 by taking the notion of "having a shape" in such a way as to exclude point-sized objects. The resulting interpretation of MD, though stronger than the official interpretation, is nonetheless very plausible. I rely on W here rather than W* because I'm more certain of the official interpretation of MD than the stronger interpretation. But if one finds the two interpretations of MD about equally plausible, and one has significantly more confidence in the possibility of W* than in the possibility of W, then one should keep this alternative version of the argument in mind in what follows.

is impossible. For *W* is stipulated to be a world in which things have subjective colors but lack shapes. If it's impossible for something to have a subjective color without having a subjective shape, and shapes just *are* subjective shapes, then it's impossible for something to have a subjective color without having a shape.

But of course, the inegalitarian does not identify shapes with subjective shapes. SMD may be true, in which case *W* is a world where everything with a subjective color also has a subjective shape. But for the inegalitarian, having a shape is not simply a matter of having a subjective shape. So for the inegalitarian, the fact (if it is a fact) that nothing can have a subjective color without having a *subjective shape* is no reason to think that nothing can have a subjective color without having a *shape*, and so provides no reason to think that *W* is impossible.

But perhaps SMD opens up another line of response for the inegalitarian.¹² In particular, the inegalitarian might claim that MD is false and that it only seems plausible because we confuse it with SMD. I find this debunking strategy unconvincing for three reasons. First, it seems that the intuition that MD is true persists even when we are given this alternative explanation of why MD seems plausible. In this respect, the proposed debunking strategy differs from other successful cases of explaining away modal appearances through suggestions of “proposition confusion,” such as Kripke’s explanation of the appearance that water could have been distinct from H₂O, which claims (on one interpretation of Kripke) that we’re confusing the proposition <possibly, water ≠ H₂O> with (something like) the proposition <possibly, the watery stuff around

¹² Thanks to an anonymous referee for this suggestion.

here \neq H_2O > (Kripke 1980). Second, we would not find analogous debunking strategies convincing for other intuitive modal claims, such as (i) it is impossible for something to be a bicycle without being spatially extended, or (ii) it is impossible for something to be both round and triangular. It's plausible that nothing could appear bicycle-ish without appearing spatially extended, and it's also plausible that nothing could appear both round and triangular. But these facts have no tendency to undermine the intuitive support for (i) and (ii). Why, then, should SMD undermine the intuitive support for MD? Third, it seems to me that we're not generally prone to confusions of this kind. For example, arguably nothing could appear red to me without appearing to be before me, but this fact (if it is a fact) does not fool me into thinking that nothing could *be* red without *being* before me. It therefore seems doubtful that we would be subject to this sort of confusion in the case at hand.

A second reason for denying the possibility of W comes from phenomenal externalism (Dretske 1995, Tye 1995, Lycan 2001a, Byrne and Tye 2006). According to phenomenal externalism, the phenomenal character of a sensory experience constitutively depends on the subject's extrinsic relations to features of her environment. According to one common version of phenomenal externalism, the phenomenal character of one's experience is determined by the properties represented by the experience, where one's experience represents a property just in case, very roughly, one's current internal state "tracks" or causally co-varies with instances of that property under normal conditions. Hence, according to this "tracking" externalist view, in order to have a square-ish sensation—i.e. an experience with the subjective character associated with our

experiences as of squares — one must token an internal state which causally co-varies (under normal conditions) with instances of squareness. But in *W*, squareness is never instantiated, nor is any other shape property. So the internal states of the subjects in *W* don't causally co-vary with instances of squareness or any other shape property. Given this version of phenomenal externalism, then, the subjects in *W* will not undergo shape phenomenology. If it is indeed a law of appearance that color experience entails shape experience, it follows that the subjects in *W* could not have color experiences. As *W* was stipulated to be a world in which subjects *do* have color experiences, it follows that *W* is not a possible world.

However, it is not clear that this line of reasoning is available to the inegalitarian. This is because the inegalitarian is committed to color relationism, and as many others have pointed out, phenomenal externalism cannot be happily married with color relationism (Tye 1995, pp. 144-5; Dretske 1995, pp. 88-93). How, for example, could the color relationist give a phenomenal-externalist account of our color experience? The color relationist holds that for an object to have a given color property *C* is for it to be disposed to produce subjective responses of type R_C in normal perceivers under normal conditions. If she proceeds to endorse a phenomenal externalist position, according to which what it is for one to token the subjective response-type R_C is for one to be in an internal state which tracks instances of *C*, then she courts vicious circularity. The resulting view, on which *C* = the disposition to produce instantiations of the property of being in some state or other that tracks instances of *C*, has the absurd consequence that *C* is one of several constituents out of which *C* itself is constructed — that *C* is a constituent

of itself. As Johnston (2001) remarks about a related view, the envisaged position “identif[ies] a property [...] with the relation to a holding of another relation that has that very property as a relatum. But a property cannot be contained within itself in this way” (p. 195, cf. Pautz 2010b).¹³ Furthermore, it seems implausible to suppose that our internal states track properties of this sort. When I have an experience of red₁₇, perhaps I am in an internal state which, under normal conditions, is caused by a certain physical property, perhaps a surface reflectance type. But it is not plausible that my internal states are caused by objects’ dispositions to produce experiences of a certain type in normal perceivers under normal conditions (cf. Prior et al. 1982, Jackson and Pargetter 1987, McLaughlin 2003).

Granted, the phenomenal-externalist argument above for the impossibility of W only requires that the externalist “tracking” account of sensory experience apply to our *shape* experience; it does not require that such an account hold true for color experience. Perhaps, then, the egalitarian can opt to retain just enough phenomenal externalism as is needed to deny the possibility of W, but not so much as to run her into the problems discussed above, perhaps by combining an externalist “tracking” account of shape

¹³ It’s worth noting that the circularity objection under consideration is importantly different from the more familiar concerns about circularity that arise for color relationism, or at least for versions of color relationism whose analyses of color take something like the form “*x* is red =: *x* is disposed to look red,” where “red” occurs on both sides of the analysis. Many relationists have convincingly argued that the apparent circularity in such analyses needn’t be vicious. For example, Cohen (2009) rightly points out that the apparent circularity in question won’t be vicious if the ultimate analysis or real definition of what it is to look red to a subject does not consist in the subject’s bearing a relation to redness or otherwise mention redness. And there are, as Cohen points out, many respectable “non-relationist” theories of color experience to which the relationist can appeal to avoid objectionable circularity (p. 170). The circularity objection I raise above applies only to the combination of color relationism with a view of color experience according to which the subjective response-type involved in the relationist’s analysis of a color *C* is a relational property consisting in a relation to *C* itself.

experience with an internalist account of color experience. But there are three serious problems with this response. First, it is extremely *ad hoc*. Second, and more significantly, it apparently precludes any *unified* answer to the question of what determines the phenomenal character of experience, one which would apply to experiences of different sensory modalities or even experiences of different sensible qualities within the visual modality. To abandon hope of such a unified account is a theoretical cost that most will be unwilling to pay.

Third, apart from these broadly theoretical considerations, in the present context there is a fatal problem with appealing to any view, such as the view under consideration, that combines an externalist account of shape phenomenology with an internalist account of color phenomenology. Imagine there is an isolated brain whose intrinsic physical state perfectly matches that of my own brain, but which, unlike my own brain, has never had any interesting causal commerce with shapes in its environment. Let us suppose further that it is not in a state that *would* be produced by an object with this or that shape if only such an object were to be placed before it, for, unlike my brain, it is not connected to the receptor cells needed to facilitate such world-to-brain exchanges. Nor, we may add, did it come about through any process, such as natural selection or intelligent design, that could have conferred on the brain's present state the biological *function* of detecting this or that shape. Rather, we may suppose it came together in the manner of a Boltzmann brain, a result of random thermal fluctuations of particles in the void. Given any reasonable form of externalism about shape experience, this brain is not enjoying any shape phenomenology. But given internalism about color phenomenology, it must be enjoying

rich color phenomenology. After all, it is an intrinsic duplicate of my brain, and I am enjoying rich color phenomenology.

So it is a consequence of the hybrid view that it's possible for a subject to have color phenomenology without having shape phenomenology. Either this consequence is true or it's false. If it is false (as seems plausible), it follows that the hybrid view is also false and so cannot be used to argue for the impossibility of W. If it is true, then this fact undermines the argument from phenomenal externalism to the impossibility of W. Recall that that argument derived the impossibility of W from two premises: (i) it is impossible to have shape phenomenology in an entirely non-spatial world; (ii) it is impossible to have color phenomenology without having shape phenomenology. Although the hybrid view, if true, could be used to support the first premise, it entails the falsity of the second premise and thereby undermines the argument from phenomenal externalism to the impossibility of W.

The third way in which the egalitarian might attempt to deny the possibility of W is as follows: W is supposed to be an entirely non-spatial world, in which an entirely non-spatial subject undergoes color sensations. But on certain physicalist views, what it is to have a given color sensation is to have a certain neurophysiological property. Now, it's plausible that an entirely non-spatial subject could not have any neurophysiological properties. For anything with neurophysiological properties must presumably have neurons among its parts. And an entirely non-spatial subject could no more have neurons as parts than it could have hands and feet as parts. So, if some such physicalist view about color sensations is correct, then W is impossible.

Unfortunately, these considerations do not help the inegalitarian. For it was inessential to the argument that the *subjects* of the relevant experiences be non-spatial. What matters to the argument is that the *objects* of the relevant experiences, those things which produce color experiences of various kinds in the subjects, be non-spatial. We may therefore amend the description of W to avoid the response above by stipulating that the subjects have whatever spatiality is needed to possess the relevant neurophysiological properties while leaving the rest of the description intact. We might, for example, suppose that the subjects are intrinsically just like human brains and that the appropriate neurophysiological states are produced in them by immaterial objects under certain conditions, in accordance with fundamental laws governing the interaction of spatial and non-spatial objects.

I have considered three grounds for denying the possibility of W—that is, three propositions (subjective egalitarianism, phenomenal externalism, type-identity physicalism about color phenomenology) that a philosopher might reasonably endorse and which apparently have the consequence that W is impossible. But as we've seen, the first two are unavailable to the inegalitarian, and the third is unhelpful to the inegalitarian.

I am aware of no other grounds for denying the possibility of W.¹⁴ On the assumption that inegalitarianism is true, we should therefore conclude that there is a possible world—either W or some variation on W—in which things have subjective color but lack shape. In other words, P, the key premise in the argument above, is true. But the inegalitarian holds that colors *are* subjective colors. So given inegalitarianism, we have not merely a possible world in which things have subjective color without shape, but a possible world in which things have *color* without shape. But it is not possible for something to have color without shape. So inegalitarianism is false.

2.3.3 From Egalitarianism to Non-relationism

The inegalitarian accepts the following two theses:

T1. Colors are response-dependent properties.

T2. Shapes are response-independent properties.

¹⁴An anonymous referee suggests a way the inegalitarian might avoid (not the result that W is possible, but) the result that objects in W are colored. I've described the color relationist as identifying colors with dispositions to produce color sensations of a certain kind (in normal perceivers under normal conditions). But some color relationists identify colors instead with dispositions of objects to *look* some way (to normal perceivers under normal conditions), where an object's looking some way to a perceiver will then be understood (at least partially) in terms of the object's producing sensations of the right kind. The crucial difference, the referee suggests, is that appealing to the looks relation may give the inegalitarian a way to deny that the objects in W are colored by maintaining that being spatially extended is a metaphysically necessary condition on something's looking a certain way. (The referee suggests that this might be because *x*'s looking some way to *S* requires *S* to visually attend to *x*, where perhaps it is a metaphysically necessary condition on visually attending to *x* that *x* be spatially extended.) Without attempting an analysis of the looks relation (or the relation of visual attention) here, I think there is good reason to deny the suggestion that being spatially extended is a necessary condition on looking some way (or being the object of visual attention). Consider Chalmers's (2006, 2012) view that shapes (more precisely, "Edenic" shapes—see fn. 10) are not really instantiated in the world, but only appear to be instantiated. I do not think this view is true, but if it were true, it seems we would still want to say, as Chalmers does, that the objects that produce roundish experiences in us *look* round, and the objects that produce reddish experiences in us *look* red (and that we sometimes visually attend to such objects). One hardly wants to say that *nothing* looks round or red in such a scenario, but what else would look round or red if not the objects that produce the relevant visual experiences?

If the argument above is successful, we must reject inequality in favor of some form of egalitarianism. We can embrace *objective* egalitarianism by accepting T2 and rejecting T1, or we can embrace *subjective* egalitarianism by accepting T1 and rejecting T2. It seems clear that the first option is by far the more reasonable option. There is a lot of controversy over T1. By contrast, T2 is about as close to uncontroversial as anything gets in philosophy. Sound methodology would therefore seem to recommend taking T2 as our fixed Archimedean point. From here we may dislodge T1, using egalitarianism as our lever.

If one needs any coaxing on this point, consider what follows if we accept a response-dependent view of shape in addition to a response-dependent view of color. A consequence of a response-dependent view of color is that our knowledge of the colors of things is not knowledge of how things in the external world are in themselves, but only knowledge of how things are in relation to us and our sensory experiences. Now, if we proceed to endorse a response-dependent view of shape, we must likewise concede that our knowledge of the shapes (and spatial features generally) of things is also not knowledge of how things are in themselves, but only knowledge of how things are in relation to us and our sensory experiences. But this seems to entail the radically skeptical conclusion that we are more-or-less entirely ignorant of how things in the extra-mental world are in themselves. Many — myself included — will find this result unacceptable.

A response-dependent view of shape — a view on which the instantiation of shape properties constitutively depends on the subjective responses of perceivers — would also seem to have the consequence that whatever properties constitutively depend on shape

also constitutively depend, at least indirectly, on the subjective responses of perceivers. But if we think of the range of properties that constitutively depend on an object's shape, this consequence looks utterly implausible. For example, whether an object has such properties as *being a leaf*, *being a star*, *being a mountain*, or *being a pine cone* depends in part on the object's spatial features, including its shape. But it's outrageous to suppose that something's being a leaf (or a star, or a mountain, or a pine cone) constitutively depends, directly or indirectly, on the subjective states of perceivers.

To summarize: I've argued that egalitarianism is true. That is, either shape and color are both response-independent, or shape and color are both response-dependent. Since shape is response-independent, I conclude that color is response-independent as well. So color relationism is false.

2.3.4 Privileged Status

Inegalitarianism, as I've characterized it, may be regarded as a specific version of a more general thesis, which we might call the *Privileged Status Thesis* (PST):

PST: Shapes have some philosophically interesting sort of privileged status *vis-a-vis* colors.

PST is very plausible, and the attraction some philosophers feel toward inegalitarianism or color relationism may derive from an attraction to PST. In light of the foregoing, it is important to note that PST is consistent with the rejection of inegalitarianism. I shall conclude by describing four ways (not all mutually exclusive) of upholding PST—four ways of according shapes a privileged metaphysical status relative to colors—which are untouched by the arguments above. First, one might uphold PST by observing that shapes

are *common sensibles*, understood as qualities that can be sensed through more than one sense modality (e.g. vision and touch), whereas colors are not. Second, one might uphold PST by maintaining that shapes, or spatial properties more generally, have a “wide cosmological role,” figuring in causal laws with sweeping generality, whereas colors have very limited causal relevance (perhaps none beyond the explanation of color perception or facts closely related to color perception, such as sorting behavior) (Campbell 1993, p. 264). Third, one might uphold PST by maintaining that shapes are more *natural* than colors in the sense of Lewis (1983), or that colors, unlike shapes, are highly disjunctive or gerrymandered properties (cf. Byrne and Hilbert 2003a, Tye 2000, Smart 1975, Jackson and Pargetter 1987). Finally, one might accept a more radical form of PST according to which *realism* is true about shape but not about color.¹⁵

2.4 Perceptual Variation

It commonly happens that an object looks different in respect of color to different perceivers, or to the same perceiver under different viewing conditions, even when viewing conditions are normal and the perceivers in question qualify as normal according to standard tests for normal color vision. For example, in the same (normal) viewing conditions, a given color chip might look unique green (that is, green without any admixture of yellow or blue) to one normal perceiver, yellowish green to another, and blueish green to another. Cases like this jointly constitute what I’ll call the phenomenon of *perceptual variation*. The phenomenon of perceptual variation (especially *normal*

¹⁵ The last option will be criticized in chapter 4, and something in the neighborhood of the third option will be criticized in §3.5, but nothing I’ve said so far rules them out.

variation—variation among normal perceivers in normal viewing conditions) makes serious trouble for any relationist view of color, or so I shall argue. This claim might be surprising. Relationists typically take the phenomenon of perceptual variation to *support* relationism. Indeed, the facts of perceptual variation are often treated as the main support for relationism. For example, the phenomenon of perceptual variation is Locke’s (1689/1996) primary motivation for accepting relationism about colors and other so-called secondary qualities. And Cohen (2009) calls the argument from perceptual variation “the core argument” (elsewhere “the most important argument”) for color relationism (p. 19). In §2.4.1, I’ll examine the argument from perceptual variation for relationism. At this stage I’ll focus on Cohen’s version of the argument in his (2009), which contains the most developed and sophisticated defense of the argument from perceptual variation to date. I shall argue that perceptual variation provides no support for relationism. More specifically, I shall argue that the phenomenon of perceptual variation turns out to play no essential role in the so-called “argument from perceptual variation.” I then proceed to argue that perceptual variation furnishes the materials for a powerful argument *against* relationism (§2.4.2).

2.4.1 The Argument from Perceptual Variation for Relationism

Consider a case like the one described above, in which a given color chip looks unique green to some normal perceivers, yellowish green to others, and blueish green to others. Call the visual experiences of these perceivers “variants.” There are three possibilities:

ALL-RIGHT: All of the variants are veridical.

ONLY-SOME-RIGHT: Some of the variants are veridical and some are not.

NONE-RIGHT: None of the variants is veridical.¹⁶

The first stage in Cohen’s argument from perceptual variation to relationism is to argue that we should accept ALL-RIGHT. In effect, his argument at this stage is an argument by elimination: we should accept ALL-RIGHT because there are problems with the other two alternatives. The problem with NONE-RIGHT is that it leads almost inexorably to the irrealist view that nothing in the external world is colored. This consequence “flies in the face of naive belief. Put bluntly, it is a notable part of the manifest image that colors are exemplified by actual, ordinary objects—fruits, lights, tables, chairs, etc., and that we learn something about the world when we visually perceive the colors of those objects. There is, then, Moorean pressure to reject irrealism about color that must be counted as a significant strike against the view” (p. 65). His objection to ONLY-SOME-RIGHT is a bit more involved. I shall postpone discussion of it until §2.4.2. For now, let it suffice to say that he finds ONLY-SOME-RIGHT objectionable, and so opts for the only remaining alternative: ALL-RIGHT.

It’s worth noting that ALL-RIGHT is fairly counterintuitive. If something looks unique green to John and yellowish green to Jane, we are tempted to ask, “who (if either) is getting the color of the chip right?,” where the underlying presupposition is that they can’t *both* be getting the color of the chip right. There is, in other words, an intuition of

¹⁶ I deviate from Cohen’s terminology here. Cohen uses “ecumenism,” “inegalitarianism,” and “eliminativism.” I avoid this terminology for two reasons. First, since I used the term “inegalitarianism” in §2.3 to designate an unrelated thesis, it might produce confusion to use the term in Cohen’s sense here. Second, “eliminativism” usually designates the thesis that nothing is colored. But the claim that all the variants are non-veridical, though it strongly suggests that nothing is colored, does not quite entail that nothing is colored. (It might be that the chip is red. Or it might be that the chip has a coarse-grained color like green (or greenish) but no fine-grained color. More on the latter possibility in 2.4.2.)

conflict, an intuition that they can't *both* be veridically representing the chip. Following Pautz (2010b), we'll call this the "Conflict Intuition." We may ultimately want to reject this intuitive judgment on theoretical grounds, but it is a very natural assumption to make. This is why the "who's getting it right?" question comes naturally here, where it seems clearly out of place in other cases, e.g. the case in which the chip feels hard to John and feels smooth to Jane. In the latter case, the "who's getting it right?" question seems out of place because there's no temptation to think they can't both be getting it right.

I shall return to the Conflict Intuition later. For now, let us follow Cohen and assume the truth of ALL-RIGHT. It's worth emphasizing that the whole point of Cohen's extended discussion of actual, empirically verifiable cases of perceptual variation in his "argument from perceptual variation" for relationism is to get us to ALL-RIGHT. But on the face of it, this looks like so much work for nothing. That's because, on the face of it, ALL-RIGHT doesn't get us any closer to color relationism than we were before learning about the phenomenon of perceptual variation. ALL-RIGHT obviously doesn't *entail* relationism. Moreover, it seems that ALL-RIGHT can be combined with a non-relationist view of color just as easily as it can be combined with a relationist view of color. This is perhaps easiest to see if we consider a physicalist version of non-relationism, such as the reflectance physicalist view that color shades are identical with reflectance types, or disjunctions of reflectances (Byrne and Hilbert 2003a, Tye 2000, Lewis 1997, Matthen 1988). Reflectance physicalism as such is entirely neutral on the question of whether certain pairs of determinate color shades, such as unique green and yellowish green, are incompatible. If the reflectance physicalist is convinced that ALL-RIGHT is the correct

response to cases of normal variation, perhaps on the basis of Cohen's argument by elimination, she should maintain that unique green and yellowish green (for example) are compatible. A natural way of implementing this suggestion within a reflectance physicalist framework is to hold that unique green and yellowish green are disjunctions of reflectances with some disjuncts in common (Byrne and Hilbert 1997). To give a simplified model: suppose the chip has reflectance R_2 , and the chip looks unique green to John and slightly yellowish green to Jane. Then both of their experiences will be veridical if, say, unique green = $R_1 \vee R_2$ and slightly yellowish green = $R_2 \vee R_3$.)

The same goes for non-reductive versions of non-relationism. The non-reductionist who is convinced that ALL-RIGHT is the correct response to cases of normal variation will simply maintain that unique green and yellowish green (for example) are irreducible qualities which happen to be compatible. Both of these pluralist forms of non-relationism are versions of what Cohen calls "selectionism," so-called because they portray different perceivers' visual systems as "selecting" one of several different (non-relational) colors the object actually has. The general idea is characterized nicely by Kalderon (2007), a defender of selectionism:

The relation between object, perceiver, and circumstance of perception does not determine the color of the object (in the way that it would if colors were relational) so much as it determines the perceptual availability of the color. Thus the relation between the chip, [John], and the circumstances of perception does not determine that the chip is unique green; rather it determines the perceptual availability of unique green for [John]. Moreover, the relation between the chip, [Jane], and the

circumstances of perception does not determine that the chip is yellowish green; rather, it determines the perceptual availability of yellowish green for [Jane]. On the [selectionist] pluralistic hypothesis, the chip is multiply colored—it is unique green all over and yellowish green all over at the same time; it is just that the perceptual availability of these sensible qualities is determined by different relations between the chip, perceivers, and circumstances of perception. According to the [selectionist], then, the [relationist] conflates the conditions of perception of a color for the perceived color. (p. 577)

In my view, the most serious problem with the selectionist combination of non-relationism with ALL-RIGHT is that it runs afoul of the Conflict Intuition. But of course, this is not a problem specifically for the combination of non-relationism with ALL-RIGHT, but a problem with ALL-RIGHT. It therefore provides no reason to prefer the combination of relationism with ALL-RIGHT to the combination of non-relationism with ALL-RIGHT, as Cohen recognizes (p. 79). Indeed, if we must accept ALL-RIGHT, then a selectionist version of non-relationism is *prima facie* preferable to relationism, for the former isn't threatened by the problems discussed in §§2.2-2.3. The selectionist can maintain that colors are intrinsic, categorical properties, just as they seem to be. The selectionist can also uphold the (objective) egalitarian thesis that color and shape are both response-independent properties.

Cohen is aware that ALL-RIGHT is consistent with non-relationist “selectionist” views. To complete the argument from perceptual variation for relationism, he takes up the task of arguing that, given ALL-RIGHT, we ought to accept relationism over the

selectionist alternative, and he does so by giving an argument against selectionism. The argument appeals to a hypothetical scenario inspired by Pautz (2006a). The scenario involves two subjects, Maxwell and Twin Maxwell, who evolved in different environments and who both qualify as normal perceivers for their species. Maxwell's and Twin Maxwell's visual systems are identical at the receptor level (e.g. they have the same kinds, proportions, number, and organization of photoreceptor cells in the retina), and their receptor responses to color stimuli are therefore identical. But they differ significantly with respect to their post-receptor processing of color stimuli. We might suppose that a given (non-relational) property P, though it produces the same receptor response in both Maxwell and Twin Maxwell, puts Maxwell's visual system into post-receptor state S_M and puts Twin Maxwell's visual system into a distinct post-receptor state S_{TM} . Cohen imagines that S_M is a state involving the activation of two opponent process channels in Maxwell's visual system (the sort of processing thought to underly our experience of binary hues like purple), while S_{TM} involves the activation of just one opponent process channel in Twin Maxwell's visual system (the sort of processing thought to underly our experience of unitary hues like (unique) green). These post-receptor differences will in turn result in behavioral differences, e.g. differences in sorting behavior and verbal reports about whether the color of the stimulus is unitary (uncomposed) or binary (composed). Cohen writes,

With this setup in hand, it is now easy to state the problem posed for the selectionist by the case of Maxwell and Twin Maxwell. On the one hand, the most plausible characterization is that Maxwell's visual state S_M and Twin Maxwell's visual state

S_{TM} differ in representational content. However, on the other hand, it appears that, given any of a wide class of popular psychosemantic theories, the selectionist will be forced to say that S_M and S_{TM} do not differ in the non-relational properties they select/represent.

We can reconstruct Cohen's argument against selectionism as follows:

C1. S_M and S_{TM} differ with respect to the colors they represent.

C2. S_M and S_{TM} do not differ with respect to the non-relational properties they represent.

C3. Therefore, the colors represented by S_M and S_{TM} (and so presumably colors generally) are not non-relational properties — that is, they are relational properties.

Premise C1 seems very plausible. It is very plausible that S_M and S_{TM} will differ phenomenally. And it's also plausible that there can be no phenomenal difference between two color experiences without a difference in the properties represented by those experiences (Tye 1995, Dretske 1995, Byrne and Hilbert 1997, Chalmers 2006, Pautz 2009),¹⁷ and it's plausible that in the case at hand the difference in properties represented would have to be a difference specifically in the *colors* represented.¹⁸

In my view, we should reject C2. C2 is supposed to be supported by the fact that it follows from "any of a wide class of popular psychosemantic theories." Here he takes as a representative example Tye's (1995, 2000) view, according to which a state S represents a property F iff S causally covaries with F in optimal conditions — roughly, the sorts of

¹⁷ Though see, e.g., Block (1990) for a contrary position.

¹⁸ Though see Shoemaker (1994) for a contrary position.

conditions in which the the subject's visual system was designed to operate. We might suppose that S_M and S_{TM} both causally covary with, or "track," the same external, non-relational property P in optimal conditions. In that case, Tye's psychosemantics seems to have the consequence that C2 is true: S_M and S_{TM} represent the same non-relational property, namely P . Moreover, Cohen claims that this consequence is not unique to Tye's psychosemantics. According to Cohen, a wide range of other psychosemantic theories have the same result, including those of Dretske (1995), Fodor (1990), Armstrong (1999), Davies (1993), Papineau (1993), and Millikan (1989).

There are a number of problems with his appeal to popular psychosemantics here. First, psychosemantics is one of those areas in philosophy where all the available theories are deeply unsatisfactory, at least if we try to take them as something more than programmatic and radically oversimplified sketches of what the correct theory might ultimately look like. So even if many extant psychosemantic theories entail C2, to conclude that C2 is true solely on this basis would evince an unwarranted level of confidence in current psychosemantic theories. A second problem is that, due to the highly schematic character of extant psychosemantic theories, it is difficult to judge whether all the theories mentioned really entail C2 (cf. Byrne and Tye 2006, p. 252). Third, it seems in any case that *if* the psychosemantic theories in question have the consequence that S_{TM} do not differ with respect to the *non-relational* properties they represent, this is only because they have the logically stronger consequence that they do not differ in *any* of the properties they represent. (As far as I can see, these theories would not predict that S_M and S_{TM} differ in the *relational* properties they represent either.) It is

therefore strange that Cohen is appealing to these theories here, because if these theories succeed in establishing C2, it is only because they succeed in establishing something that entails the falsity of C1.

It's also worth noting that some popular psychosemantic theories that Cohen does not mention, specifically those within the "interpretationist" tradition (Davidson 1973, Lewis 1974, Dennett 1987), would likely deliver the verdict that C2 is false. On an interpretationist psychosemantics, contents are assigned to states partly on the basis of what would rationalize the subject's behavior. Since S_M and S_{TM} produce different behavior, it's likely that, for any non-relational property F such that assignment of F as content S_M would rationalize Maxwell's behavior, assignment of F to S_{TM} would *not* rationalize Twin Maxwell's behavior, and *vice versa*. So on an interpretationist approach to psychosemantics, it's plausible that S_M and S_{TM} would differ in the non-relational properties they represent.

In any case, given that even our best psychosemantic theories have not matured past infancy, the way to proceed in the case at hand is *not* to take a top-down approach, starting with a worked-out psychosemantic theory and then deriving a verdict on the particular case. Rather, we ought to make our best intuitive judgment about the particular case given the details of the case. Taking the latter approach, it seems easy enough to sketch a plausible account of the case according to which S_M and S_{TM} represent distinct non-relational properties. (The following is partly inspired by a suggestion made, but not endorsed, by Pautz (2006a).) S_M is supposed to involve the activation of two opponent process channels, whereas S_{TM} is supposed to involve the activation of just one opponent

process channel. As a result, Maxwell judges the represented color of the stimulus to be binary (composed) and Twin Maxwell judges the represented color of the stimulus to be unitary (uncomposed). Let's call the pair of opponent process channels in Maxwell's visual system the "A-B channel" and the "C-D channel." A generalization of the opponent-process theory of human color vision (Hurvich and Jameson 1957) suggests the following picture: Positive ($>$ base-line) activation of the A-B channel represents the stimulus as having some determinable hue magnitude A, with higher activation levels representing larger determinate values of A, and negative ($<$ base-line) activation levels represent the stimulus as having some determinable hue magnitude B, with lower (higher absolute value) activation levels representing larger determinate values of B. The determinable hue magnitudes A and B might be familiar hue magnitudes like reddishness and greenishness, or they might be hue magnitudes alien to human color space. In any case, for the non-relationist, these will be response-independent hue magnitudes. On a physicalist version of non-relationism, we might suppose that they are highly unnatural or "anthropocentric" physical magnitudes (such as the degree of difference between an object's propensity to reflect long- and medium-wavelength light (cf. Bradley and Tye 2001). On a non-reductive version of non-relationism, we might suppose they are relatively natural (perhaps primitive) magnitudes distinct from any magnitude of interest to the physical sciences. The same story will hold, *mutatis mutandis*, for Maxwell's C-D channel. Similarly, Twin Maxwell's visual system will also have a pair of opponent channels, which we'll call the A*-B* channel and the C*-D* channel, which function in an analogous fashion. I'll assume that $A \neq A^*$, $B \neq B^*$, $C \neq C^*$, and $D \neq D^*$.

Now, we're supposed to imagine that Maxwell's state, S_M , involves the activation of both opponent channels. We'll suppose the activation levels on each are positive. In that case, the activation state of the A-B channel represents the stimulus as (say) A to degree x ($x > 0$), and the activation state of the C-D channel represents the stimulus as (say) C to degree y ($y > 0$). From here, it's natural to suppose that the composite state S_M semantically binds these contents together and represents the stimulus as having the conjunctive property *A to degree x & C to degree y* .

Turning to Twin Maxwell, we're supposed to imagine that his state, S_{TM} , involves the activation of only one opponent process channel—let's say the A*-B* channel— together with baseline (zero-level) activity on the other channel. So Twin Maxwell's state will presumably represent the object as (something like) *A* to degree z ($z > 0$) & C* to degree 0*. The conjunctive property represented by S_M , i.e. *A to degree x & C to degree y* , may be coextensive with the conjunctive property represented by S_{TM} , i.e. *A* to degree z & C* to degree 0*. They may even be coextensive across nearby possible worlds, or even across all possible worlds. But given a sufficiently fine-grained conception of properties (e.g. that defended in §3.5.6), it's plausible that they will nonetheless be distinct properties. They are, after all, constructed from distinct hue magnitudes. Since these conjunctive properties are response-independent (and so “non-relational” in the sense relevant to Cohen's argument), we get the result that C2 is false.

Apart from the question of whether Cohen's argument against selectionism is ultimately successful, there are two features of his argument worth highlighting: first, the argument does not rely on the (actual) phenomenon of perceptual variation. Second, it is

not an argument against selectionism specifically; it is, rather, an argument against non-relationist views generally. The argument, in other words, does not rely on the assumption that ALL-RIGHT is true. Now, either the argument succeeds or it doesn't. If it doesn't, then although we may follow Cohen in accepting ALL-RIGHT, we still have no reason to prefer relationism to non-relationism. If it does succeed, then it succeeds in establishing relationism all by itself, without the assistance of ALL-RIGHT. In that case, ALL-RIGHT turns out to be an idle wheel in the argument for relationism. Either way, ALL-RIGHT is doing no work. But the whole point of Cohen's extended discussion of the (actual) phenomenon of perceptual variation was to establish ALL-RIGHT. So it looks like the phenomenon of perceptual variation likewise is doing no work at all in Cohen's "argument from perceptual variation" for relationism.

I conclude that the phenomenon of perceptual variation provides no support for relationism. Given that the phenomenon of perceptual variation is, for many relationists, the primary motivation for relationism, this is bad news for the relationist. But it gets worse. Not only does perceptual variation provide no support for relationism. It furnishes the materials for a powerful argument *against* relationism.

2.4.2 The Argument from Perceptual Variation Against Relationism

In broad strokes, the argument from perceptual variation against relationism proceeds as follows.

P1. NONE-RIGHT is false.

P2. ALL-RIGHT is false.

P3. If relationism is true, then either NONE-RIGHT is true or ALL-RIGHT is true.

C. Therefore, relationism is false.

Let us consider the premises in turn. P1 is equivalent to

SOME-RIGHT: some variant is veridical.

The case for SOME-RIGHT is straightforward:¹⁹

SR1. The chip is green.

SR2. If the chip is green, then it is some determinate shade of green.

SR3. If the chip is some determinate shade of green, then some variant is veridical.

SR4. Therefore, some variant is veridical.

The justification for SR1 comes from the fact that the chip *looks* green; it visually *seems* or *appears to be* green. Presumably if the chip looks green to you, this fact provides at least *prima facie* justification to believe that it really is green. Here I am presupposing something like the Phenomenal Conservative principle mentioned in §2.2: if it seems to one that p, then one has at least *prima facie* justification to believe that p. In other words, if it seems to one that p, then absent sufficiently strong defeaters, one has justification to believe that p. The opponent of SOME-RIGHT may well have no problems with the Phenomenal Conservative principle. But she will maintain that there *are* defeaters. Most of these will be considered in §4 when I defend realism. For now, I want to point out that there at least aren't any defeaters for SR1 stemming from the facts of

¹⁹ It might be thought that in the present dialectical context, I can take SOME-RIGHT for granted. After all, on the face of it, denying SOME-RIGHT is more-or-less tantamount to accepting the irrealist view that the objects in our environment aren't really colored, which the relationist denies. But as we'll see below, it turns out that there are reasonable versions of relationism that reject SOME-RIGHT while avoiding the irrealist view that the objects in our environment aren't colored at all. For this reason, it's worth spending some time supporting SOME-RIGHT.

perceptual variation themselves. The epistemic status of SR1, which attributes the *coarse-grained* color green to the chip, therefore differs significantly from that of a proposition attributing a *fine-grained* color to the chip, such as:

U. the chip is unique green.

Suppose the chip looks unique green to you. Given the Phenomenal Conservative Principle, you have *prima facie* justification to believe that the chip is unique green. But then suppose you learn the relevant facts about perceptual variation. You learn, for example, that the chip looks yellowish green to other normal perceivers—perceivers who navigate their environment just as well as you, who do just as well on color discrimination tests as you do, and so on. Assuming the chip can't be both unique green and yellowish green, it's plausible that learning these facts undercuts your justification to believe U. But in a case like the one I've described, where the chip looks unique green to some normal perceivers, yellowish green to other normal perceivers, and blueish green to others, it will look *green* to all normal perceivers in normal conditions of illumination. (In any case, it will look *greenish* to all normal perceivers, and the argument will go through just as well if we substitute “greenish” for “green.”) So we have *prima facie* justification to believe the chip is green, and this time the facts of perceptual variation do not undercut our justification, for all the variants agree that it is green.

Consider an analogy, adapted from a familiar case in the peer disagreement literature (Kelly 2010). You and your friend, both normal perceivers in equally favorable viewing conditions, are watching the final seconds of a close horse race between horse A and horse B. As A and B cross the finish line, it looks to you as though A finishes at

approximately the same time as B, but more specifically it looks to you as though A finishes *just barely* before B. Put another way: A appears to you to have the coarse-grained property of finishing at approximately the same time as B, and also appears to you to have the fine-grained property of finishing just barely before B. (Note that the property of finishing just barely before B stands in something like a determinate-determinable relation with the property of finishing at approximately the same time as B. Finishing just barely before B is one of several ways of finishing at approximately the same time as B; other ways include finishing just barely after B and finishing at exactly the same time as B.) You then find out that it also looked to your friend as though A finished at approximately the same time as B, but more specifically it looked to him as though A finished just barely *after* B. Here, it's fairly plausible that upon learning about how things seem to your friend, you should give up, or at least significantly reduce your confidence in, your initial fine-grained judgment. But there's no pressure in this case to give up your coarse-grained judgment. You should continue to believe that A has the coarse-grained property of having finished at approximately the same time as B, even if you can't reasonably take a stand on which fine-grained "determinate" of this property A has. Similarly, if other normal perceivers agree on the coarse-grained color of the chip, and disagree only on its fine-grained color, then at most we should give up our judgments about the fine-grained color, not our judgments about its coarse-grained color. So we ought to accept premise SR1.

Turning to the other two premises: Premise SR2 is an instance of a very plausible general metaphysical principle about the relation between determinates and determinables, which I shall call the Principle of Determination:

Principle of Determination: For any object x and determinable property F , if x has F , then x has some determinate of F .

Finally, premise SR3 is justified by the setup of the case, which stipulates that there are enough variants to cover all the determinate shades of green (at least within a given sizable subregion of the green portion of color space).

P2, the second premise of our argument against relationism, asserts that ALL-RIGHT is false. Our justification to accept P2 comes from the *Conflict Intuition* introduced in §2.4.1. If a chip looks unique green (all over) to John and yellowish green (all over) to Jane, there's a strong intuitive that they can't *both* be getting the color of the chip right; at least one of their experiences must be misrepresenting the chip. The Conflict Intuition that arises in cases like this straightforwardly yields the falsity of ALL-RIGHT.

Relationists typically deny what I'm calling the Conflict Intuition. Is the appeal to the Conflict Intuition in an argument against relationism therefore question-begging? No. Asserting P as a premise in an argument for C is question-begging only if the justification for accepting P depends on our having antecedently accepted C. So appeal to the Conflict Intuition in an argument for the denial of relationism is question-begging only if our justification to accept the Conflict Intuition depends on our having antecedently rejected relationism. But the claim that John and Jane can't both be getting the color of the chip right in the above scenario is simply a pre-theoretic intuition. Our justification to accept it

does not depend on our having already rejected relationism, nor indeed does it depend on any prior substantive theoretical commitments.

I've argued that both NONE-RIGHT and ALL-RIGHT are false. If I'm correct, it follows that ONLY-SOME-RIGHT is true. We should only be prepared to deny NONE-RIGHT and ALL-RIGHT on the basis of my arguments if we're satisfied that there are no objections of comparable (or greater) strength against ONLY-SOME-RIGHT. It's therefore worth pausing to consider Cohen's objections to ONLY-SOME-RIGHT. His objection to ONLY-SOME-RIGHT is that, although there is not "any inconsistency in the thought that one perceptual variant is veridical at the expense of the others, [...] it's hard to see what (besides *ad hoc* stipulation) could make this the case" (p. 25). Because ONLY-SOME-RIGHT requires *ad hoc* stipulation; we should therefore reject ONLY-SOME-RIGHT because "we should take measures to avoid *ad hoc* stipulation when possible" (p. 25).

I find this objection to ONLY-SOME-RIGHT deeply puzzling. To begin with, his claim that "it's hard to see what (besides *ad hoc* stipulation)" could make it the case that one variant is veridical at the expense of the others seems to suggest that *ad hoc* stipulation *could* make it the case that one variant is veridical at the expense of the others. If that were so, Cohen would be fated to lose this debate. For in that case, I could win the debate simply by using my powers of *ad hoc* stipulation to make my position true.

More importantly, as many reviewers of Cohen's book have pointed out (Pautz 2010b, Tye 2012, cf. Byrne and Hilbert 2003a), it's not at all clear why ONLY-SOME-RIGHT is supposed to require *ad hoc* stipulation. Cohen's justification for this claim seems to be that "there is no independent, well-motivated reason for singling out any single

variant as veridical (at the expense of the others)” (p. 24), so the only way to single out any particular variant as veridical would be through *ad hoc* stipulation. I agree with Cohen that in many cases of perceptual variation, there is no independent, well-motivated reason for singling out any particular variant as veridical. If a chip looks unique green to John and yellowish green to Jane, and both are normal perceivers viewing the chip in normal viewing conditions, there may well be no independent, well-motivated reason for singling out John’s experience as the veridical variant at the expense of Jane’s, and also no independent, well-motivated reason for singling out Jane’s experience as the veridical variant at the expense of John’s.²⁰ So if I were to single out a particular variant—say, John’s experience—as veridical, then perhaps this would amount to “*ad hoc* stipulation.” (In any case, it would amount to *something* objectionable.) Fortunately, the proponent of ONLY-SOME-RIGHT needn’t single out any particular variant as veridical at the expense of the others. She needn’t say *of* any variant that *it* is veridical and the others aren’t. ONLY-SOME-RIGHT is just the existential claim that *some* variant is veridical and others aren’t. Asserting this existential proposition doesn’t require any stipulation, *ad hoc* or otherwise. Still less does the *truth* of this existential claim require *ad hoc* stipulation. Rather, it requires some variant to be veridical and others to be non-veridical. The veridicality or non-veridicality of an experience is a matter of whether the object of the

²⁰ I’m assuming that Cohen means that there is no independent and well-motivated reason for asserting/believing *of* any variant that *it* is veridical at the expense of the others. If he intends the *de dicto* reading—that there is no independent and well-motivated reason for asserting/believing the proposition that one of the variants is veridical and the others aren’t—then the use of the phrase “singling out any single variant as veridical” is singularly inapt. If I say to a room of people, “Exactly one of you stole my pencil,” but I don’t say *of* anyone in the room that *he* stole my pencil, then I clearly haven’t singled out any single person as the thief. More importantly, if he intends the *de dicto* reading, then his claim is false for reasons to be given in the next (main text) paragraph.

experience has the property or properties the experience represents it as having, and neither the properties represented by experience nor the (relevant) properties of objects are a matter of stipulation.

While there may be no independent and well-motivated reasons to accept any particular instance of this existential claim, the same cannot be said of the existential claim itself. The premises in the arguments given in the last two sections—SR1-SR3 and the Conflict Intuition—jointly entail the existential claim, and I’ve shown that each of these premises is well-motivated. This much should be granted even if one thinks that we should ultimately reject one of these premises on the basis of further considerations.

Setting aside Cohen’s confusing remarks about “*ad hoc* stipulation,” the crux of the objection seems to be Cohen’s claim that it’s hard to see what could make it the case that one of the variants is veridical at the expense of the others. But I don’t understand what’s hard to see here. Consider again our case in which a color chip looks unique green to John and yellowish green to Jane. Consider the hypothesis—call it H—that one of their experiences is veridical and the other is non-veridical. What could make H the case?

Well, either of the following would do the trick:

Case 1: John’s experience is veridical and Jane’s experience is not veridical.

Case 2: Jane’s experience is veridical and John’s experience is not veridical.

Of course, my opponent will press on: “But what could make it the case that (say) Case 1 obtains?” Answer: the following four conditions would certainly do the trick (cf. Byrne and Hilbert 2007b, p. 88-9; Byrne 2006, p. 337): (The first pair would ground the first conjunct of Case 1, and the second pair would ground the second conjunct.)

- (i) The chip looks unique green to John.
- (ii) The chip is unique green.
- (iii) The chip looks unique yellow to Jane.
- (iv) The chip is not unique yellow.

Likewise, *mutatis mutandis*, for Case 2. (Compare: suppose a certain rod looks straight to John but looks bent to Jane. What could make it the case that John’s experience is veridical while Jane’s is not veridical? Answer: the following four conditions would suffice: (i*) the rod looks straight to John, (ii*) the rod is straight, (iii*) the rod looks bent to Jane, (iv*) the rod is not bent.)

It’s difficult to deny that if facts (i)-(iv) obtained, they would make it the case that Case 1 obtains, and hence that “one variant is veridical at the expense of the others.” But Cohen might press the same question again for (i)-(iv). Although (i)-(iv) would obviously suffice to make ONLY-SOME-RIGHT true, he might say that we should not accept the conjunction of (i)-(iv) because it’s hard to see what could make (i)-(iv) obtain. Now, we needn’t bother with the question of what could make (i) and (iii) obtain, because there is no dispute about the truth of (i) and (iii). Cohen is not prepared to reject (i) or (iii) even if we are unable to explain what could make them the case. And since the “what could make it the case that P?” question is here being used to motivate the rejection of P, we needn’t consider this question for instances of P which are not in dispute. So the concern must be that it’s hard to see what could make (ii) and/or (iv) obtain—in other words, that it’s hard to see what could make it the case that

- U. the chip is unique green.

and/or what could make it the case that

NY. The chip is not yellowish green.

Now, there may be puzzles about what could make NY the case, but these are just instances of highly general puzzles about the grounds for negative facts. To take another instance of the same puzzle: the room is 74 degrees Fahrenheit. It is *not* 80 degrees Fahrenheit. What makes it the case that—what grounds the fact that— the room is not 80 degrees? I'm not sure. One possibility is that negative facts like this are brute or explanatorily fundamental facts. Another possibility is that this negative fact is explained by the positive fact that the room is 74 degrees, perhaps together with the exclusion principle that being 74 degrees is incompatible with being 80 degrees. Another possibility is that the negative fact is explained by whatever positive facts ground the fact that the room is 74 degrees (e.g. that the molecules in the room are moving in such-and-such specific way), perhaps together with the exclusion principle that these facts are jointly incompatible with each collection of facts that would ground the room's being 80 degrees. Whatever we say about this case we can apply, *mutatis mutandis*, to NY. There is no *special* puzzle about what could make NY the case.

So if there is any difficulty in seeing what could make it the case that one of the variants is veridical at the expense of the others, it must be because there is some difficulty in seeing what could make it the case that the chip is unique green. But what's supposed to be the problem here? If unique green is a certain reflectance type, then what makes it the case that the chip is unique green is the way the chip alters the incident light, which in turn is explained by the microphysical structure of the chip's surface, the laws of

electromagnetism, and so forth. (But what could make it the case that unique green is that particular reflectance type? Bad question. Identity facts don't admit of explanation (cf. Lewis 1986, pp. 192-3).) On the other hand, if unique green is an irreducible quality, then one possibility is that *nothing* makes it the case the chip is unique green; in other words, the fact that the chip is unique green is explanatorily basic. (Explanation has to end somewhere!) Another possibility is that the chip's being unique green is grounded in the way its surface alters the incident light, even though being unique green is not identical to a way of altering incident light. (Compare: Whisker's furriness is grounded in his microphysical properties, but the property of being furry is (probably) not identical to any microphysical property—i.e. any property definable in the vocabulary of microphysics.) Another possibility, which I'm inclined to accept (see chapter 5), is that the chip's being unique green is grounded (not in its dispositions to alter incident light, but) in certain categorical microphysical features of the chip (which themselves ground its dispositions to alter the incident light).

Until we're given some reason to think that such features of the chip could not make the chip unique green, I conclude there is no difficulty in seeing what could make it the case that one variant is veridical while there others aren't, and hence no reason (as yet) to reject ONLY-SOME-RIGHT.

The third and final premise of the argument from perceptual variation against relationism is the conditional: if relationism is true, then either NONE-RIGHT is true or ALL-RIGHT is true. To support this, I shall consider a range of different versions of relationism and show that each of them either has the consequence that NONE-RIGHT is

true or the consequence that ALL-RIGHT is true. I shall then motivate the claim that the same holds not just of the versions of relationism I consider, but for all reasonable forms of relationism. As we'll see, although it *is* possible to devise a version of relationism that does not entail ALL-RIGHT or NONE-RIGHT, such versions of relationism turn out to have deeply unattractive features.

On a simplistic form of relationism, when the chip looks green to John, John's experience represents the chip as having the disposition to look green. Likewise, when the chip looks yellowish green to Jane, Jane's experience represents the chip as having the disposition to look yellowish green. Does the chip have these properties? It depends on how we interpret "the disposition to look unique green" and "the disposition to look yellowish green." Different interpretations yield different versions of relationism:

ALL: John's experience (Jane's experience) represents the chip as having the disposition to look unique green (yellowish green) to *all* perceivers under *all* conditions.

SOME: John's experience (Jane's experience) represents the chip as having the disposition to look unique green (yellowish green) to *some* perceivers under *some* conditions.

Given the empirical facts of perceptual variation, ALL yields the result that NONE-RIGHT is true, for there is no determinate color C such that the chip is disposed to look C to all perceivers under all conditions. On the other hand, SOME yields the result that ALL-RIGHT is true.

Other relationist views make restrictions to *normal* perceivers and *normal* conditions. On such views, when a chip looks unique green to John, the property his visual experience represents the object as having is the disposition to look unique green to normal perceivers in normal conditions. Likewise, when the chip looks yellowish green to Jane, the property her visual experience represents the chip as having is the disposition to look yellowish green to normal perceivers in normal conditions. Does the chip have these properties? Again, it depends on how we interpret “the disposition to look unique green [yellowish green] to normal perceivers under normal conditions.” Different versions of relationism result from different interpretations of such expressions. Here are two:

NORM-ALL: John’s experience (Jane’s experience) represents the chip as having the disposition to look unique green (yellowish green) to *all* normal perceivers under *all* normal conditions.

NORM-SOME: John’s experience (Jane’s experience) represents the chip as having the disposition to look unique green (yellowish green) to *some* normal perceivers under *some* normal conditions.

NORM-ALL, together with the empirical facts of normal variation, entail NONE-RIGHT. The chip is not disposed to look unique green to all normal perceivers in all normal conditions; Jane is a counterexample. Nor is the chip disposed to look yellowish

green to all normal perceivers in all normal conditions; John is a counterexample.²¹ On the other hand, NORM-SOME, together with the empirical facts of normal variation, entails ALL-RIGHT. (I'm assuming here that all the variants at issue are experiences of normal perceivers in normal conditions.)²²

Other relationist views hold that the color represented by a given color experience is relativized to particular perceivers (the subject of the experience) and particular viewing conditions (the fully specific condition of the subject's current experience). If we let C_{John} be the fully specific viewing condition in which John finds himself and we let C_{Jane} be the fully specific viewing condition in which Jane finds herself, then this form of relationism would hold that:

²¹ It's worth noting that although NORM-ALL entails NONE-RIGHT, it does not entail the irrealist thesis that nothing is colored. Although the chip doesn't present any particular *fine-grained* color appearance to all normal perceivers under all normal conditions, the chip does present a common *coarse-grained* color appearance to all normal perceivers in normal conditions. In particular, it looks *green* (or at least greenish) to all normal perceivers under all normal conditions. So NORM-ALL will have the result that the chip has a coarse-grained color but no fine-grained color. In my view, NORM-ALL is the most attractive form of relationism among those considered in this section. Unlike ALL, it allows us to accept realism about coarse-grained colors like red, green, blue, and purple. Unlike SOME, NORM-SOME, and SPEC (introduced below), it does not violate the Conflict Intuition. And unlike SPEC, it can accommodate the fact that sometimes two perceivers, or a single perceiver in different viewing conditions, perceptually represent the same color. The main problem with NORM-ALL is that it seems to be inconsistent with the Principle of Determination mentioned in §2.4.2.

²² There are other ways to interpret expressions like "the disposition to look unique green (yellowish green) to normal perceivers under normal conditions," e.g. as denoting the disposition to look unique green (yellowish green) to *most* normal perceivers under *most* normal conditions, or to some normal perceivers under all normal conditions, or to all normal perceivers under some normal conditions, or any of the other results of permuting the quantifiers *all*, *some*, and *most*. But each of the corresponding versions of relationism will entail either ALL-RIGHT or NONE-RIGHT for reasons similar to those just given for NORM-ALL and NORM-SOME. (The same seems to hold if we work with other natural-language quantifiers, such as *few*, *many*, *6% of*, the GEN quantifier, and any other non-gerrymandered quantifier that comes readily to mind.)

SPEC: John's (Jane's) experience represents the chip as having the disposition to look unique green (yellowish green) to John (Jane) in C_{John} (C_{Jane}).

SPEC, which happens to be the version of relationism endorsed by Cohen, yields the result that ALL-RIGHT is true. For the chip has the disposition to look unique green to John in John's specific viewing condition, and it also has the disposition to look yellowish green to Jane in Jane's specific viewing condition, so both experiences veridically represent the chip. The same will be true of all variants in such cases of perceptual variation.

Now, as I've formulated ALL, SOME, NORM-ALL, and SPEC, each of these is a form of what I earlier called *dispositionalist* relationism (§2.2). But nothing here hangs on the dispositionalist formulation. The non-dispositionalist counterparts of each of these views (i.e. the result of replacing "the disposition to look" with "occasionally looks" in the formulation of each version) will also entail the truth of either ALL-RIGHT or NONE-RIGHT. Nor does anything hang on formulating the views in terms of the way objects (are disposed to) *look*. We could also replace expressions like "look unique green" with expressions like "produce unique-green-ish sensations," "produces an experience of a unique-green-prime region of the visual field" (Peacocke 1984), and the resulting forms of relationism would have the same result.

Each of the version of relationism we've considered—ALL, SOME, NORM-ALL, NORM-SOME, and SPEC—have the result that either ALL-RIGHT is true or NONE-RIGHT is true. It seems to me that these five versions of relationism (along with relevantly similar views such as those mentioned in the previous paragraph and in fn. 22) exhaust the *reasonable* relationist positions. Now, there are other possible relationist positions, and

some of these would even have the result that ONLY-SOME-RIGHT is true, and hence avoid the result that either NONE-RIGHT or ALL-RIGHT is true. For example, one could accept the following position:

SPEC-JOHN: John's experience represents the chip as having the disposition to look unique green to John in C_{John} . Jane's experience represents the chip as having the disposition to look yellowish green to John in C_{John} .

Since the chip is disposed to look unique green to John in C_{John} , and isn't disposed to look yellowish green to John in C_{John} , SPEC-JOHN yields the result that ONLY-SOME-RIGHT is true. But no one could take such a proposal seriously. It's absurd to suppose that Jane's experience represents color properties that are relativized specifically to *John* and *his* viewing conditions. And indeed, it seems that any version of relationism that has the result that ONLY-SOME-RIGHT is true will be deeply unattractive for much the same reasons. Since all *reasonable* forms of relationism have the result that either NONE-RIGHT or ALL-RIGHT is true, I conclude that premise P3 is true. Given P1 and P2 (defended above), it follows that relationism is false.

Chapter 3: Against Physicalism

3.1 Introduction

Physicalist theories of a class of sensible qualities identify the qualities in that class with response-independent physical properties. “Physical property” here is intended in a fairly narrow sense to mean a property that admits of real definition in terms of the canonical vocabulary of the physical sciences. Properties that merely supervene on, but are not identical to, such properties do not qualify as “physical” in my sense.

In objecting to physicalist theories of the secondary qualities, I focus once again on color, though as in the previous chapter, many of my criticisms are applicable to other secondary qualities as well. Moreover, as before, it is antecedently plausible that if physicalism is false about color, it is also false about the other secondary qualities. I’m not aware of anyone who accepts a non-physicalist theory of color together with a physicalist view of (say) flavor, scent, or timbre. So if we can show that physicalism about color is false, then it’s a fair bet that physicalist theories of other secondary qualities are false as well.

In this chapter, I discuss three objections (better, three kinds of objection) to physicalist theories of color: the *structure-mismatch* objection(s), the *categoricity* objection, and a class of objections that may be called *disjunctive-property* objections, with a special focus on one member of this class, which I call the *similarity-grounding* objection.

3.2 The Structure-Mismatch Objection:

The structure-mismatch objection is one of the most important objections to physicalist theories of color and other secondary qualities. However, my discussion of it will be extremely brief, as I have nothing of significance to add to the work of Hardin (1988), Maund (1995), Pautz (2006b), and others who have developed and defended the objection in great detail. In broad strokes, the structure-mismatch objection begins with the empirical premise that the colors in human color space jointly exhibit certain structural features which are not possessed by any collection of physical properties, or at least any collection of physical properties that could reasonably be identified with the colors in human color space. Now, if a collection C of properties exhibits a structural feature (or any other feature, for that matter) which is not possessed by any collection of X-properties, then by Leibniz's law it follows that C is not identical to any collection of X-properties. It therefore follows from the empirical premise above that colors are not identical to any collection of physical properties.

One important aspect of the structure of human color space is its *resemblance* structure. There is a large class of apparent second-order truths about the colors, such as that red is more similar to orange than it is to green, or that scarlet is more similar to crimson than it is to canary yellow, that collectively embody the apparent resemblance structure of color space. This resemblance structure does not seem to match that of any collection of physical properties that could reasonably be identified with the colors. Other structural features of human color space, such as the distinction between unitary colors (red, green, blue, yellow) and binary colors (e.g. orange, purple), also seem not to be

present in any relevant collection of physical properties. As Hardin (1997) writes, “the unitary-binary structure of the colors as we experience them corresponds to no known physical structure lying outside nervous systems that is causally involved in the perception of color” (p. 300). Related difficulties arise for physicalist theories of other secondary qualities as well (Pautz 2006a, 2010a; Hartmann 1997, ch. 12).

As with other apparent second-order truths about the colors, it is possible to hold that the relevant appearances are illusory. But the fact that colors appear to stand in such-and-such resemblance relations gives us *prima facie* grounds to think they really do stand in these resemblance relations, and therefore gives us *prima facie* grounds to reject color physicalism.

3.3 The Categoricity Objection

In §2.2 I objected to dispositionalist forms of relationism on the grounds that it contradicts an apparent second-order truth about color, namely:

Categoricity: Colors are categorical properties.

Just as there are dispositionalist forms of relationism, there are dispositionalist forms of physicalism, i.e. forms of physicalism according to which colors are identical to certain dispositional physical properties. The most popular form of dispositionalist physicalism (and probably the most popular form of physicalism) is what may be called “light-dispositionalism” (Byrne and Hilbert 2003a, Tye 2000, Lewis 1997, Matthen 1988).

According to light-dispositional theories of color, colors are dispositions to alter light in certain ways. For example, according to one straightforward version of light-dispositionalism, colors are surface spectral reflectances— complex dispositions to reflect

a certain proportion of incident light at each wavelength within the visible spectrum. The apparent truth of Categoricity gives us *prima facie* grounds to reject light-dispositionalism along with any other dispositionalist view of color, physicalist or otherwise.

3.4 Disjunctive Property Objections

In this section, I consider a class of objections, which I call *disjunctive-property objections*. A disjunctive-property objection is any instance of the following argument schema:

D1. If color physicalism is true, then color shades are disjunctive (physical) properties.

D2. Color shades have feature F.

D3. No disjunctive property has feature F.

D4. Therefore, color physicalism is false.

(D1, which is common to all instances of the schema, is an empirical premise, the evidence for which will be reviewed in §3.5.1.) We get an instance of this schema by plugging in some specific property for F in D2 and D3. I begin by discussing very briefly a handful of instances of this schema, each of which is, I think, somewhat promising (though I am not prepared to endorse any of them outright). Then in §3.5, I develop and defend at length one further instance of this schema.

F = Being Causally Efficacious: It's plausible that, typically, scarlet things look scarlet because they are scarlet. If so, it seems to follow that colors are causally efficacious, at least with respect to our color experiences. But it's sometimes said that

disjunctive properties are never causally efficacious (Watkins 2005, pp. 37-9; cf. Johnston 1992, pp. 235-6). Given D1, it then follows that color physicalism is false.

F = Being an Object of Acquaintance: We are acquainted with colors if we are acquainted with any properties at all. But arguably, one cannot be directly acquainted with disjunctive properties. More cautiously, it can be argued that color experience of any determinate color shade C does not involve acquaintance with any disjunctive property, on the following grounds: just as (arguably) one can only be acquainted with a conjunctive property if one is acquainted with each of its conjuncts (e.g. one can only be acquainted with the conjunctive property being red and round if one is acquainted with redness and also acquainted with roundness), one can only be acquainted with a disjunctive property if one is acquainted with each of its disjuncts. From this it follows that if some determinate color shade C is a disjunctive property, then in being acquainted with C, one is acquainted with some properties which are more specific than C, namely C's disjuncts. But C is a determinate color shade. In experiencing C, we are not experiencing any color shade which is more fine-grained or more specific than C. So it seems implausible that, in experiencing C, there is any more specific property with which we are acquainted.²³

F = Existence: Proponents of sparse properties will want to deny the very existence of disjunctive properties (Armstrong 1978). But a proponent of sparse

²³ I am ignoring conjunctions of C with other sensible properties that one's experience predicates of the same object, such as shapes. Such properties will, of course, be more specific than C, and may also be objects of our acquaintance, but they will not be disjuncts involved in any credible physicalist reduction of C.

properties might still have reason to accept the existence of sensible qualities like colors. When one has a hallucination as of something red₁₇, while there is no concrete physical object of which one is aware, many have found it plausible that we are nonetheless aware of or acquainted with *something* (Johnston 2004, Hawthorne & Kovakovich 2006). Unless we are prepared to embrace the ontological extravagancies of sense data theory, the most plausible candidate for this something is the property red₁₇. And if one is aware of or acquainted with red₁₇, it seems to follow that red₁₇ exists. If disjunctive properties don't exist, then given D1, it follows that color physicalism is false.

F = Being wholly present in its instances: When one visually experiences some object as having a specific color, it is somewhat natural to describe one's phenomenology by saying that the color itself seems to be wholly present in the object of one's experience. One might be tempted to say that this is a category mistake, that colored *objects* have locations, but colors (like all properties) are not located anywhere. I share this temptation. But don't we experience colors as having locations? When the surface of a tomato looks red, doesn't the redness (and not just the tomato's surface) appear to be located *out there*, present in the object's surface?)

If indeed we experience colors as being wholly present in the objects of visual experience, then in cases of veridical color perception, the color really is wholly present in the object of one's experience. If this is correct, then the colors we visually experience must be properties of a sort which can be wholly present in their instances. But in general it is not plausible that disjunctive properties are wholly present in their instances.

Suppose *o* has *P* and lacks *Q*. Then *o* has the disjunctive property $P \vee Q$. Assuming that a

complex property is wholly present in an object only if its constituent properties are, and that an object instantiates any property which is wholly present in it, it follows that $P \vee Q$ is not wholly present in o . So if colors are wholly present in their instances, then given D1, we ought to conclude that color physicalism is false.

3.5 The Similarity-Grounding Argument

In this section, I consider—and defend at length—one further instance of the argument schema above, which I call the *similarity-grounding argument*:

SG1. If color physicalism is true, then color shades (i.e. the fine-grained colors we visually experience) are disjunctive (physical) properties.

SG2. Color shades are similarity-grounding properties—roughly, properties the sharing of which makes for genuine similarity.

SG3. No disjunctive property is a similarity-grounding property.

SG4. Therefore, color physicalism is false.

Most of what follows will consist of a clarification and defense of SG1-SG3. In §3.5.1, I review familiar empirical considerations which suggest that any tenable form of color physicalism will have to identify color shades with disjunctive physical properties. In §§3.5.2-3.5.4, I offer a much more precise characterization of the notion of a similarity-grounding property in terms of the notions of metaphysical grounding and what I call “weak similarity,” contrasting it with some related notions, including Lewisian “naturalness” (Lewis 1983). In §3.5.5, I defend the claim that it is evident in our ordinary experience of colors that they are similarity-grounding properties. In §§3.5.6-3.5.9, I clarify the notion of a “disjunctive property” and argue that disjunctive properties are

(probably) never similarity-grounding properties. Moreover, I argue that even if there are some exceptions to this principle, the kinds of physical disjunctive properties which might reasonably be identified with color shades are plausibly not among them.

3.5.1 Color Physicalism Leads to Color Disjunctivism

Color physicalism comes in two main varieties: light-dispositional physicalism (Byrne and Hilbert 2003a, Tye 2000, Lewis 1997, Armstrong 1999, Matthen 1988) and categorical physicalism (Jackson and Pargetter 1987, Armstrong 1968a, Smart 1975).

According to the former, colors are dispositions to alter light in certain ways. According to the latter, colors are categorical (micro-)physical properties of objects, perhaps those chemical-structural properties which serve as the categorical physical bases of objects' dispositions to alter light in the way they do and influence our perceptual systems in the way they do. In this section, I briefly review familiar empirical reasons for thinking that each of these forms of physicalism leads to “disjunctivism”—the view that color shades are disjunctive properties.

A simplistic version of light-dispositionalism would identify color shades with surface spectral reflectances (SSRs). The SSR of a surface is given by the proportion of incident light that it is disposed to reflect at each wavelength within the visible spectrum. It is widely agreed by both proponents and opponents of light-dispositionalism that this simplistic version of light-dispositionalism is untenable due to the phenomenon of *metamerism*. Let us say that two objects are a metameric pair just in case they have very different SSRs, but normal perceivers experience them to be identical in respect of color relative to some normal illuminant. The existence of metameric pairs creates difficulties

for the identification of color shades with specific SSRs. Suppose o_1 and o_2 are a metameric pair which, in some normal condition of illumination, you experience as having a determinate color C. Provided we reject irrealism—and physicalism, like other reductive theories of color, is largely motivated by the desire to avoid irrealism—then it seems we have no reason to think that your experience of either member of the metameric pair is non-veridical. But if your experiences of both o_1 and o_2 are veridical, then both o_1 and o_2 have the color shade C that they appear to have. Since o_1 and o_2 have drastically different SSRs, it follows that C cannot be identical to any specific SSR. Moreover, as Byrne and Hilbert (2003a) note,

The various reflectances that are perceptually equivalent (with respect to a given illuminant) are not just minor variants of each other. Surfaces with grossly different reflectances can perceptually match even under fairly normal illuminants. (p. 11)

Because there is no specific SSR or natural type of SSR that can plausibly be identified with any given color shade, the light-dispositional physicalist is apparently forced to identify color shades with long disjunctions of specific SSRs. Light-dispositional physicalists tend to embrace this conclusion: “colors are reflectance types, or disjunctions of reflectances [...] In the terminology of Lewis (1983) they are not very ‘natural’” (Byrne and Hilbert 2007a, p. 75).

As in the above passage, light-dispositional physicalists sometimes call the relevant disjunctions of reflectances “reflectance types.” But this label is potentially misleading. Calling these disjunctions of reflectances “reflectance types” is rather like

talking about the “length type” that something has just in case it is either 3.2 cm long, 79.672 cm long, or 122.38 cm long, or the “shape type” that something has just in case it is either an isosceles trapezoid, a regular chiliagon, or a concave hexagon. In other words, the disjuncts of the relevant disjunctions of reflectances do not form anything resembling a natural class; they do not jointly exhibit the kind of unity we normally expect of a collection that is said to constitute a “type.”

Metamerism is not the only phenomenon that pushes the light-dispositionalist toward disjunctivism. Another comes from objects whose perceived color is determined not (only) by the way they reflect light but by the way they transmit or emit light, such as filters, volumes, and colored lights. Unless the physicalist is prepared to convict ordinary color experience of light transmitters and emitters of systematic error, the physicalist will have to account for the apparent fact that a non-reflective object can exemplify the same color shade as a reflective object. This will require identifying color shades with physical properties which are shared by reflective and non-reflective objects alike, such as the “productance types” of Byrne and Hilbert (2003a). Doing so will almost certainly commit the light-dispositionalist to an even more radically disjunctive view of color than is required just to accommodate the phenomenon of metamerism.

The route to disjunctivism from categorical physicalism is even more straightforward. The class of categorical physical properties which produce color experiences as of a given color shade will, in general, be even more diverse and dis-unified than the class of reflectances associated with such experiences. Indeed, the bewildering diversity in the categorical physical causes of color experiences as of a given color shade is what

motivates many physicalists to identify colors with higher-level dispositions to alter light in certain ways, which are multiply realized by a range of categorical physical properties. After surveying the wide variety of categorical physical properties that standardly cause an object to look blue, Hardin (1988) writes,

That variety suggests it would be vain to suppose that objects sharing a common color resemble one another in physical structure [...] It should thus be clear that, if blue things have any physical feature in common at all, it will reside not in the physical microstructure of those things, but rather in their dispositions to radiate light of a particular character from their surfaces. (pp. 4, 6)

For both categorical and light-dispositional forms of physicalism, then, there are strong empirical grounds to accept disjunctivism. Since these are the only two serious physicalist options, I conclude that SG1 is true: if color physicalism is true, then color shades are disjunctive (physical) properties.

3.5.2 Grounding

Premise SG2 asserts that color shades are *similarity-grounding* properties, properties the sharing of which makes for genuine similarity. Before defending this premise, it will be necessary to get clear on what it is saying, which in turn will require getting clear on what “making for” and “similarity” mean in this context. Let us take these expressions in turn.

To say of some shareable property F that it is a property the sharing of which *makes for* similarity is not merely to say that, necessarily, any two things which share F are similar. It is rather to say that, necessarily, any two things which share F are *ipso facto* similar—that they are similar in virtue of the fact that each has F. Here and throughout

this chapter, “in virtue of” is meant to express the converse of the relation of *metaphysical grounding* (Fine 2001, Schaffer 2009, Rosen 2010). To say that a fact A *metaphysically grounds* a fact B (or, alternatively, that B holds *in virtue of* A) is to say, roughly, that A (non-causally) makes it the case that B obtains. Like causation, grounding is an explanatory relation. If A grounds B, then A serves as the explanans in some complete metaphysical explanation for why B obtains. However, metaphysical grounding is importantly different from causation in at least two respects. First, whereas causation is paradigmatically *diachronic*, relating events or facts which obtain at different times, grounding is paradigmatically *synchronic*, relating events or facts which obtain at the same time. Second, as Hume is supposed to have taught us, causes do not metaphysically necessitate their effects. By contrast, grounds do metaphysically necessitate the facts they ground. In other words, if A grounds B, then it is metaphysically necessary that if A obtains, then B obtains. Nonetheless, it is important not to confuse metaphysical grounding with mere metaphysical necessitation. The latter differs from the former in at least two important respects. First, unlike metaphysical grounding, metaphysical necessitation is *not* an explanatory relation. If A metaphysically necessitates B, it does not follow that A explains why it is that B obtains. For example, it is metaphysically necessary that if Socrates’s singleton exists, then Socrates exists, and it is metaphysically necessary that if the proposition that grass is green is true, then grass is green. But Socrates’s existence is not explained by the existence of Socrates’s singleton, nor is the truth of the proposition that grass is green what makes it the case that grass is green. In fact, in these cases the order of explanation plausibly runs in the other direction.

Socrates's singleton exists because Socrates does, and the proposition that grass is green is true because grass is green. The second important difference between metaphysical grounding and metaphysical necessitation is that the former is hyperintensional, whereas the latter is merely intensional. Consider the following two claims:

S1. (Socrates is a philosopher or Socrates is a prime number) in virtue of the fact that Socrates is a philosopher.

S2. Socrates is a philosopher in virtue of the fact that Socrates is a philosopher.

Intuitively, S1 is true, but S2 is false. (The truth of S1 and the falsity of S2 follow respectively from two plausible and widely accepted principles in the theory of grounding— that disjunctive facts are grounded in their true disjuncts (Fine 2001, pp. 21-2; Rosen 2010, p. 117), and that grounding is irreflexive (Schaffer 2009, Rosen 2010, Fine 2010, Correia 2010)). However, S1 only differs from S2 in the substitution (on the left) of co-intensional sentences—sentences which agree in truth value at every possible world.

Using our notion of metaphysical grounding, we can draw a distinction between those facts that are grounded in other facts and those facts that are ungrounded (brute, fundamental). For our purposes, it is important to observe that there are some kinds of facts which must be grounded if they are to obtain at all. For example, facts about the wrongness of particular actions plausibly fall within this category. If a particular action is wrong, there will always be some non-moral features of the action which *make* it wrong (e.g. that it was done with the sole intention of causing suffering). The fact that an action is wrong is never a brute fact. Importantly, the same holds for facts about *similarity*. If

two things are similar, it will always be the case that they are similar in virtue of sharing some property or properties. Suppose that Joe's car is similar to Fred's car. The fact that they are similar is surely not a brute fact about them. If they are similar, then there will be facts about these cars that *make* them similar. For instance, they might be similar because —similar in virtue of the fact that—each is a mid-sized sedan, each is charcoal gray, each has a manual transmission, and so forth.

3.5.3 Similarity

As mentioned above, a sharable property F is a similarity-grounding property just in case, necessarily, if any two things share F, they are similar in virtue of the fact that each has F. I assume that “in virtue of” is reasonably clear in light of the above discussion. But how should we understand that notoriously slippery word “similar” in this context? It will not do to simply take the “ordinary sense” of the word, because the word has no context-invariant and precise sense within ordinary usage. Within ordinary usage, whether the predicate applies to a pair of objects in a given context often depends on which respects of similarity are of interest to the conversational participants. The predicate might apply to the pair <George Bush, Will Ferrell> in a context in which we are primarily concerned with physical appearance, but not in a context in which we are primarily concerned with personality. Moreover, even if we fix on a definite respect of similarity, “similar” will still exhibit the sort of vagueness which is characteristic of other gradable adjectives. Just as it is a vague (and contextually variable) matter how tall something must be in order to fall within the extension of “tall,” or how sharp something must be to fall within the extension of “sharp,” it is a vague (and contextually variable) matter how similar a pair of

things must be (in the relevant respect of similarity) in order to fall within the extension of “similar.” If we are to use the predicate “similar” for serious theoretical purposes, it will have to be semantically sanitized. This will require assigning the predicate a more stable and precise meaning, which, in turn, will require making a decision about which respects of similarity matter, and how similar things must be in the relevant respects in order to fall within the predicate’s extension. Moreover, if the resulting precisification of “similar” is to be theoretically interesting, it should not be an arbitrary or unprincipled precisification.

Two natural ways of satisfying these desiderata suggest themselves. The first is to give “similar” an extremely strong interpretation; the second is to give it an extremely weak interpretation. On the strong interpretation, “similar” expresses what we might call *strong similarity*, where two things are strongly similar just in case they are perfectly similar in every respect, i.e. perfect qualitative duplicates. On the weak interpretation, “similar” expresses what we will call *weak similarity*, where two things are weakly similar just in case they are at all similar in any respect. Intuitively, strong similarity and weak similarity are, respectively, the logically strongest and logically weakest relations which have a claim to be called similarity relations. The strong interpretation is clearly ill-suited for our purposes, for two things can share a color shade while failing to be perfect qualitative duplicates. Let us, then, adopt the weak interpretation. Hence, our official definition of “similarity-grounding property” is this:

F is a similarity-grounding property =_{df} F is a shareable property such that, necessarily, if any two things share F, they are weakly similar in virtue of the fact that each has F.

Hereafter, unqualified uses of “similar” are to be understood as expressing weak similarity.

It is worth pausing to address a worry about this interpretation of similarity that has doubtless occurred to the reader. It might be thought that this weak notion of similarity, though principled and reasonably precise, is theoretically useless, since every pair of things will be similar in this sense. Take two things as intuitively dissimilar as you like—say, an ant and a nebula. They will nonetheless be similar in the sense at issue, in virtue of their sharing such properties as *being a physical object*, *being spatially extended*, and so forth. But this is not a problem for our purposes, for we are not primarily concerned with the question of which pairs of things are similar. Rather, given two things which are similar, we are concerned with the question of which of their shared properties *make* them similar. It may be true—even necessarily true—that everything is similar to everything else, perhaps because some (necessarily) universally instantiated property such as *existence* is a similarity-grounding property. But it does not follow that every property shared by a given pair of objects is such that those objects are similar in virtue of sharing that property. Here it is important to recall the difference between metaphysical grounding and mere metaphysical necessitation. Supposing that it is a necessary truth that everything is similar to everything else, then it will follow that, for any property F, the sharing of F by x and y metaphysically necessitates the similarity of x

and *y*. But as we saw in §3.5.2, this does not entail that *x* and *y* are similar in virtue of sharing *F*.

3.5.4 An Aside on Naturalness

David Lewis's characterization of *natural* properties in some ways closely resembles my characterization of similarity-grounding properties. Natural properties, Lewis tells us, are those whose "sharing makes for resemblance" (1983, p. 347). But while the two notions are related, it will be helpful to point out some differences between them in order to better understand the notion of a similarity-grounding property, as well as to explain why I have so far avoided talk of naturalness.

The first difference lies in the breadth of their associated theoretical roles. The theoretical role of naturalness has connections to laws of nature, causation, metasemantics, duplication, and much else besides. I suspect that similarity-grounding properties have some of the features that Lewis ascribes to natural properties. They may be more eligible to figure in the contents of our utterances and mental states, for example, and they may be more eligible to figure in the laws of nature. But because I have no need of such strong and controversial assumptions for my purposes, it is useful to introduce a new concept which does not carry the theoretical baggage carried by the concept of naturalness. Secondly, Lewis and I mean different things when we say of a property that sharing it makes for similarity or resemblance. In particular, Lewis does not understand "makes for," as I do, in terms of metaphysical grounding. (The preference for claims of metaphysical necessitation and supervenience over claims of metaphysical grounding is one of the hallmarks of Lewisian metaphysics (Lewis 1983, p. 58; 2001, p. 384).) A third

difference is that Lewis typically treats naturalness as a gradable notion. But similarity-grounding is meant to be an all-or-nothing affair. Lewis does have a non-gradable notion in the vicinity: perfect naturalness. But the perfectly natural properties are generally taken to be a very sparse lot, perhaps comprising only fundamental microphysical properties (in addition to some alien properties). But it seems plausible that the similarity-grounding properties are fairly abundant. Personhood is, plausibly, a similarity-grounding property, though it will certainly not figure in the physicist's final catalogue of basic properties. Lewisian naturalness is therefore not an adequate surrogate for similarity-grounding, because there is no notion in the vicinity of naturalness in Lewis' philosophy which is both non-gradable and reasonably abundant.

3.5.5 Color Shades are Similarity-Grounding Properties

I hope that the notion of a similarity-grounding property is now tolerably clear. Let us now proceed to consider premise SG2, according to which color shades are similarity-grounding properties. Imagine you are looking at a uniformly colored red chip in good lighting conditions. You visually experience the chip as having a certain determinate shade of red—call it R. Focusing on R, it is conceivable that the object before you which appears to have R does not really have this quality. It is perhaps even conceivable that nothing actually has R. But what seems utterly inconceivable is that two things could share R without being similar thereby, without being similar in virtue of that very fact. It is plainly evident to us in an ordinary experience of R that, whatever else is true about R, any pair of things which have R are made similar by the sharing of R. That is, R is a similarity-grounding property. And the same can be said of other color shades as well.

3.5.6 Disjunctive Properties do not Ground Similarity

I have claimed that color shades are similarity-grounding properties. To this it might be replied: “So what? *Every* property is a similarity-grounding property. After all, what does similarity consist in if not the sharing of properties? And so long as we are concerned with weak similarity, shouldn’t the sharing of *any* property be sufficient to ground similarity?” However, it is not plausible to suppose that every property is a similarity-grounding property, even when weak similarity is at issue. In particular, it seems that *disjunctive* properties fail to ground similarity. As Armstrong (1978) writes,

Suppose a has a property P but lacks Q, while b has Q but lacks P. It seems laughable to conclude from these premises that a and b are identical in some respect. Yet both have the “property” P or Q.²⁴ (p. 20)

Suppose A is a aardvark and B is a bread-box. It would be odd—indeed, laughable—to say that A and B are similar in virtue of the fact that each is either an aardvark or a bread-box. Of course, I do not deny that A and B are similar. They are. They are similar in virtue of many of their shared features, e.g. in virtue of the fact that each is spatially extended, that each is made of matter, that each is a denizen of Earth, and so forth. Nor do I deny that it is a necessary truth that any two things which share the property of being either an aardvark or a bread-box are similar. (Recall the distinction between

²⁴ In this passage, Armstrong is giving an argument against the *existence* of disjunctive properties or universals. The argument rests on the assumption that *bona fide* properties are, in my terminology, similarity-grounding properties. *Contra* Armstrong, I am assuming that disjunctive properties exist, though I deny that they have a certain important feature which Armstrong takes to be a condition on property existence. But the assumption that disjunctive properties exist is merely a convenience. See fn. 26 below for a brief indication of how to reformulate the present argument in terms which are friendly to proponents of a sparse or ultra-sparse (nominalist) theory of properties.

metaphysical grounding and mere metaphysical necessitation.) But it is evident that things which share this property are not similar because each is an aardvark or a bread-box. The reason why things which share the disjunctive property are similar is that it's impossible to be either an aardvark or a bread-box without having some other properties, such as spatial extension, which themselves are similarity-grounding properties.

From this example we ought to conclude that at least some disjunctive properties fail to ground similarity. But of course, there was nothing very special about the example chosen. The same points could be made about just about any disjunctive property which springs to mind, e.g. being angry or a chandelier, being whimsical or poisonous, or being a student or a trash bin. More to present purposes: the same points could be made about those disjunctive physical properties with which the physicalist might identify color shades. Suppose o_1 and o_2 are a metameric pair, each of which appears to you to have a determinate color shade R , but which have, respectively, drastically different surface spectral reflectances P_1 and P_2 . Suppose for simplicity that the color physicalist identifies R with the disjunctive property $P_1 \vee P_2$. Just as it is not plausible to say that A and B are similar in virtue of the fact that each is either an aardvark or a bread box, it is equally implausible to say that o_1 and o_2 are similar in virtue of the fact that each has either P_1 or P_2 .²⁵ Once again, this is not to deny that o_1 and o_2 are similar. They are. They are similar in virtue of the fact that each reflects light, each produces experiences of such-and-such

²⁵ Similar points could of course be made about the kinds of disjunctive properties with which a categorical physicalist might identify the colors.

type in normal humans, each is a physical object, and so forth. But the fact that each has either P_1 or P_2 is no part of what makes them similar.

The above considerations strongly suggest the following generalization, expressed in premise SG3 of the argument above: *No* disjunctive property is a similarity-grounding property. In what follows, I address some objections to SG3 and attempt to clarify what is meant by a “disjunctive property.”²⁶

3.5.7 Structured Properties and Disjunctive Properties

In order to clarify what is meant by a “disjunctive property,” it will be helpful to begin by considering an objection which contends that every property is trivially a disjunctive property. (Why is this an objection? Because together with SG3, this would entail that no property is a similarity-grounding property, which contradicts SG2.) The objection begins by assuming a coarse-grained conception of properties, according to which necessarily coextensive properties are identical. Now consider an arbitrary property H . Take any other property F , and form the disjunction of the conjunction of F and H with the conjunction of F 's complement and H . So, if H is the property of being a horse and F is the property of being fast, this yields the property of being a fast horse or a non-fast

²⁶ In speaking of “disjunctive properties,” I am presupposing a view of properties as both real and abundant—abundant enough that most English predicates, including complex disjunctive predicates, correspond to properties. This background assumption, though a significant convenience, is ultimately inessential. We could alternatively elucidate SG3 in terms friendlier to a sparse or ultra-sparse (nominalist) view of properties. In these terms, SG3 could be recast as a claim about disjunctive predicates, namely: For any disjunctive predicate ϕ , there are no true metaphysical explanations of the form “ α and β are similar in virtue of the fact that $\phi(\alpha)$ and $\phi(\beta)$.” If SG3 is reformulated in this way, then we would need to reformulate premises SG1 and SG2 in a corresponding fashion to preserve the validity of the argument. This in turn would require giving nominalist-friendly redefinitions of color physicalism, color disjunctivism, and similarity-grounding properties, each of which was earlier defined in abundant realist terms. I will leave it as an exercise for those readers who have a taste for the tedious business of nominalist paraphrasing to convince themselves that such reformulations can indeed be given.

horse. The resulting property is a disjunctive property, in that it is the disjunction of some properties (i.e. $\lambda x(Hx \wedge Fx)$ and $\lambda x(Hx \wedge \neg Fx)$). Moreover, the resulting property is necessarily coextensive with H, and therefore identical with H given the coarse-grained conception of properties. So H is a disjunctive property. As H was arbitrary, it follows that every property is a disjunctive property.

My response to this objection is that we must reject the coarse-grained conception of properties. The coarse-grained conception of properties may be well suited for some theoretical purposes. But it is not suited to play the theoretical role associated with at least one good sense of “property.” Though there are many aspects of this role which coarse-grained properties are ill-suited to play, I will focus on the one most relevant for our purposes: namely, coarse-grained properties are ill-suited to play the theoretical role of properties *vis-à-vis explanation*. Explanations are hyperintensional in a way that requires a hyperintensional conception of properties. In typical explanations, we explain why things are the way they are by citing the properties of things. As has been observed by others, there seem to be cases in which necessarily coextensive properties differ in their explanatory significance. Elliot Sober (1982), for example, convincingly argues that there are cases in which we can truly explain why something happened by citing the trilaterality of an object but not by citing its triangularity, even though trilaterality and triangularity are necessarily coextensive. To give a somewhat simpler and more mundane example: in explaining why Joe left the party, we might cite a certain property of Joe, such as his tiredness. But given the truth of

E1: Joe left the party because he was tired,

it does not follow that we can explain why Joe left the party by citing his being tired-and-metallic-or-tired-and-non-metallic, or by citing his being tired-or-non-self-identical. For, barring some inconceivably bizarre circumstance, it will simply not be true that

E2: Joe left the party because he was tired-and-metallic-or-tired-and-non-metallic,
or that

E3: Joe left the party because he was tired-or-non-self-identical.

But the properties cited in the explanans of E2 and E3 are necessarily coextensive with the property attributed to Joe in the explanans of the true explanation E1. The point is not limited to causal explanations like E1-E3. The same holds for non-causal explanations, such as metaphysical explanations of similarity. It may be true that Dasher and Splasher are similar because—similar in virtue of the fact that—each is a horse. But it does not follow that—and it is not true that—they are similar because each is either a fast horse or a non-fast horse. (Recall the discussion in §3.5.2 about the hyperintensionality of metaphysical grounding.) So on any conception of properties which enable properties to play their intended theoretical role vis-à-vis explanation, it will turn out that being a horse is not identical to the disjunctive property being a fast horse or a non-fast horse.

The most obvious conception of properties fit to play this theoretical role is a conception of properties as *structured* entities. The relation between coarse-grained properties and structured properties is very much analogous to the relation between possible-worlds propositions—sets of worlds or their characteristic functions—and structured propositions (Salmon 1986, Soames 1987). On a structured view of propositions, propositions have a quasi-sentential structure, with propositional

constituents occupying nodes in this structure. Likewise, structured properties can be thought of as having quasi-predicational structure; that is, their structure is analogous to the syntactic structure of complex predicates.²⁷ This structured conception of properties has two important consequences. First, it allows us to individuate properties more finely than does the coarse-grained conception of properties—i.e. more finely than sets of possibilia. Vagueness aside, any structured property will determine a unique set of possibilia,²⁸ i.e. the set of possible objects which instantiate the property, just as any structured proposition will, vagueness aside, determine a unique set of worlds, i.e. the set of worlds at which the proposition is true. But sets of possibilia do not determine a unique structured property, just as sets of worlds do not determine a unique structured proposition. More precisely, if some structured property F is instantiated by exactly the possibilia in S , then there will be (infinitely many) other structured properties instantiated by exactly the possibilia in S (e.g. $\lambda x(F x \wedge \text{Exists}(x))$, $\lambda x((F x \wedge \text{Exists}(x)) \wedge \text{Exists}(x))$, ...). The second important consequence of the structured conception of properties is that it allows us to make a meaningful distinction between disjunctive and non-disjunctive properties. Recall that structured properties are supposed to be structurally analogous to complex predicates. The question of whether a complex predicate is disjunctive is straightforwardly settled by whether or not it has a disjunctive

²⁷ I will leave unsettled the question of just how tight the analogy is. We probably ought to say that structured properties are somewhat less finely structured than complex predicates. For example, while “ $\lambda x(\text{Dog}(x) \vee \text{Brown}(x))$ ” is a different predicate from “ $\lambda x(\text{Brown}(x) \vee \text{Dog}(x))$,” it would be somewhat odd to maintain that they express distinct properties.

²⁸ To handle complications that arise if we allow for transworld identity, I really ought to say “a unique function from worlds (or perhaps world-time pairs) to sets of individuals in those worlds (or at those world-time pairs). But I will ignore these complications in what follows.

structure (i.e. whether or not the main connective of the largest formula within the scope of the λ -operator is a disjunction symbol). Likewise, the question of whether a structured proposition or property is disjunctive is straightforwardly settled by whether or not it has a disjunctive structure, where having a disjunctive structure will be a matter of whether an appropriate constituent (perhaps the disjunction truth function?) occupies the relevant position within the structure of the property. Given this independently motivated structured conception properties, the objection above fails. Not every property is a disjunctive property, even if every property is necessarily coextensive with some disjunctive property.

3.5.8 Extrinsic Similarity Grounding?

The next objection I would like to consider suggests that a disjunctive property can ground similarity in cases where its disjuncts have some common extrinsic causal influence, e.g. on human perceptual systems. Suppose that the color physicalist identifies some color shade C with the disjunctive physical property $D = P_1 \vee P_2 \vee \dots$. Let us furthermore assume that the P_i out of which D is constructed are all and only the physical properties which cause such-and-such state in normal human perceivers under normal conditions. Everything that has D will therefore have a similar influence on human perceivers. It might be suggested that, due to this contingent and extrinsic similarity among the things which share D , we ought to say that D is a similarity-grounding property, albeit contingently so.

But this suggestion is confused. Any pair of things which share D will also share the property of normally causing perceptual states of such-and-such kind. The latter

property may well be a similarity-grounding property. (Not all similarity-grounding properties are intrinsic.) It follows from this that any pair of things which share D will be similar. But from the fact that D is coextensive with a similarity-grounding property, we cannot conclude that it is itself a similarity-grounding property. In a world where the only spatially extended things are aardvarks and bread boxes, being an aardvark or a bread box will be coextensive with the similarity-grounding property of being spatially extended, but that will not make the former a similarity-grounding property.

Moreover, it follows from the definition of a similarity-grounding property that there cannot be properties which only *contingently* ground similarity. To say of some shareable property F that it is a similarity grounding property is to say that:

SG. Necessarily, if any two things share F then they are similar in virtue of the fact that each has F.

To claim that this is a contingent fact about F, one must hold that SG is not necessary, and this is to deny the validity of the inference from $\Box\phi$ to $\Box\Box\phi$. But this inference is valid in S5,²⁹ which is standardly taken to be the correct logic for metaphysical modality.

3.5.9 Potential Counterexamples to SG3

The next objection I want to consider maintains that while disjunctive properties are typically not similarity-grounding properties, in special cases they can be. The strongest reason to think this comes from determinable properties. Determinable properties like having mass or being red seem to be similarity-grounding properties. Any two things

²⁹ More generally, it is valid in any modal logic which can be modeled using a transitive accessibility relation.

which have mass are similar in virtue of the fact that each has mass, even if they differ in their determinate mass. Likewise, any two red things are similar in virtue of the fact that each is red, even if they differ in their determinate shade of red. But arguably, determinable properties are disjunctive properties—specifically, disjunctions of all their absolute determinates. If so, then SG3 is false; at least *some* disjunctive properties are similarity-grounding properties. And if some disjunctive properties are similarity-grounding properties, this at least opens up the possibility that the disjunctive physical properties with which color physicalists identify color shades are among them.

There are three points to make in response to this suggestion: first, much of the motivation for identifying determinables with disjunctions of their absolute determinates evaporates once we reject the coarse-grained conception of properties. For, while it may be that determinable properties are necessarily coextensive with disjunctions of their absolute determinates, this does not entail that determinables are identical with such disjunctions unless we assume a coarse-grained view of properties. Second, there is at least one good positive reason to think that determinable properties are not identical to disjunctions of their absolute determinates (besides the dialectically unacceptable reason that determinables, unlike disjunctive properties, are similarity-grounding properties). In particular, it seems that many determinable properties are not even necessarily coextensive with the disjunction of their absolute determinates. (Necessary coextensiveness may not be a sufficient condition for property identity, but it follows from Leibniz's Law that it is a necessary condition.) Consider a dappled red cloth which is composed of a collection of small-ish patches, each of which has a determinate shade

of red, with no two patches sharing the same shade of red (cf. Liebesman 2011, p. 435). The cloth itself has the determinable property redness. But while we can attribute determinate shades of red to many of its small-ish parts, there seems to be no determinate shade of red we can correctly attribute to the cloth itself.³⁰

The third and most important point to make in response to this objection is this: let us grant for the sake of argument that determinable properties are identical to disjunctions of their absolute determinates and that, therefore, some disjunctive properties ground similarity. Even so, this concession does not help the color physicalist. That's because even if determinable properties are a special kind of disjunctive property, we obviously cannot construct a genuine determinable by disjoining just any collection of properties. *Bona fide* determinables like having mass or being colored are importantly different from the result of disjoining just any motley assortment of properties, e.g. the result of disjoining red₁₇, having a mass of 2.531 kg, and being 15.52 inches long. Disjunctive properties of the latter kind, like the property of being an aardvark or a bread box, do not ground similarity. So if one identifies determinables with disjunctions of their absolute determinates, and therefore accepts that some disjunctive properties ground similarity, one must answer the question: what distinguishes similarity-grounding disjunctive properties from other disjunctive properties—those which are *merely* disjunctive? It seems to me that any plausible answer to the above question must say something in the ballpark of the following:

³⁰ This would be a counterexample to the Principle of Determination mentioned in §2.4.2. But it is clearly not the sort of counterexample that could help my opponent in that context.

Those disjunctive properties which ground similarity are precisely those which are formed by disjoining some collection of properties that exhibit (what we may call) “internal unity.” A collection of properties is internally unified in the relevant sense just in case, very roughly, they jointly constitute a topologically connected sub-region of some natural quality space. (In the case of very general determinables, such as mass or color, the relevant region may be an *improper* subregion of the quality space.) A quality space consists of a large number of determinate properties (the points of the space), which can be naturally ordered along one or more dimension(s) of similarity so as to induce something akin to metrical structure on the space, a structure which then allows for an intelligible distinction between connected and non-connected regions of the space. This proposal delivers the intuitively correct result on the cases. It delivers the intuitively correct result that the determinable height property *being 7 to 9 feet tall* (which is formed by disjoining the properties in a certain connected region of “height space,” i.e. the set of determinate height-properties in the 7’-9’ range), as well as the determinable red (which is formed by disjoining the properties in a certain connected region of color space, i.e. the set of determinate shades of red), will count as similarity-grounding. It also delivers the intuitively correct result that the disjunctive property *being 5’3” or 6’2” or 7’4”* (which is formed by disjoining the properties in a non-connected region of height space), as well as the disjunctive property being red₁₇ or yellow₂₁ or green₁₁ (which is formed by disjoining the properties in a non-connected region of color space), will not count as similarity-grounding properties.

Let us say that a disjunctive property is an “internally unified disjunction” just in case it is formed by disjoining the properties in some internally unified collection, in the sense defined above. And let us say that a disjunctive property is a “motley disjunction” just in case it is not internally unified. I agree with the above proposal that the only disjunctive properties which could plausibly ground similarity are internally unified disjunctions. It is not plausible that motley disjunctions ground similarity. The problem for the color physicalist is that the only physical properties which could reasonably be identified with the color shades are motley disjunctions. This is very clear in the case of categorical physicalism. For any given color shade C, there seems to be no natural internal unity at all to the collection of categorical properties which are causally involved in the production, in normal circumstances, of C-experiences. And it is also true for the light-dispositionalist. The reflectance curves which are metamers for a given color shade (relative to some normal illuminant) are a motley assortment. As light-dispositional physicalists are wont to admit, the disjunctive reflectance properties which are identified with color shades are *anthropocentric* properties (Hilbert 1987, Byrne and Hilbert 2007a, Tye 2000): If we set aside their common influence on human perceivers, there is nothing about the disjuncts involved in these properties which recommends classing them together. The only disjunctive physical properties which could reasonably be identified with the color shades are motley disjunctions. Therefore, none of the disjunctive physical properties which could reasonably be identified with color shades are similarity-grounding properties. But color shades *are* similarity-grounding properties. So we ought to conclude that color physicalism is false.

Chapter 4: Realism

4.1 Introduction

In the previous two chapters, I addressed the question of reductionism: are the sensible qualities reducible to physical or response-dependent properties? I argued for the non-reductionist conclusion that at least many of the sensible qualities, including many of the so-called “secondary qualities,” are not reducible to physical or response-dependent properties of external objects. This raises the question: do they belong to external objects at all? This is the question of *realism*. Let realism about a class of sensible qualities be the thesis that qualities in that class are instantiated by ordinary physical things. “Ordinary things” include things like roses, ice cubes, strawberries, pressure waves, and odor clouds —roughly, those things that scientifically informed common sense takes to belong to the extra-mental world. Realism about a given class of sensible qualities therefore rules out *eliminativist* views (Hardin 1988, Maund 1995, Chalmers 2006, Pautz 2011, ms.), according to which nothing at all has the relevant qualities, as well as *projectivist* views, according to which the only things that possess the relevant qualities are non-ordinary things like sense-data, sensory experiences, or patches of the visual field (Boghossian & Velleman 1989, Jackson 1977). I shall use the more general term “irrealism” to designate the negation of realism.

In this chapter, I offer two positive reasons to accept realism about every major class of sensible qualities. The first is that the various sensible qualities *seem* to be instantiated in our environment (and in our bodies). Here I rely on (a weakened version

of) what Michael Huemer (2001) calls the “Phenomenal Conservative” principle: if it perceptually seems to one that *p*, then one has at least *prima facie* justification to believe that *p*. The second positive reason to accept realism comes from the independently motivated thesis of phenomenal relationism. Roughly speaking, phenomenal relationism holds that what it is to undergo a (sensory) experience with a given phenomenal character is to stand in a certain relation, which might variously be called “sensory awareness,” “sensory acquaintance,” or “sensory representation,” to a certain sensible quality or complex of sensible qualities. In §4.9, I shall argue that, given phenomenal relationism, the assumption of realism about the sensible qualities makes the “hard problem” of consciousness much more tractable. Realism is therefore supported on abductive grounds.

Many philosophers have held that the rejection of reductionism leads—or should lead—to irrealism. I disagree. We have strong reasons to accept realism, and these reasons aren’t much vitiated by the failure of reductionism. But the possibility of a non-reductive realism has often been ignored or dismissed by philosophers at the outset. Many philosophers just take it as given that the scientific image, perhaps supplemented with primitive experiential facts, constitutes an exhaustive characterization of reality. So if the secondary qualities can’t be found in the scientific image (augmented with phenomenology), they must have no place in the world at all. The following remarks from Boghossian and Velleman (1991) are typical:

The dispute between realists about color and [ir]realists is actually a dispute about the nature of color properties. The disputants do not disagree over what material objects are like. Rather, they disagree over whether any of the uncontroversial

facts about material objects—their powers to cause visual experiences, their dispositions to reflect incident light, their atomic makeup, and so on—amount to their having colors. The disagreement is thus about which properties colors are and, in particular, whether colors are any of the properties in a particular set that is acknowledged on both sides to exhaust the properties of material objects (p. 67).

This may be an accurate characterization of the dispute between irrealists and *reductionist* realists about color. There is surely an intuitive sense in which the reductionist realist, whether physicalist or relationist, agrees with the irrealist about what material objects are like. But however accurate the above remarks are concerning the dispute between irrealists and reductionist realists, they certainly do not apply to *my* dispute with the irrealist. It's often hard to say exactly when we have a first-order disagreement about "what material objects are like" as opposed to just a second-order disagreement about the nature of a certain class of properties. But my dispute with the irrealist is a clear case of the former. We may well agree on all relevant second-order claims about the natures of such qualities as color, flavor, pitch, and so on. And we agree that none of "the uncontroversial facts about material objects—their powers to cause visual experiences, their dispositions to reflect incident light, their atomic makeup, and so on—amount to their having colors." But we do not agree about whether such properties exhaust the properties of material objects. (Such properties therefore do *not* constitute "a particular set that is acknowledged on both sides to exhaust the properties of material objects," because there is no such set.) We emphatically do *not* agree about what material objects are like.

4.2 The Empirical Justification for Realism

Since the early 17th century, it has been widely held that scientific or empirical investigation has shown that the so-called secondary qualities are not instantiated in the material world. Galileo writes,

Hence I think that these tastes, odors, colors, etc., on the side of the object in which they seem to exist, are nothing else than mere names, but hold their residence solely in the sensitive body; so that if the animal were removed, every such quality would be abolished and annihilated (Galilei 1623/1954).

The Galilean view is alive and well among philosophers today, even if somewhat less popular than it was in the early modern period. And among scientists, the view that empirical investigation supports irrealism is as popular as ever.

However common this view has been over the past few centuries among scientists and philosophers, on its face it looks entirely backwards. Everyday observation provides abundant *prima facie* empirical justification to believe, for example, that grass is green, that female opera singers sometimes produce loud, high-pitched sounds, that ripe strawberries are sweet, and a great many other propositions that straightforwardly entail realism about the major classes of secondary qualities. While it may be that *scientific* investigation, understood as a kind of controlled, systematic, and rigorous form of empirical investigation, has not provided *further* support for realism about, say, color or flavor, this is only to be expected, given the nature of realist claims. Claims of realism concerning color, flavor, scent, pitch, loudness, and other secondary qualities are what I'll call *mundane* propositions. A mundane proposition is any proposition that is abundantly

supported by empirical observations of the sort that ordinary humans are likely to make in their day-to-day lives. That is, they are propositions whose empirical support comes from mundane or casual observation, as opposed to the careful and painstaking observational methods employed in scientific inquiry. Examples of mundane propositions include: that there are cups; that dogs bark; that it sometimes rains. Anyone who leads a reasonably normal life will have had ample exposure to cups, barking dogs, and rainy weather, and so will have amassed plenty of empirical justification to believe each of these propositions. But while these claims are overwhelmingly supported by empirical observation, *scientific* investigation has not significantly added to our empirical justification for accepting these claims, nor could it. This is because controlled, systematic, and rigorous observation can add little support to what is already overwhelmingly confirmed by casual everyday observation.

One characteristic feature of mundane propositions is that, while it is easy to verify them by observation, it is difficult to *argue* for their truth in any direct way. Consider how you might try to convince a skeptic about cups of the mundane truth that there are cups. One possibility is that your interlocutor has led a very abnormal life, and has simply never encountered a cup. In this case, the way to proceed is not to argue the point; rather, you ought to simply present your interlocutor with a cup; if her perceptual faculties are in working order and she is competent with the relevant concepts, that ought to settle the matter. Another possibility is that, despite having had, like you, ample exposure to cups, your interlocutor denies the existence of cups anyway, perhaps because she has been convinced by the writings of mereological nihilists that there are no

mereologically complex entities (Dorr 2002; Sider 2013). How can we respond to this sort of interlocutor? Within debates about mereological nihilism, we find a common maneuver. The anti-nihilist, unable to provide a direct defense of the mundane truths denied by the nihilist, gives something like the following speech:

Common sense tells us that there are cups. We ought to believe the dictates of common sense unless there are very strong reasons to believe otherwise, and the nihilist has not provided such reasons. So, we ought to believe that there are cups.

In giving this speech, the anti-nihilist resorts to a kind of “epistemic ascent.” Finding herself at a loss about how to argue directly for the mundane proposition that there are cups, she instead motivates an epistemic principle from which she can argue that the proposition in question has some favorable epistemic status—in this case, the status of being such that we ought to believe it. I find myself in a similar situation with respect to realism about the sensible qualities. There isn’t much³¹ I can say to argue directly for the mundane proposition that some things in the external world have color or flavor or loudness or pitch. So I will defend realism the best way I know how—indirectly, by way of epistemic ascent.

4.3 The Perceptual Argument

Let us first lay out a general argument *template*, beginning with the following epistemic principle:

³¹ But there is something; see §4.9.

T1: If it perceptually seems to one that p, then one has at least *prima facie* justification to believe that p.

In chapter 2, I appealed to the Phenomenal Conservative principle: if it seems to one that p, then one has at least *prima facie* justification to believe that p. T1 is a weaker version of the Phenomenal Conservative principle, one restricted to *perceptual* seemings. (It is therefore noncommittal about the epistemic significance of other kinds of seemings, such as intellectual or intuitive seemings.) From here, the argument template proceeds as follows:

T2: It perceptually seems to us that ____.

T3: There are no (sufficiently strong) defeaters for the belief that ____.

T4: So, we are justified in believing that ____.

For each sensible quality class Q, we can plug in a suitable formulation of realism about Q for “____,” yielding an argument to the conclusion that we are justified in believing (the relevant formulation of) realism about Q. In the case of color, for example, we might fill in the template in the following way:

PA1: If it perceptually seems to one that p, then one has at least *prima facie* justification to believe that p.

PA2: It perceptually seems to us that there are colored things.

PA3: There are no (sufficiently strong) defeaters for the belief that there are colored things.

PA4: So, we are justified in believing that there are colored things.³²

For the sake of concreteness, I devote much of what follows (§§4.4-4.6) to defending this particular instance of the perceptual argument template, rather than trying to defend all instances of it at once. But as I shall argue in §§4.7-4.8, the defense generalizes with only a few minor complications to other instances of the template. I do not expect the second premise to meet resistance from anyone. It is clear, for example, that lemons, tomatoes, and grass at least *appear* to be colored. Similar claims hold for the other secondary qualities as well: we experience ice cubes (or perhaps the body parts that make contact with them) as cold, honey as sweet, the chirps of blue jays as high pitched, and so on. For this reason, I focus exclusively on the first and third premises.

4.4 Premise PA1

Premise PA1 is intended to articulate a fairly weak anti-skeptical principle in the epistemology of perception. It is important to distinguish PA1 from another anti-skeptical principle which is more commonly invoked in defense of realism, a principle we might call *Mooreanism*:

Mooreanism: If it belongs to common sense that p, then one has prima facie justification to believe that p.

If we replace PA1 above with Mooreanism and replace PA2 above with “it belongs to common sense that there are colored things,” the result is what I’ll call the Moorean argument. While most realists invoke something like the Moorean argument in defense of

³² “Us” in PA2 and “we” in PA4 should be understood to refer to humans with normally functioning visual systems. The quantifier in PA2 and PA4 should be understood as restricted to ordinary material things in the external world.

their view,³³ the Moorean argument is significantly weaker than the perceptual argument, for at least two reasons. First, Mooreanism enjoys considerably less *prima facie* plausibility than PA1. The point is made forcefully by Sider (2013): “On the face of it, Mooreanism is utterly implausible. Why should the inherited prejudices of our forebears count for anything?” (p. 10). As Sider notes, even those philosophers who think that our philosophical theorizing ought to respect common sense feel the initial force of this challenge. By contrast, one cannot raise an analogous challenge to PA1 without thereby calling into question the whole of empirical science. Second, even if we set aside staunch anti-Moorean positions like Sider’s, there is a more moderate (and perhaps more reasonable) attitude toward the epistemic status of common sense which makes trouble for the appeal to Mooreanism in this context. According to this view—expressed in various forms by Lycan (2001b, p. 41), Gupta (2006, p. 178), and Kelly (2008)—that a proposition belongs to common sense may make a significant epistemic difference when doing armchair philosophy; but history shows us that science is much less limited in its capacity to overturn common sense. Witness, for example, the discovery that simultaneity is relative to an inertial frame, that my cat and I are distant cousins, or that the earth is currently spinning on its axis at about 1000 miles per hour and hurtling through space around the sun at a rate of 18.5 miles per second. So in the context of scientific inquiry or scientifically informed philosophy, common sense counts for significantly less than it does in armchair philosophy. But irrealists typically take their position to be supported by empirical science, not armchair philosophical arguments. So the appeal to common sense

³³ See, e.g., Tye (2000; ch. 7), Lewis (1997, p. 325), Cohen (2009, p. 65).

in this context carries significantly less weight than it does in other philosophical contexts which are further removed from scientific inquiry.

The appeal to PA1 is dialectically more effective against the irrealist than the appeal to Mooreanism. The irrealist may reasonably be skeptical about the deliverances of common sense. But she cannot be radically skeptical about the deliverances of perception without undermining all of empirical science, and with it the standard arguments against realism. So we can safely assume in this context that any theory in the epistemology of perception which will be acceptable to my opponent must be compatible with *empirical anti-skepticism*—the thesis that we are justified in accepting many of the deliverances of empirical science. (Here I understand a “theory” in the epistemology of perception as any collection of general claims characterizing the epistemic significance of perceptual experience.) In what follows, I argue that any acceptable theory in the epistemology of perception that is compatible with empirical anti-skepticism will endorse PA1, or at least something near enough to PA1 to serve my argumentative purposes. In particular, I shall argue that, given reasonable assumptions, PA1 (or something near enough) follows from the dominant internalist and externalist approaches to the epistemology of perception (§§4.4.1-4.4.2); I then consider and respond to an objection to PA1 (§4.4.3).

4.4.1 Internalist Theories of Justification: Dogmatism and Rationalism

Among broadly “internalist” theories in the epistemology of perception, the two most popular theories are Dogmatism and Rationalism.³⁴ Roughly speaking, I take internalism to be the thesis that facts about what one has justification to believe supervene on one’s (past and present) non-factive mental states. Externalism can then be understood as the denial of internalism. PA1 follows more-or-less straightforwardly from both Dogmatism and Rationalism. Dogmatists and Rationalists will both agree that, if it perceptually seems to one that p, then one has *prima facie* justification to believe that p. Dogmatists add to PA1 the claim that this justification does not rest on any antecedent justification to believe anything else. In particular, it does not rest on any antecedent justification to believe that our perceptual faculties are reliable. Rationalists, by contrast, add to PA1 the claim that one’s (*prima facie*) justification to believe that p (when it perceptually seems to one that p) rests on one’s antecedent justification to believe that one’s perceptual faculties are reliable, but also add that one has the latter justification “by default” or *a priori*. My own sympathies lie with Rationalist theories in the epistemology of perception, but the appeal of PA1 is not restricted to those with Rationalist sympathies.

Note that because PA1 is logically weaker than both Dogmatism and Rationalism, an objection to one of the latter may not be an objection to PA1. For example, the most famous objections to Dogmatism, the well known bootstrapping and Bayesian objections (White 2006), are best construed not as objections to (1) but rather as objections to what Dogmatism *adds* to (1). There are other theories in the epistemology of perception

³⁴ For a classic defense of Dogmatism, see Pryor (2000). For defenses of Rationalism, see Wright (2004) and White (2006).

(including Rationalism) to which these objections don't apply, despite their endorsement of PA1.

4.4.2 Externalist Theories of Perceptual Justification: Process Reliabilism

Among broadly “externalist” theories of justification, perhaps the dominant approach is that of Process Reliabilism (hereafter just “Reliabilism”) (Goldman 1979). The basic thesis of Reliabilism may be formulated as follows:

Reliabilism: One's belief that *p* is *prima facie* justified if and only if it is produced by a reliable belief-forming process.³⁵

How does PA1 fare on a Reliabilist theory of epistemic justification? Two complications arise in trying to answer the question. First, Reliabilism is a thesis about *doxastic* justification, i.e. a thesis about the conditions under which a belief is (*prima facie*) justified, whereas PA1 is a claim about *propositional* justification, i.e. the conditions under which one has (*prima facie*) justification to believe a proposition, which one might have even if one does not in fact believe the proposition. This is not a major complication, however. It's plausible that positive facts about doxastic justification entail corresponding positive facts about propositional justification. That is, if one's belief that *p* is doxastically justified, it follows that one has propositional justification to believe that *p*. On the other hand, negative facts about doxastic justification do not entail corresponding

³⁵ Some formulations of the Reliabilist thesis omit the “*prima facie*” qualification. I include it here in part to simplify matters: PA1 is itself a claim about *prima facie* justification, and it is easier to assess Reliabilism's relevance to PA1 if we add this qualification. But the main reason for the addition is that the view is utterly implausible without it. For without the qualification, the thesis predicts that a belief which is formed by a reliable process is justified even if one possesses overwhelming defeating evidence against the belief.

negative facts about propositional justification. That is, if one's belief that *p* is *not* doxastically justified, it does not follow that one lacks propositional justification to believe that *p*. However, the absence of doxastic justification is often a good indicator of the absence of propositional justification. For our purposes, it is fair to assume (since it is dialectically to my disadvantage) that if the common belief that there are colored things lacks doxastic justification, then we also lack proposition justification to believe that there are colored things.

The second complication that arises in assessing the relevance of Reliabilism to PA1 is that, if Reliabilism is true, then the justificatory status of our perceptual beliefs depends to a great extent on how the external world actually is. If it turns out that we are in a skeptical scenario, then given Reliabilism, it will turn out that our perceptual beliefs are not *prima facie* justified. Because of this, it's tempting to think that there will be no non-question-begging route from Reliabilism to the truth of PA1. The temptation is motivated by the following line of thought: suppose that color realism is false. Then our perceptual beliefs about the colors of things in our environment, including the general belief that there are colored things, are uniformly false. It would seem, therefore, that each of them is the result of an unreliable process. So given Reliabilism together with the falsity of color realism, it follows our perceptual beliefs about the colors of things are not *prima facie* justified. It's then reasonable to infer that we lack propositional justification to believe their contents, from which it follows that PA1 is false. Hence any route from Reliabilism to PA1 will have to assume the truth of color realism. But this would clearly be question-begging in the current context. More generally, the basic problem is that

Reliabilism entails a tight connection between the justificatory status of our beliefs and the truth of our beliefs, and this makes it difficult to answer questions about the justificatory status of our perceptual beliefs without making assumptions about their truth. Thus, if a Reliabilist approach to justification is correct, then it would seem that my strategy of indirectly defending color realism by way of epistemic ascent is problematic. The whole reason for resorting to epistemic ascent was that there seems to be no more direct way of defending color realism. But if I cannot defend the claim that we are justified in believing color realism without assuming color realism, then it would seem that the indirect defense of color realism can't get off the ground.

Despite appearances, I believe that we *can* make the inference from Reliabilism to PA1 plausible without begging any questions against the irrealist. Recall that in the present dialectical context, any admissible view must meet the empirical anti-skeptical constraint. This plausibly entails that our perceptual beliefs about spatiotemporal matters are, in large part, justified. After all, if our perceptual beliefs about spatiotemporal matters are not justified, then it is hard to see how we could be justified in accepting the results of empirical science, much of which characterizes the world in terms of its spatiotemporal structure. Given Reliabilism, this means that the belief-forming process responsible for our perceptual beliefs about spatiotemporal matters is reliable. From this it follows, given Reliabilism, that any belief, whether it concerns spatiotemporal matters or not, produced by the belief-forming process responsible for our perceptual beliefs about spatiotemporal matters is itself (*prima facie*) justified.

Now consider some perceptual belief on whose truth value the realist and the irrealist disagree, e.g. my perceptual belief that my socks are blue. Without assuming the truth of color realism, it is plausible, given the weak anti-skeptical assumptions at play in the current context, that this belief is the product of a reliable belief-forming process. Note that when we assess whether a belief was the outcome of a reliable process, we need to make some choices about how to individuate the relevant processes. The token series of events leading to the formation of a given belief is a token of indefinitely many process types, some very general, some very specific. For any belief, there will be some way of individuating processes such that the belief in question counts as the outcome of a very reliable process, and there will be other ways of individuating processes such that the belief in question counts as the outcome of a very unreliable process. One of the most pressing challenges for Reliabilist theories of justification is the “generality problem”: which way of individuating process types is the one which is relevant to determining the justificatory status of our beliefs. If Reliabilism is true, then the generality problem has an answer. In other words, there is some privileged way of individuating process types—perhaps one which is maximally “natural” in the sense of Lewis (1983)—such that the reliability of those processes is what matters for justification. It’s plausible that on the privileged natural way of individuating belief-forming processes, my belief that my socks are blue is a result of the same process that produces my perceptual beliefs about spatiotemporal matters. Perhaps the relevant belief-formation process type is just the *perceptual* belief-formation process: the process which involves a kind of (non-deviant) causal transition from a perceptual state to a belief state with a suitably related content.

Or perhaps the relevant process type is something narrower, perhaps the *visual* belief formation process (or the close-up-in-good-lighting visual belief formation process): the process which involves a kind of (non-deviant) transition from a perceptual state of the visual system (of a nearby object in good lighting) to a belief state with a suitably related content. In either case, it will turn out that our beliefs about the colors of things, including our belief that there are colored things, are products of the same belief-forming process as most of our perceptual beliefs about spatiotemporal matters. It then follows from Reliabilism, given our anti-skeptical assumptions, that my belief that my socks are blue is *prima facie* justified as well.

Suppose, however, that given some well motivated answer to the generality problem, we cannot show that Reliabilism yields the truth of PA1 without assuming the truth of color realism. (This would be the case if, for example, we held that the relevant process type responsible for my belief that my socks are blue is the process that involves a (non-deviant) transition from a *color-representing* state of the visual system to a belief about the colors of things in my environment.) In other words, suppose that, given this version of Reliabilism together with the assumption of color irrealism, it turns out that PA1 is false. Even if this is so, there is good reason to think that this would stem from a defect in Reliabilism rather than a defect in PA1. Perhaps the best known problem with Reliabilism is that it seems to deliver the wrong results about justification in skeptical scenarios (Cohen 1984; Pollock 1984). Intuitively, my brain-in-vat twin, who has all the same experiences as me and forms all the same beliefs on the basis of experience, though systematically *mistaken* in his beliefs about the external world, is no less *justified* in

holding those beliefs than I am. But Reliabilism seems to entail that the beliefs of my envatted twin are not just systematically mistaken, but systematically unjustified. If Reliabilism, when combined with the assumption that color realism is false, delivers the falsity of PA1, we should suspect that this is due to a defect in Reliabilism, not a defect in PA1. For to suppose that color realism is false is to suppose that we are in a kind of skeptical scenario, a scenario in which perceptual experience is systematically mistaken about how the world is, and Reliabilism gives notoriously problematic results when applied to such skeptical scenarios.

4.4.3 Abductivism/Classical Foundationalism

To conclude my defense of PA1, I would like to consider an approach to the epistemology of perception that often seems to be tacitly presupposed by opponents of realism and which appears to present a challenge for PA1.

In objecting to realism about color (more precisely, “Edenic” color, understood as the primitive color qualities phenomenally presented in color experience), Chalmers (2006) remarks that “the hypothesis that objects have [Edenic colors] seems quite unnecessary in order to explain color perception” (p. 67). Remarks like this are commonly made by color irrealists, with the implicit suggestion that, because the supposition that objects are colored is unnecessary to explain our color experiences, perceptual experience therefore does not provide us with justification to believe that objects are colored. It is natural to suppose that behind this thought lies something like the following view, closely related to classical foundationalism, about our justification for holding beliefs about the external world:

Abductivism: If p is a proposition concerning the external world (i.e. a proposition whose truth requires the existence of an external world with a certain character), then one has justification to believe p only if p figures in the best explanation for one's perceptual experiences.

Abductivism appears to present a challenge for PA1. Assuming that our color experiences can be fully explained without the hypothesis that there are colored things, then one can appeal to Abductivism to argue that we lack (perhaps even *prima facie*) justification to believe that there are colored things. More generally, there may be many propositions p—including but not limited to the proposition that there are colored things—such that p perceptually seems to us to be the case, yet p does not figure in the best explanation for our sensory experiences. If this is correct, then given the truth of Abductivism, there will likely be many counterexamples to PA1.³⁶

I do not think that the Abductivist threat to PA1 is very serious, for two reasons. First, Abductivism itself is a highly suspect approach to the epistemology of perception. Second, the assumption of explanatory irrelevance—i.e. the assumption that the hypothesis that colors are instantiated in the external world does not figure in the best explanation of our sensory experiences—is doubtful. Let us take these in turn:

Against Abductivism: Abductivism, like its close cousin, classical foundationalism, presents a picture of epistemic justification on which our basic stock of evidence, from which all of our (*a posteriori*) beliefs derive their justification, consists in

³⁶ Why only “likely”? Because, as stated, Abductivism is a claim about (*ultima facie*) justification while PA1 is a claim about *prima facie* justification. Lack of (*ultima facie*) justification does not straightforwardly entail lack of *prima facie* justification.

facts about our occurrent phenomenology. Nowadays this phenomenal conception of evidence is largely repudiated by epistemologists, including contemporary foundationalists, who typically deny the classical foundationalist assumption that beliefs about the external world are absent from our stock of foundational beliefs.³⁷ Our basic evidence should be taken to include facts about the external world, including facts about the gross features of our immediate environment.

Abductivism, like classical foundationalism, is presumably motivated by the following, loosely Cartesian, line of thought:

Our ordinary beliefs about the external world are dubitable and corrigible. By contrast, our beliefs about our occurrent phenomenology are especially epistemically secure, perhaps even indubitable or incorrigible. So whatever justification attaches to our ordinary beliefs about the external world must derive from our beliefs about our occurrent phenomenology.

But there are two major problems with this line of thought. First, in general our beliefs about our occurrent phenomenology are not especially epistemically secure. I may have some very general beliefs about my occurrent phenomenology which are more certain than any of my beliefs about the external world, e.g. my belief that I am currently undergoing a conscious experience. But my beliefs about more specific features of my occurrent phenomenology, which is the only kind of phenomenological belief from which I could hope to derive anything interesting about the external world, are not especially secure. We are far more reliable in discerning the gross features of objects in our

³⁷ For an influential attack on the phenomenal conception of evidence, see Williamson (2000, ch. 4, 8).

immediate environment than we are in discerning the phenomenal features of our occurrent experiences. To illustrate: hold a quarter out in front of you and tilt the quarter so that it is at an angle with respect to your line of sight. Does the coin now look oval-shaped? Now stretch your arm out as far as you can, so as to move the coin further away from you. Does it look smaller now than it did before? The answer is not obvious. My own inclination is to say that there is no sense in which the coin looks oval-shaped when it is tilted or looks smaller when it is further away, but other philosophers disagree (Hume 1739/1978; Noë 2004). Now consider the corresponding non-phenomenological questions about the coin itself: when you tilt the coin, does the coin itself become oval-shaped, or does it remain circular? When you move the coin away from you, does the coin itself get smaller, or does it remain the same size? While it is difficult to settle phenomenological questions about the shape and size the coin *looks* to have, we have no trouble at all answering non-phenomenological questions about the shape and size of the coin itself on the basis of perception. We are rightly much more confident in our external-world beliefs about the coin itself than in our phenomenological beliefs about our experience of the coin (cf. Schwitzgebel 2006).

The second problem with the above motivation for Abductivism is that it confuses epistemic *security* with epistemic *priority*. Even if beliefs about our occurrent phenomenology were especially secure relative to our beliefs about the external world, the inference from this to the claim that they play a foundational role in justifying our beliefs about the external world is a non-sequitur. Epistemic security and epistemic priority come apart about as often as they go together. They will, for example, nearly

always come apart in cases where one belief is derived from another by way of deductive inference. Suppose I come to believe that Harry plays the harp. On the basis of this belief, I form the belief that someone plays the harp. The belief that Harry play the harp is epistemically prior to the belief that someone plays the harp, but the former is logically stronger than, and therefore less epistemically secure than, the latter.

Against explanatory irrelevance: The assumption of explanatory irrelevance involved the the Abductivist challenge to PA1 is also doubtful. A full explanation of our perceptual experiences must involve an explanation of why they have the content that they have. And it is doubtful that we can adequately explain the color content of our experiences without positing colored things in the world. Color experiences represent color properties. If we assume that color properties are instantiated in the world, we can begin to give an explanation for how we manage to perceptually represent these color properties. We might, for example, explain the color content of our perceptions by appeal to causal covariation between certain internal states of ours and instantiations of color properties (Dretske 1988; Tye 1995). To be sure, even with the hypothesis that colors are instantiated in the world, no one has managed to give a fully worked out explanation for how our color experiences represent the colors they do; all we have are programmatic sketches of explanations. But without the hypothesis that colors are instantiated in the world, no one has managed to offer even so much as a programmatic sketch of how our color experiences represent the colors they do. This gives us reason to think that the best explanation for some aspects of our color experiences, namely their content, will involve the hypothesis that color properties are instantiated in the external world. Similar

considerations suggest that the same holds for other sensible qualities as well. (These points are closely related to the abductive argument for realism in §4.9.)

This claim is further supported by the assumption of empirical anti-skepticism. If one holds that the correct explanation of the color content of our experiences makes no use of the hypothesis that color properties are instantiated in the external world, then it's hard to see what principled reason there could be to hold that the correct explanation of our spatiotemporal experiences will require the hypothesis that spatiotemporal properties are instantiated. But for the reasons given in §4.4.2, any view according to which we are not justified in believing things to have even the approximate spatiotemporal features they seem to have is plausibly a view that fails to satisfy the empirical anti-skeptical constraint. So such a view is not available to my opponent. Given empirical anti-skepticism, then, it is doubtful that Abductivism presents a serious threat to our justification to believe realism about the sensible qualities.³⁸

The above remarks rely on the assumption (which I will continue to make throughout this paper) that perceptual experiences have content; that is, perceptual experiences have accuracy conditions in accordance with which they can be assessed as veridical or falsidical. Indeed, this seems to be a truism. Nonetheless, this assumption is occasionally denied by some (though certainly not all) “direct realist” philosophers of mind (Travis 2004). The above way of responding to the challenge from Abductivism will therefore not be convincing to direct realists of this persuasion. But this is no cause for concern. The considerations which incline one toward direct realist views in the

³⁸ I discuss related issues in more detail in §4.5.1.

metaphysics of perception will typically incline one away from Abductivist views in the epistemology of perception. More importantly, direct realists deny the assumption of explanatory irrelevance involved in the Abductivist challenge to PA1. According to direct realism, the phenomenal character of a normal perceptual experience of a ripe tomato is partly constituted by the redness of the tomato. For direct realists, then, the instantiation of color properties therefore figures in the explanation of (the phenomenal character of) our color experiences. So given direct realism, the challenge from Abductivism will not get off the ground.

4.5 Objections to (Non-Reductive) Realism: Skeptical Challenges

In light of the testimony of our senses, color realism carries the status of a default presumption. Irrealists may well agree that perception provides us with *prima facie* justification to believe that there are colored things. But such irrealists will hold that there are considerations that defeat this justification. In §§4.5-4.7, I consider what I take to be the most powerful objections to color realism. Many objections to realism in some way rely on the assumption that reductionism is false. For if some form of reductionism is correct, there can be no serious doubt about realism. No one doubts that material objects really have the sorts of physical properties and response-dependent properties with which reductionists identify colors. But of course, given the non-reductionist position defended in chapters 2-3, my defense of realism cannot take refuge in reductionism. In what follows, I therefore take for granted the falsity of reductionism, in agreement with my irrealist opponent.

One common objection to realism, at least in its non-reductive forms, is that (non-

reductive) realism leads to skepticism about the colors of things; in other words, (non-reductive) realism has the result that we do not know the colors of things in our environment.

It's worth pausing to ask why this conclusion would be a problem for realism. Compare: suppose you're a realist about moral desert; you believe that properties like *deserving to be punished* and *deserving to be rewarded* (as well as their determinates, such as *deserving to be punished severely* and *deserving to be rewarded lavishly*) are really instantiated by individuals. Suppose I then persuade you of the following conditional: if realism about moral desert is true, then we do not know what individuals in our community deserve; that is, we do not know the extent to which any given person deserves to be punished or rewarded, or indeed whether any given person deserves to be punished or rewarded at all. Should this lead you to reject realism about moral desert? It's not clear that it should. Similarly, if it could be shown that color realism has the result that we don't know the colors of things in our environment, it's not entirely clear why this should lead one to reject color realism.

It might be thought that if we don't know the colors of things in our environment, then we don't know the more general proposition that things in our environment are colored. And perhaps it is irrational to believe something if we're convinced that we don't know it. (Why would this be irrational? Well, an application of conjunction-introduction would then yield a Moore-paradoxical belief of the form: p and I don't know that p . And perhaps it's irrational to hold such Moore-paradoxical beliefs.) The problem is that it's not clear why we should accept that we don't know the general proposition that things in

our environment are colored, even if it's granted that we don't know propositions attributing specific colors to things. (Recall the argument in §2.4.2 that it's possible to know that a given object is green, and even that it (therefore) has some determinate shade of green, even though we can't know which shade of green it has.)

In any case, most realists (myself included) hold not only that things are colored, but that we can know (at least approximately) the colors of things in our environment. I shall therefore grant to my objector the assumption that realism stands or falls with the latter claim.

Skeptical challenges to realism typically begin by arguing that there are possible communities whose members perceive things as having different colors than we do. There is, for example, a possible community of color perceivers whose color experiences are red-green inverted with respect to ours—i.e. a community whose members experience ripe tomatoes and firetrucks as green and unripe tomatoes and fresh grass as red. Plausibly, it cannot be the case the members of *both* communities are having veridical color experiences, since it's plausible that nothing can be both red (all over) and green (all over) at the same time. But now there is a skeptical worry. What reason do we have to think that our community is the one whose members perceive colors veridically?

As stated, the argument isn't especially compelling. The mere possibility of individuals whose perceptual representations of objects are systematically inconsistent with ours is not sufficient to raise serious doubt about the reliability of our perceptual faculties. Presumably there is a possible community whose visual experiences are left-right inverted relative to our own (whose visual experiences of the a lowercase "d" would

be like our experiences of a lowercase “b,” and *vice versa*), or who visually experience objects as uniformly stretched in a certain direction relative to the way we experience things (cf. Thompson 2010; Chalmers 2012, ch. 7; forthcoming-a). (One might deny the possibility of such a community on broadly “externalist” grounds (cf. §2.3), but the same considerations should then lead one to deny the possibility of red-green inverted communities.) But surely the mere possibility of such a community should not drive us to skepticism about the reliability of our spatial perception. Why should the possibility of communities whose color perceptions conflict with ours be any different?

Typically those who raise skeptical challenges for non-reductive realism will try to bolster the argument given above by invoking the claim that communities whose color perceptions are inconsistent with ours could easily have arisen through natural selection (Pautz ms., Chalmers 2006). As Chalmers writes,

Evolutionary processes will be indifferent among perceivers in which apples produce phenomenally red experiences, perceivers in which apples produce phenomenally green experiences, and perceivers in which apples produce phenomenally blue experiences. Any such perceiver could easily come to exist through minor differences in environmental conditions or brain wiring. If we accept the reasoning above, only a very small subset of the class of such possible perceivers will normally have veridical experiences, and there is no particular reason to think that we are among them (p. 69).

How is skepticism about the colors of things supposed to result from these considerations? The last sentence above might be taken to suggest that, unless we can provide some positive reason to think that our perceptual experiences are veridical, we

aren't justified in believing things to be the way they seem to be. But this claim is doubtful. It's plausible that we have a (defeasible) default entitlement to believe our senses to be reliable (Wright 2004, White 2006). More charitably, we can interpret Chalmers as claiming that the above evolutionary considerations defeat any justification we might have had to believe our color experiences to be veridical. On one interpretation, Chalmers is arguing roughly as follows:³⁹ Let V be the proposition that our color perceptions are generally veridical. Chalmers relies on the following two claims:

S1. Our visual system arose through evolutionary processes which could easily have produced creatures whose visual systems represent colors in a manner that is inconsistent with our actual color representations.⁴⁰

S2. The probability of V given S1 is very low.

If we are to derive a skeptical conclusion from S1 and S2, we'll have to rely on something like the following "epistemic" premises:

S3. Anyone who accepts S1 and S2 possesses a defeater for V.

S4. Anyone who possesses a defeater for V possesses a defeater for the beliefs formed on the basis of color experience, including beliefs about the colors of things in her environment.

³⁹ The following reconstruction is modeled very loosely on Plantinga's (1993, ch. 12) famous evolutionary argument against naturalism.

⁴⁰ In light of the facts of variation in the perception of fine-grained color shades among normal perceivers discussed in §2.4, it would perhaps be better to say "in a manner that is *grossly* inconsistent without our actual color representations." The question of whether an inconsistency is gross is a matter of the distance between the corresponding points in color space. For example, an experience of an object x as scarlet is grossly inconsistent with an experience of x as sky blue, but not grossly inconsistent with an experience of x as crimson. For simplicity, I ignore these complications in what follows.

Given natural assumptions about the relation between knowledge and epistemic defeat, it then follows that anyone who accepts S1 and S2 does not know the colors of things in her environment. The realist who is persuaded of S1 and S2 is then saddled with skepticism.

It appears that this argument is valid, or at least it could be made valid by adding some fairly uncontroversial premises. Moreover, S2 and S4 seem to be plausible.

However, as I shall argue, S1 is very likely false, and S3 is questionable as well. Let us take these in turn:

4.5.1 Against S1

Is there any reason to suppose that evolutionary processes could easily have produced creatures whose visual systems represent colors in a manner that is inconsistent with our color representations? It is of course plausible that evolutionary processes could easily have produced creatures whose visual systems represent a different range of colors than do our visual systems. Indeed, we have more-or-less conclusive empirical grounds for thinking that evolutionary processes *have* produced many such creatures; we know (about as well as we know any truths about other minds) that goldfish, pigeons, mantis shrimp, and many other animals experience a different range of colors than humans do. Of course, that the color representations of other animals are often *different* from ours does not entail that their color representations are *inconsistent* with ours. But my opponent might offer the following argument from difference to inconsistency (cf. Pautz ms, p. 16-7): Suppose that when I look at a certain flower petal, I have an experience as of determinate color c_1 , and when a pigeon looks at the flower petal, it has a visual experience as of a different determinate color c_2 , perhaps one that falls outside of human

color space. In general, determinates under a common determinable exclude one another. Just as *having mass 4.26 kg* and *having mass 28.9 kg* (two determinates under the determinable *mass*) exclude one another, and *being a 5"×6" rectangle* and *being a circle of radius 3"* (two determinates under the determinable *shape*) exclude one another, it is plausible that c_1 and c_2 (two determinates under the determinable *color*) exclude one another.

The argument is unconvincing. There are at least three ways the realist might respond to this line of reasoning. The first (and to my mind, weakest) line of response is to deny that there is any genuine determinable common to c_1 and c_2 . One way to implement this response is to hold that, strictly speaking, c_2 and other qualities outside of human color space visible to non-human perceivers are not *colors*. Another way of implementing this response is to allow that c_2 and other such qualities are colors, but to deny that color is a genuine determinable, holding that the only genuine determinables in the vicinity are *human color*, *pigeon color*, and so forth.

Second, she might argue that determinates under *highly general* determinables such as color do not always exclude one another. This view is not entirely unmotivated: one might hold that all determinate properties, e.g. red_{17} or *being a 5"×6" rectangle*, are determinates of *being* (or *existence*), where the latter is understood as something like a maximally general determinable. But determinates (even absolute determinates—those without any determinates “under” them) under the determinable *being* do not exclude one another, for it is possible for something to be both red_{17} and a 5"×6" rectangle. Arguably

the determinable *color*, when understood as covering properties outside of human color space, is sufficiently general that its determinates don't always exclude one another.

The third (and to my mind, strongest) line of response runs as follows: first, note that it is certainly not the case that determinates of a common determinable *always* exclude one another. Red and scarlet are both determinates of the determinable *color*, but they do not exclude one another. (In this case, the latter entails the former.) *Being a rhombus* and *being a rectangle* are determinates of the determinable *shape*, but they do not exclude one another. (In this case, neither entails the other, but they can be co-instantiated, in particular by squares.) We can draw the following lesson: in modeling determinate/determinable relations, we can associate properties with regions of a quality space whose points correspond to absolute determinates. Determinates d_1 and d_2 under a common determinable will not exclude one another if the intersection of the regions associated with d_1 and d_2 is non-empty. Hence the relevant exclusion principle, if it is to be at all plausible, will have to be reformulated in something like the following way:

Determinate Exclusion (DE): If d_1 and d_2 are determinates under a common determinable D and are associated with non-overlapping regions of the quality space associated with D , then d_1 and d_2 exclude one another.

Let us assume that DE is true, setting aside the worry raised above that it may fail for highly general determinables. The question we must ask is whether (the sets of absolute determinates associated with) c_1 and c_2 overlap. They will not if c_1 and c_2 are themselves absolute determinates, for in that case their associated sets will be distinct and (therefore)

disjoint singleton sets. But the realist can reasonably hold that the color properties we visually represent are never absolutely determinate. It is widely accepted that the colors we visually represent *sometimes* fail to be absolutely determinate, as when we represent the colors of things in the periphery of our visual field. It is not unreasonable to hold that the colors we visually attribute to objects in the foveal region of our visual field, though surely far more determinate, are still not absolutely determinate. By analogy, consider the case of length representation. When I look at an iron bar, my visual experience represents the bar as having a certain length. But the length attributed to the bar by my experience is not surely not an absolutely determinate length. For this would have the deeply implausible consequence that there is some real number r such that the experience is veridical iff the bar is *exactly* r inches long—not a trillionth of an inch shorter or longer. Visual representation is simply not that exacting. And what holds for length representation plausibly holds for our perceptual representations of color as well. The realist ought to say that c_1 and c_2 are not absolute determinates, but rather are *fairly* determinate properties whose associated sets of absolute determinates partially overlap, and hence that c_1 and c_2 do not exclude one another.⁴¹

I conclude that there is no easy transition from actual variation in color perception across the animal kingdom to S1. Furthermore, there are good philosophical reasons to

⁴¹ Are the shared determinates colors? Well, they fall under the determinable *color*, at least on the assumption that c_1 and c_2 do and the determinate-determinable relation is transitive. Whether we call them colors or not, they are presumably not properties experienced by any actual creature. One possibility is that the relevant absolutely determinate properties are just highly specific *physical* properties (Yablo 1995). It is consistent with non-reductionism about the color properties we experience that these color properties are determinables of physical properties, so long as the relevant color properties are not reductively analyzable in terms of physical properties.

reject S1. Note that S1 entails that evolutionary processes could easily have brought about creatures whose visual systems systematically *misrepresent* the colors of things. But this goes against the dominant philosophical tradition concerning the nature of representation and the determinants of content. According to the dominant tradition—represented by such disparate theories of intentionality as interpretationism (Lewis 1974, Davidson 1973, Williamson 2008, ch. 8, Dennett 1987), causal covariation theories (Dretske 1988, Tye 1995, Stalnaker 1984), teleosemantic theories (Millikan 1984), and Fodor’s (1987) theory of asymmetric dependence—veridical representation is in some sense the default. Each such theory comes with what I’ll call a “presumption of veridicality,” in the sense that, on each such theory, normal representation is veridical representation; misrepresentation constitutes a sort of aberration from normality. To illustrate, consider a simplistic causal-covariation theory of perceptual representation. According to this theory, an organism *o* perceptually represents *F* just in case *o* tokens a state type *S* such that, under normal conditions, *S* is tokened iff *F* is instantiated and because *F* instantiated (cf. Cutter and Tye 2011, p. 91). The simple causal-covariation theory entails that it would be difficult for there to arise a species whose members normally had radically inaccurate color representations. In my view, while the simple tracking theory is surely overly simplistic, the correct view of perceptual intentionality is likely somewhere in the vicinity. And if anything in the vicinity of the simple causal covariation theory is true, then S1 is probably false. More generally, if *any* theory which carries a presumption of veridicality is correct—which will be the case if any of the theories listed above is even roughly on the right track—then S1 is probably false.

To my knowledge there has been no serious attempt to offer even a rough sketch, in reductive or non-reductive terms, of the conditions under which we represent a property or content which does not come with a presumption of veridicality. Furthermore, even my opponent has strong reason to believe that the correct theory of perceptual intentionality will come with a presumption of veridicality. Recall that, in the present dialectical context, we are assuming that all parties to the dispute accept empirical anti-skepticism—the claim that we are justified in accepting many of the deliverances of empirical science. Anyone who accepts empirical anti-skepticism ought to allow that our perceptual representations of spatiotemporal properties are typically at least approximately veridical. For, as was mentioned in §4.4.2, if we think that perception gets things radically wrong even with respect to spatiotemporal matters, then it is hard to see how empirical science, much of which characterizes the world in spatiotemporal terms—could come out unscathed.⁴² But if our perceptual representations of spatiotemporal properties are typically veridical, then there must be some explanation for this remarkable correspondence between our spatiotemporal representations and the spatiotemporal facts. And it is difficult to see how this remarkable correspondence could be explained unless the correct theory of perceptual intentionality, the principles of which will figure in the

⁴² Even setting aside the justification of empirical science, the costs of irrealism about spatiotemporal properties are utterly exorbitant. It's *a priori* that whatever is (say) a table, or a house, or a butter knife, or (insert almost any English noun that comes to mind) is spatially extended. So if nothing is spatially extended, it follows that there are no tables, houses, butterknives, ... Indeed, there wouldn't even be simples arranged table-wise, house-wise, or butter-knife-wise, for it's *a priori* that the *xs* are arranged table-wise (or ...) only if the *xs* are spatially related. Fodor (1990) writes that if mental states aren't literally causally efficacious, "then practically everything I believe about anything is false and it's the end of the world" (p. 156). But rejecting realism about spatiotemporal properties would require an even more drastic revision to our web of beliefs than would rejecting the causal efficacy of the mental.

complete explanation of this correspondence, carries a presumption of veridicality. So anyone who accepts empirical anti-skepticism has good reason to think that the correct theory of perceptual intentionality, whatever it is, will be such as to make veridical representation the default. From this it follows that all parties to the debate have good reason to think that S1 is probably false.

It's worth noting that objections to S1 apply to an alternative objection to realism which also makes use of S1 (or some similar premise). The envisaged alternative argument goes as follows: realism (let's assume) is committed to V. It follows from V that there is a systematic and widespread correlation between the colors things appear to have and the colors they actually have. Such a systematic and widespread correlation calls out for explanation. If S1 is true, then it would appear that such an explanation will be hard to come by. S1 entails that evolutionary processes could easily have brought about creatures whose visual systems usually misrepresent the colors of things, and if we are sympathetic to S1, we will likely think that evolutionary processes are far more likely to bring about such creatures. Suppose this is correct. Then if realism (and hence V) is also correct, there would appear to be no explanation for why we evolved a visual system that accurately represents the colors of things, rather than one of the vast range of possible visual systems that typically do not accurately represent the colors of things. So given S1, it appears that realism, with its commitment to a systematic and widespread correlation between the colors of things and their apparent colors, is committed to a kind of pre-established harmony or vast cosmic coincidence—surely an unpalatable consequence.⁴³

⁴³ For related arguments, see Pautz (2011) and Byrne and Hilbert (2007a, pp. 96-8).

In my view, this is a somewhat more promising line of objection to realism than skeptical challenges of the sort raised by Chalmers. However, due to its reliance on S1, it nonetheless inherits the problems with S1 discussed above. I agree that there must be some explanation of the correlation implicit in V. But in light of the remarks made above, it is very likely that such an explanation exists. The full explanation for the correlation will make reference to the general principles involved in the correct theory of perceptual representation—the principles involved in the metaphysical explanation for why it is that representational states have the content they do. If these principles come with a presumption of veridicality, then there will be an explanation for the correlation implicit in V. And as I have argued, all parties to the debate have strong reason to think that they do.

4.5.2 A Challenge to S3

In my view, S1 is by far the weakest link in the above argument. However, S3 may also be challenged. Why should we think that if I accept S1 and I accept that the probability of V given S1 is very low, then I possess a defeater for V? It is obviously not true in general that if I accept p and I accept that the probability of p given q is very low, that I have a defeater for q. Suppose I pull a card at random from a well shuffled deck. I look at the card and see that it is the ace of spades. The probability that I pulled an ace given one of my beliefs—namely, that I just pulled a card at random from a well-shuffled deck—is low. But I do not thereby have a defeater for my belief that I pulled an ace, since the probability that I pulled an ace given *all* my evidence—which includes the proposition that I just pulled the ace of spades—is very high. Plausibly, what matters epistemically is

what is likely conditional upon *all* my evidence. In order to establish that V is unlikely given all my evidence, my opponent would have to assume that facts about the colors of things in my environment don't belong to my total evidence. For suppose they do. My evidence also presumably includes various facts about the colors things *look* to have. So my total body of evidence includes many pairs of propositions of the form: *o has color c*; *o looks to have color c*. If that's the case, then it would seem that the probability of V conditional upon my total body of evidence should be very high. But it would be question-begging in this context for my opponent to assume that facts about the colors of things do not belong to my evidence. Assuming that my evidence includes what I know (Williamson 2000, ch. 9), this would amount to the assumption that I don't know the colors of things in my environment. But this is precisely the skeptical conclusion my opponent is trying to establish!

I conclude that S1 is very likely false and S3 may not survive critical scrutiny. The above interpretation of Chalmers's evolutionary skeptical challenge fails.

4.5.3 Evolution, Safety, and Luck

Byrne and Hilbert (2007a) offer an alternative interpretation of Chalmers's evolutionary skeptical challenge: "[Non-reductive realism] leads to skepticism, according to Chalmers, for the straightforward reason that if [non-reductive realism] is true, our beliefs about the colors of things could easily have been false, and hence do not amount to knowledge" (p. 88). On this interpretation, the relevant facts about the evolution of our visual systems are supposed to lead us to the conclusion that our beliefs about the colors of things fail the *safety* condition on knowledge, and hence, even if true, do not constitute knowledge. The

safety condition on knowledge is typically formulated in something like the following way:

Safety: One knows *p* only if one could not easily have believed *p* falsely—i.e. only if there are no nearby possibilities in which one falsely believes *p* (Cf. Williamson 2000, p. 128).⁴⁴

This version of Chalmers's evolutionary skeptical challenge is even more problematic than the previous interpretation. First of all, in order to establish that our color beliefs could easily have been false, the argument from safety must rely on S1 or some similar premise, so it inherits the most serious problems with the previous argument. Second, even if we accept S1, and so we accept that alternative color perceivers could easily have arisen in evolutionary history, it is not plausible that our beliefs about the colors of things, if true, could easily have been false—at least not in any sense which would render them unsafe. Worlds in which evolution goes very differently, in which instead of humans there arises a species of hominids whose color representations diverge drastically from our own, are not relevant to considerations of safety. This is so for three reasons: first, in assessing whether my belief that grass is green is safe from error, we consider other nearby worlds in which I believe that grass is green and ask whether what I believe is true in those worlds. But the worlds with alternate evolutionary histories are almost certainly worlds in which I don't exist, and so a fortiori

⁴⁴ A fully adequate formulation of the safety condition will likely have to add qualifications which restrict the relevant set of nearby possibilities to those in which one forms the belief that *p* using the same method one uses in the actual world to form the belief that *p*. But such niceties needn't concern us here.

are not worlds in which I have the belief that grass is green. Secondly, suppose I do exist in some of these worlds, despite having a very different evolutionary history, and suppose that in some of these worlds I have non-veridical color experiences and so form mostly false beliefs about the colors of things. None of this gives us the slightest reason to suppose that any of my actual color beliefs could easily have been false. Consider again my belief that grass is green. Since I have very different color experiences in these worlds, I presumably won't form the belief that grass is green in these worlds. And if my actual belief that grass is green is true, I certainly don't believe *falsely* that grass is green in these other worlds. (Grass presumably doesn't change color across these nearby possibilities.) So these nearby possibilities are irrelevant to the question of whether my actual belief that grass is green is safe from error. Formally, the problem arises from a scope confusion in the claim that:

E: My beliefs about the colors of things could easily have been false.

There are (at least) two syntactic disambiguations of E (read " $\diamond\phi$ " as "it could easily have been the case that ϕ ," and read "Gen" as the binary quantifier involved in the logical form of generic sentences (cf. Leslie 2008; Asher and Morreau 1995):

E1: $\diamond(\text{Gen } x [x \text{ is one of my beliefs about the colors of things}][x \text{ is false}])$.

E2: $\text{Gen } x (x \text{ is one of my beliefs about the colors of things}) \diamond(x \text{ is false})$

To the extent that considerations of alternative evolutionary histories support E at all, they support E1, not E2. But it is E2, not E1, which is relevant to the question of whether my actual beliefs are safe from error.

Third, and most importantly, these worlds are simply not nearby possible worlds by any metric which could be relevant for applying the safety criterion. Granted, it's not always clear which worlds we should count as "nearby possibilities" in assessing the safety of a belief. Suppose that unbeknownst to me, I am driving through a county with mostly fake barns, but I also happen to be in a local district which has only real barns. Upon looking at one of these barns, I form the belief that there is a barn before me. In assessing the safety of this belief, should we consider worlds in which I am looking at a different barn in the same district? Worlds in which I am looking at a fake barn in a different district within the same county? It is not easy to say exactly which worlds are relevant. But one thing is clear: worlds in which evolution went very differently, in which there are hominids around with very different neural wiring from that of actual humans, are certainly not relevant to assessing the safety of my belief. This is true whether we are assessing the safety of my belief that there is a barn before me or my belief that there is a red thing before me.

Pautz (ms.) offers a similar argument that non-reductive realism results in skepticism by appeal to the following principle:

Independence: If objects have primitive colors, what color vision system evolves in a given lineage is completely independent of what primitive colors objects possessed prior to the evolution of color vision (p. 12).

From Independence, Pautz argues that it is extremely unlikely that humans' visual systems evolved so as to represent the actual colors of things. So even if, by chance, we happened to have evolved a visual system which typically represents the actual colors of

things, our beliefs about the colors of things will only be true “by blind luck,” and therefore do not constitute knowledge (p. 17).

The most serious problems with this argument, in my view, are those it shares with the first interpretation of Chalmers’s argument above. In particular, the objections to S1 in §4.5.1 apply to Independence as well. Independence will fail on just about any theory of perceptual representation that carries a presumption of veridicality. On causal-covariation theories of perceptual representation, for example, what color vision system evolves in a given lineage will be partly dependent on which color properties are causally related to internal states of members of that lineage, and therefore will be dependent on which color properties things actually have. (I am assuming that color vision systems here are to be individuated partly in terms of the range of colors they represent, and not merely in terms of their internal physical structure.)

But let us focus on the second part of Pautz’s argument—namely, the inference from Independence to the claim that our beliefs about the colors of things, if true, are only true by “blind luck” and so do not amount to knowledge. It is widely agreed that luck can undermine knowledge. This is one of the lessons we’re supposed to have learned from Gettier’s (1963) counterexamples to the analysis of knowledge as justified true belief. But not all forms of luck undermine knowledge. Consider the case of Theodore. When Theodore’s mother was pregnant with him, she participated in a drug study with ninety-nine other pregnant women. Ninety-five of the participants were selected at random to take the drug, and the other five were given a sugar pill. Unfortunately, the drug caused a peculiar cognitive defect in the children of women in the first group, with

the result that, when the children came of age, their arithmetic beliefs were mostly false. Fortunately for Theodore, his mother, by “blind luck,” was one of the five women who were given the sugar pill rather than the harmful drug. Theodore is now a normal, reasonably intelligent adult with arithmetical abilities comparable to those of other normal adults. In some sense, Theodore’s beliefs about simple arithmetic are true by luck. He is lucky to have reliable arithmetic belief-forming faculties. But there is no temptation at all to say that, on that account, Theodore’s arithmetic beliefs do not amount to knowledge. The kind of luck involved here is not knowledge-undermining luck. If Independence is true, our case is analogous to that of Theodore. If indeed our visual system accurately represents the colors of things, we are in some sense lucky to have true beliefs about the colors of things; we are lucky to have reliable chromatic belief-forming faculties in just the way that Theodore is lucky to have reliable arithmetic belief-forming faculties. But this is not knowledge-undermining luck. If Independence is true, our situation is more like that of Theodore and less like that of Gettier’s Smith.

4.6 Objections to (Non-Reductive) Realism: Causation and Simplicity

The next objection to non-reductive realism is the *causal-exclusion* objection. The causal-exclusion objection, which is roughly analogous to the more familiar causal-exclusion objection to dualism and non-reductive physicalism about mental properties (Kim 1993), may be formulated as follows: science tells us that the properties of objects on which our sensory experiences causally depend are complex physical properties. For example, our color experiences are causally dependent on physical properties of objects, such as their dispositions to reflect, transmit, and emit light, and the micro-structural

properties that ground these dispositions. Our auditory experiences of pitch, loudness, and timbre are causally dependent on physical features of pressure waves, such as their fundamental frequency, amplitude, and wave shape. Our thermoceptive experiences of heat and cold are causally dependent on the temperature and conductivity of their objects. And so on. If the secondary qualities are distinct from any physical property, as the non-reductive realist supposes, then (barring overdetermination) it would appear that they do not have any causal relevance with respect to our experiences. This would be a troubling result. If our experiences do not causally depend on the secondary qualities of things, then it would appear that we do not really *perceive* the secondary qualities of things. And this conclusion would seem to undercut our justification for believing that the objects of experience have the secondary qualities they appear to have.

To this objection we may add a related but distinct objection, which we may call the “simplicity objection” (Chalmers 2006, p. 67; Pautz ms.). If non-reductive realism is true, then physical objects have secondary qualities that are distinct from any of the properties attributed to them by the physical sciences. Hence the non-reductive realist seems to be committed to a more metaphysically complex world than one who identifies secondary qualities with certain physical properties or denies that they are instantiated. As Pautz (ms.) writes,

[Non-reductive realism] is ontologically inflationary. It requires a kind of Dualism at the surfaces of objects. [Irrealism] avoids such a Dualism. It has the virtue of simplicity. This provides a strong reason to accept [Irrealism] over [Non-reductive realism] (p. 48).

A non-reductive realist might be tempted to respond to the simplicity objection as follows:

Considerations of simplicity are to be used to decide among theories that equally well explain, or at least accommodate, our evidence or data. Our data minimally include our direct observational data, which in turn minimally include propositions concerning the distribution of sensible qualities in our environment. The simplicity objection uses considerations of simplicity not to decide among theories that equally well explain (or accommodate) our observational data, but rather to *deny* our observational data. And this is an illegitimate use of Ockham's razor.

However, we cannot dismiss the simplicity objection so easily. Considerations of simplicity properly play a much larger role in rational inquiry than the above reply suggests. Consider an early astronomer observing the motions of celestial bodies across the sky. Perhaps he takes his direct observational data to include a large collection of propositions like "Altair (absolutely) moves in such-and-such pattern" and "Venus (absolutely) moves in so-and-so pattern." Initially he accepts a fairly inelegant geocentric theory, rife with epicycles and equant points, to account for the retrograde motion of the planets. But eventually it dawns on him that there is a much simpler and more elegant heliocentric alternative on which the planets revolve around the sun along elliptical orbits with the sun at one of the foci. This alternative is, strictly speaking, inconsistent with what he took to be his data, but he notes that if he systematically revises his putative data, replacing attributions of absolute motion with corresponding attributions of apparent (observer-relative) motion, the resulting body of propositions will be consistent with, and

elegantly explained by, the heliocentric theory. It would be rational for our astronomer to revise his beliefs on matters of (putative) observational data in light of such theoretical considerations. Likewise, it might be rational for us to systematically revise our direct observational beliefs about the distribution of sensible qualities in our environment, retaining merely our beliefs about the *apparent* distribution of sensible qualities, if doing so would result in a significantly simpler and more elegant picture of the world.

It will not be possible to respond adequately to the causal exclusion and simplicity objections without saying more about how the secondary qualities are related to physical reality. That is the task of chapter 5. I shall therefore postpone my response to these objections until then.

4.7 Beyond Color

I have defended in detail the specific instance of the perceptual argument template which pertains to realism about color. In this section, I argue that this defense generalizes to the other sensible quality classes as well, with only a few minor complications.

For each major sensible quality class, we can fill in the perceptual argument template in the appropriate way to yield an argument in support the conclusion that we are justified in believing realism about that class. The issues that arise in defending the other instances of the perceptual argument template closely parallel the issues discussed in detail in §§4.5-4.6 that arise for the case of color. For example, in each case we can note that the qualities in question *appear* to be instantiated in our environment (or in our bodies) and that this fact gives us *prima facie* justification to believe they really are instantiated in our environment (or bodies). And in each case, there will be concerns

about causal exclusion and simplicity that arise at least for non-reductionist forms of realism about the relevant class of qualities. In many cases, there will also be skeptical concerns analogous to those discussed in §4.5. My responses to these worries will exactly parallel my responses in §§4.5.1-4.5.3 and (for causal exclusion and simplicity) in §5.3.

Are there any relevant disanalogies? In particular, are there objections to realism about pitch, loudness, timbre, flavor, scent, heat, cold, or other sensible qualities which do not have analogues in the case of color? In what follows, I consider and respond to three objections that might be raised to realism about other sensible qualities which do not have analogues in the case of color. In §4.7.1, I consider objections to realism about allegedly non-spatial qualities. In §4.7.2, I consider objections to realism about “sensational qualities,” the qualities we experience when we undergo bodily sensations. In §4.8, I consider an objection that alleges that modern physics undermines (something like) realism about shape.

4.7.1 Non-Spatial Qualities

Hume held that impressions derived from sensory modalities other than vision and touch are entirely non-spatial. Many contemporary philosophers agree that, at least in some sensory modalities, perception is non-spatial. This view is perhaps most natural for olfactory perception (Lycan 2000). But it is also often suggested that auditory perception is non-spatial as well. O’Shaughnessy (2002) writes, “We absolutely never immediately perceive sounds to be at any place. (Inference from auditory data being another thing)” (p. 446). If these claims are correct, then this suggests a potential objection to realism about the qualities experienced in olfactory and auditory perception. Suppose it is

correct that when I experience a smoky scent or a low pitch, these qualities aren't experienced as instantiated at any location in my environment. It might be inferred from this that the relevant qualities are therefore not objective or public, perhaps on the Kantian grounds that spatiality is a necessary condition for objectivity (Kant 1781/1999), or perhaps on the grounds that it would require a significant modification of the scientific view of the objective world to countenance properties which are instantiated without being instantiated at any spatial location. And if these qualities are not instantiated in the objective (public) world, then realism about these qualities is false.

There are two problems with this line of argument. First, the suggestion that olfactory and auditory experience are non-spatial is questionable. This is especially clear in the case of audition. Intuitively, we hear sounds as being located in certain positions relative to us (O'Callaghan 2007; Pasnau 1999). If you hold your left hand out at your side and snap your fingers, you will distinctly experience the sound as being on your left. It does not seem as though we merely infer or "work out" the location of a sound event on the basis of the raw auditory data. Rather, we simply *experience* sounds as having certain locations, just as we visually experience objects as having certain locations—though plausibly the locative features ascribed by auditory experience are far less determinate than those ascribed by visual experience. It is admittedly less phenomenologically obvious that olfactory experience has locative content. Nonetheless, there is good reason to think that olfactory experience does represent (highly non-specific or determinable) locative properties. Consider an ordinary experience as of a smoky odor. Plausibly, this olfactory experience represents a certain quality—call it "smokiness"—as

being instantiated “here-ish” (Batty 2010). The view can be motivated by appeal to our judgments about veridicality. Suppose I have an olfactory experience of exactly this sort, but it turns out that the relevant smoky quality is not instantiated anywhere in my vicinity, but is instantiated somewhere else—say, on the other side of the world. If queried, we would intuitively judge my olfactory experience to be non-veridical. (Admittedly, it is somewhat less natural to assess olfactory experiences for veridicality than it is to assess, say, visual or auditory experiences for veridicality. But on reflection, we can make sense of the notion of olfactory illusions or hallucinations, which suggests that we have a grip on the notion of the veridicality conditions for olfactory experiences.) But if all there is to the content of my olfactory experience is the information that smokiness is instantiated, then, since by hypothesis smokiness *is* instantiated (on the other side of the world), my experience should be veridical. Since my experience is not veridical, it follows that there is more information built into the content of my experience. And what could this be other than *locative* information?

The second problem with the above objection to realism concerning auditory and olfactory qualities is more serious: even if we grant the assumption that audition and olfaction are non-spatial modes of perception, realism about auditory and olfactory sensible qualities would not commit us to properties which are instantiated without being instantiated at any spatial location. It does not follow from the claim that we do not experience pitch or scent as being instantiated at a location that these qualities, if instantiated, are not instantiated at a location. Nor would their instantiation at a location imply that the relevant perceptual experiences are guilty of error. It is one thing for an

experience to lack locative content—to fail to represent a quality as having a location; it is another for the experience to have negative locative content—to represent a quality as not having a location. Even if we accept the claim that the qualities experienced in olfactory and auditory experience are not experienced as having a location, it is entirely implausible to suppose that we experience them as having no location. I see no objection to saying that the relevant qualities are instantiated at spatial locations in our environment, even if the content of our experience is entirely noncommittal about their locations.

4.7.2 Bodily Sensations

Consider the class of sensory experiences that might be called “bodily sensations.” Bodily sensations include the experience of an orgasm, the experience of a pain in your leg, the experience of an itch on your elbow, the experience of a tickle on your toe, and so forth. It is sometimes said that such experiences are “raw sensations,” experiences devoid of representational content, and that they therefore differ importantly from typical visual, auditory, and tactile experiences. Of course, everyone should allow that when we undergo a bodily sensation, we are aware of certain qualities, which we might call *sensational* qualities, including painful, pleasant, and itchy qualities. But according to the view in question, because our experiences of these qualities lack representational content, we do not sensorily represent these qualities as instantiated in our environment or in our bodies; they do not *seem to us* to be—we do not experience them as—instantiated in the objective world. If this view is correct, then sensory experience does not give us justification to believe that they are instantiated in the objective world.

But the view of these experiences as “raw sensations” is highly dubious. Consider the itchy quality you experience when, as we say, you feel an itch in your elbow. Whether or not this quality is really exemplified in the region of your elbow, don’t you experience it as exemplified there? As you undergo the experience, doesn’t this quality *seem to you* to belong to a certain localized region in the vicinity of your elbow? And if you experience this quality as belonging to a certain bodily region—if it seems to you to belong to that bodily region—then you have *prima facie* justification to believe that it really *does* belong to that region. So you have *prima facie* justification to accept realism about such itchy qualities.

More generally, it seems clear that whenever we experience a sensational quality, we experience it as belonging to a certain bodily region or location. When a feather tickles my foot, the ticklish quality I experience is a quality my experience locates *in my foot*. The same holds for the pain quality I experience when a needle pricks my foot (cf. Cutter and Tye 2011, Hill 2009, ch. 6; Chalmers 2006, Pautz 2010a). On such occasions, we therefore have *prima facie* justification to believe that such sensational qualities are really instantiated in the relevant bodily regions.

Are there any reasons to reject realism about the sensational qualities beyond those that apply to realism about color and other secondary qualities? Here I examine one interesting challenge to realism about the sensational qualities from Chalmers (2006, pp. 114-5).⁴⁵ I will consider the objections as it applies to pain, though analogous objections are available for other sensational qualities.

⁴⁵ For a closely related argument with a slightly different target, see Pautz (2010a, pp. 348-9).

I invite you to pinch your arm until you feel a mild twinge of pain. Now focus on the painful quality—we'll call it *q*—which seems to be instantiated at the location of the pinch. Upon reflecting on the nature of *q*, Chalmers claims, two facts become evident:

Q1. It is not possible for *q* to be instantiated without being experienced.

Q2. *q* is (nonetheless) intrinsic.

From Q2 it follows that *q* is not a response-dependent property; it cannot be identified with anything like the disposition to produce experiences of a certain kind. Q1 and Q2 are somewhat in tension with one another. Putting them together, we get the result that *Q* is “an intrinsic property that stands in a necessary connection to distinct intrinsic properties of experience” (p. 114). From here it can be argued that *q* is necessarily uninstantiated.

Chalmers writes,

If this property could be instantiated, problems would follow. It is not clear that there can be necessary connections between distinct existences of this sort. It seems plausible that for any conceivable or possible situation in which an intrinsic property is instantiated in one's ankle, it should be conceivable or possible that the property is instantiated in an arbitrarily different context. But it is not conceivable or possible that there is perfect pain without pain experience. The natural conclusion is that perfect pain cannot be instantiated: there is no possible world in which there is perfect pain, and on reflection it is not even conceivable that there is perfect pain. In effect, the instantiation of perfect pain places incoherent requirements on the world (p. 114).

If we understand “perfect pain” as referring to *q*, then we can interpret Chalmers as arguing for the following claim:

Q3. If it is possible for q to be instantiated, then [since q is intrinsic] it is possible for q to be instantiated without being experienced.

It follows from Q1-Q3 that q is necessarily uninstantiated. Hence realism about q is false. The above considerations plausibly generalize to all sensational qualities, including the pleasant qualities experienced during orgasm or while eating a delicious strawberry, and the itchy quality one experiences when one's elbow itches.

One option at this point is to simply accept the conclusion that realism about the sensational qualities is false. This might not be a major concession; we would not thereby be forced to abandon realism about most of the major sensible quality classes. However, this would be overly hasty. There are many possible lines of response to the above argument. I will limit my discussion to the option of denying Q1, though other options are worthy of exploration.⁴⁶

Q1 admittedly enjoys some intuitive support. For rejecting it seems to commit one to the bizarre claim that there could be a pain that was experienced by no one at all. But Q1—which is equivalent to the claim that, necessarily, if q is instantiated, then q is experienced—is just one half of an intuition which is not friendly to irrealism, namely:

Esse est percipi (EEP): Necessarily, q is instantiated if and only if q is experienced.

EEP is supported by the intuitive thought that, (necessarily) if one seems to be in pain, then one is in pain; and if one is in pain, then one seems to be in pain (cf. Hill 2009, ch.

⁴⁶ Incidentally, the theist has an easy response to the argument. The argument relies on the assumption that if F is an intrinsic property and F is possibly instantiated, then possibly F is instantiated without being experienced. But this assumption will fail if there is some necessary being x such that it lies in the nature of x that for any property F (intrinsic or extrinsic), if F is instantiated, then x consciously experiences F as instantiated. It's natural for the theist to hold that God satisfies this condition.

6). But it's not clear that the irrealist about sensational qualities can accept EEP. For EEP seems to entail that realism about *q* is true, assuming, what the irrealist accepts, that *q* is experienced. It is somewhat *ad hoc* for the irrealist to reject the EEP intuition while retaining the half of the intuition which is congenial to her view.

In my view, we ought to reject Q1, and with it EEP. I propose that the attraction to Q1 arises not from reflection on the quality we experience upon pinching ourselves. Rather, the attraction to Q1 arises from the intuitive thought that there cannot be a pain which is not experienced by anyone (just as there could not be an itch or a tickle which was not experienced by anyone). This thought is certainly correct in some sense, so long as “pain” (“itch,” “tickle”) here picks out—as it often does—a kind of experience, rather than the quality which our pain (itch, tickle) experiences are experiences *of*. For in this case, the claim just requires that every experience (of a certain sort) is experienced by someone. As Tye (1995, ch. 3) points out, taken in the right way, this is a triviality, just as it is a triviality that every laugh is laughed by someone and every smile is smiled by someone. But this triviality does not support Q1, for “*q*” in Q1 is stipulated to refer not to our pain experience, but rather to the quality that our pain experience is an experience *of*.

4.8 Shape and Special Relativity

Thus far I've focused on defending realism about the secondary qualities. Although there are very few irrealists about so-called primary qualities like shape, some have thought that modern physics casts doubt on realism the primary qualities. For example, in recent work, Chalmers (2006, 2012, forthcoming-a) has argued that what he calls “Edenic shapes”—roughly, those shape properties phenomenally presented in spatial experience,

and others under the same determinable—are not instantiated in our world. His argument is based largely on considerations of Special Relativity (SR). Edenic shapes, he maintains,

[...] might have been instantiated in the garden of Eden, and perhaps in a classical Newtonian world. But I think that there is little reason to think that they are instantiated in our world. Certainly, it is not easy to see how there could be Edenic squareness in a relativistic world. (2012, p. 333)

In this section, I defend realism about Edenic shape against Chalmers’s challenge from SR.

4.8.1 Edenic Content and Edenic Shape

To a first approximation, what Chalmers calls “Edenic properties” are the properties directly presented in the phenomenology of perceptual experience. For example, Edenic colors and Edenic shapes are, respectively, the colors and shapes directly presented in the phenomenology of color and shape experience.⁴⁷

We can define the notion of an Edenic shape somewhat more precisely in terms of Chalmers’s notion of the “Edenic content” of experience. Edenic content has two defining characteristics: first, the Edenic content of an experience is a *phenomenal* content. The

⁴⁷ Chalmers uses the “Edenic” qualifier to distinguish these properties from what he calls “ordinary” or “imperfect” colors and shapes. Ordinary colors and shapes are picked out as whatever physical properties serve as the normal causal basis for color and shape experience. On Chalmers’s view, the properties that play this role in actuality are radically different from the Edenic properties presented in experience. Edenic properties are so-called because, according to Chalmers’s fable, they were instantiated and played something like this role in the Garden of Eden. (As should be clear, the non-reductive realism defended in this chapter and the previous two chapters concerns *Edenic* properties. On my view, things in the actual world are like things in Chalmers’s Eden.) Chalmers is a (reductive, and specifically physicalist) realist about ordinary shape and color and a (non-reductive) irrealist about Edenic shape and color.

notion of phenomenal content can be defined in terms of the more familiar notions of *phenomenal character* and *representational content* as follows: C is a phenomenal content of an experience E iff necessarily, any experience with the same phenomenal character as E has representational content C (Chalmers 2006, pp. 50-1). The second defining characteristic of Edenic content is that it is Russellian, in that it has properties among its “constituents.” In this way, Edenic content differs from Fregean content, which contains senses or modes of presentation that may pick out different properties for different perceivers or in different environmental contexts.⁴⁸ The notion of a property’s being a “constituent” of a content is, of course, highly theoretical. For our purposes, the cash value of the claim that a property F is a constituent of (“belongs to,” “figures in,”) a content C is this: necessarily, any experience with content C attributes F to its object. From here, we may define an Edenic shape as a shape property that figures in the Edenic content of some experience.

Chalmers rejects realism about Edenic shape (RES). We turn now to his reasons for doing so.

⁴⁸ More precisely, Chalmers suggests that Edenic content is only partially Russellian. Russellian contents, as traditionally understood, are built up out of properties and concrete individuals. Chalmers suggests that the Edenic content of an experience is Russellian in its “predicative” components, but includes a demonstrative mode of presentation in place of the concrete object of the experience. The exclusion of concrete objects from Edenic content reflects the possibility of phenomenally identical experiences of different concrete objects. Plausibly, Edenic contents will also need to include indexical modes of presentation to pick out the subject and time of the experience as well. More on this in §4.8.3.

4.8.2 Chalmers's Case Against RES

Chalmers's reasons for rejecting RES are perhaps best presented against the backdrop of his Lorentz-Earth thought experiment, a physically possible scenario that relies on the phenomenon of Lorentz contraction:⁴⁹

According to the special theory of relativity, if an object is at about 0.87 times the speed of light relative to us, its length will contract by a factor of 2 in the direction of travel. [...] We can now imagine Lorentz Earth, which is a physical duplicate of Earth, except that it is traveling at 0.87 the speed of light relative to Earth, say, on the plane of its equator. Then according to special relativity, where Earth is roughly spherical, Lorentz Earth is compressed so that it is roughly ellipsoidal.

(forthcoming-a, p. 19)

From here, we are to suppose that Albert is standing on the north pole of Earth looking at an object, O_1 , which we would call a "square." Similarly, Twin Albert is standing at the north pole of Twin Earth looking at an object, O_2 , which he would call a "square," but which is a 2:1 rectangle as measured from Earth's frame of reference.

When looking at O_1 , Albert has a "square-ish" experience, which represents O_1 as having a certain Edenic shape: Edenic squareness. Does O_1 really have this property?

Chalmers answers "no," suggesting that none of the SR-friendly properties of O_1 can plausibly be identified with Edenic squareness. His reasoning seems to run as follows.

Given the phenomenon of Lorentz contraction, all the SR-friendly shape properties we

⁴⁹ This thought experiment is originally introduced to argue that our "ordinary" (non-Edenic) shape concepts are Twin-Earthable, though he also makes reference to it when arguing against RES.

can ascribe to an object like O_1 will involve relations to frames of reference. Some of these are properties common to O_1 and O_2 , such as what Chalmers calls “rest-squareness,” where something has rest-squareness just in case it is a square in the frame of reference in which it is at rest. Others are peculiar to O_1 , such as “ F_1 -squareness”—the property of being a square in F_1 (where F_1 is Earth’s frame of reference)—which O_1 has and O_2 lacks. But, Chalmers claims, “neither seems a good match for what is directly presented in our categorical shape experience” (forthcoming-a, p. 22).

This is clear enough in the case of F_1 -squareness. Given the physical symmetry of Albert’s and Twin Albert’s circumstances, we may reasonably suppose that their experiences are phenomenally identical. Assuming that shape experience has Edenic content, it follows that their experiences attribute the same Edenic shape, Edenic squareness, to their objects. Supposing that Edenic squareness is F_1 -squareness therefore leads to the conclusion that Twin Albert’s experience attributes F_1 -squareness to its object. But that is absurd. It is incredible to suppose that Twin Albert’s experience would represent a property involving a relation to a reference frame to which neither he nor the object of his experience belongs, and relative to which he is traveling at 0.87 the speed of light.

Another reason to think that Edenic squareness could not be F_1 -squareness or anything like it comes from the intuition that Edenic squareness might have been instantiated in a classical Newtonian world, where the notion of a frame-relative shape has no meaningful application (2012, p. 333). This is also a reason to think that Edenic

squareness could not be rest-squareness or any other shape property that involves relations to reference frames. It appears, then, that there are no SR-friendly properties of O_1 with which we could reasonably identify Edenic squareness. Of course, analogous considerations apply beyond the specific case at hand, suggesting the general conclusion that Edenic shapes cannot be identified with any SR-friendly shape properties.

We might reconstruct Chalmers's case against RES in summary form as follows:

P1. The only shape properties instantiated in a relativistic world like our own involve relations to frames of reference (e.g. F_1 -squareness, rest-squareness).

P2. Edenic shapes do not involve relations to frames of reference.

C. Therefore, Edenic shapes are not instantiated in our world.

In the following section, I argue that SR, properly understood, provides no grounds for accepting P1.

4.8.3 A Defense of RES

Special Relativity is “first and foremost [...] a theory of the geometry of space-time” (Maudlin 2012, p. xii). I shall argue in this section that analysis of this geometry reveals that that there is no more tension between RES and SR than there is between RES and the classical Newtonian picture of the world. Given the assumption that there is no tension between RES and the classical Newtonian picture—an assumption Chalmers concedes (2012, p. 333)—I shall conclude that there is no tension between RES and SR.

Setting aside SR, what would we intuitively what to say about the Edenic content of Albert's experience? To a first approximation, we might say that Albert's experience represents its object as square—not square relative to this or that frame of reference, but

absolutely square, where “absolute-squareness” is the sort of property an object could have in a Newtonian world. But there is more to the content of his experience than this. Even setting aside the representation of color and location, visual experience does not merely represent its object as having a shape; it represents its object as having a shape *at a given time*, presumably the time of experience. If at noon I have an experience of a lump of clay as spherical, it does not suffice for the veridicality of my experience if the lump is spherical at some other time. It must be spherical at noon—at the time of the experience. This intuitive truth about the content of shape experience reflects a truism about shapes themselves: shapes are properties that belong to persisting material objects only *at or relative to* a time.⁵⁰

Intuitively, then, Albert’s experience (in its Edenic content) represents its object as being (absolutely) square at t_1 , where t_1 is the time of experience.⁵¹ In assessing whether Albert’s experience could be veridical in a relativistic setting, three questions arise: (i) How should we understand the notion of a (moment of) *time*? (ii) How are we to understand “the time of Albert’s experience?” (iii) What is involved in a persisting object’s having a shape *at* a time? Before answering these questions within a relativistic setting, though, it will be instructive to consider how we would answer these questions

⁵⁰ This does not mean that nothing has a shape *simpliciter*. It might be that momentary stages of persisting things or certain three-dimensional regions of space-time have shapes *simpliciter* (cf. Lewis 2002). More on this below.

⁵¹ It is probably best to think of the Edenic content of Albert’s experience as involving an indexical mode of presentation (“NOW”), which functions much like a Kaplanian character (Kaplan 1989), picking out different times when evaluated at different contexts of experience. Given that phenomenally identical experiences may occur at different times, an indexical mode of presentation is needed in place of an individual time to ensure that Edenic content is phenomenal content (cf. footnote 48).

within a Newtonian setting—or, rather, a *neo*-Newtonian setting with a Galilean space-time.

There are two features of the intrinsic geometry of Galilean space-time of special significance for our purposes. The first is that it hosts a relation of absolute simultaneity between events or space-time points. This relation allows for a privileged foliation of space-time into “simultaneity hyperplanes,” understood as equivalence classes under this relation—or, better, mereological fusions of the members of such classes. Within Galilean space-time, simultaneity hyperplanes are the obvious choice for a physical structure with which to identify moments of time. If we make the harmless idealizing assumption that Albert’s experience is an instantaneous occurrence (perhaps a momentary neural event), we can then answer question (ii) by identifying the time of experience with the simultaneity hyperplane on which Albert’s experience falls.

The second important fact about Galilean space-time is that each simultaneity hyperplane has the topological, affine, and metrical structure of three-dimensional Euclidean space. An important consequence of this fact is that the geometric structure of a given simultaneity hyperplane licenses the description of certain of its subregions as spherical, others as square, others as 2:1 rectangles, and so forth. To give the flavor of how this will go: if we let D be a metric function that adequately represents the objective metrical structure of a simultaneity hyperplane h , then a three-dimensional subregion R of h is a (closed) sphere iff there is there is a point p on h and a constant k such that for all q on h , q belongs to R iff $D(p, q) \leq k$.

Persisting objects in Galilean space-time can be associated with a four-dimensional region representing the object's path through space-time, which we'll call its "world path."⁵² Let us call the region that results from intersecting the world path of a persisting object *x* with a given time (simultaneity hyperplane) *t* "the region *x* occupies at *t*." Turning to question (iii)—what is involved in a persisting object's having a shape *at* a time—the foregoing suggests the following principle:

Correspondence : The shape a persisting object *x* has at a time *t* is the shape of the region that *x* occupies at *t*.

Correspondence has a great deal of intuitive support. Intuitively, a persisting lump of clay is spherical at a given time if and only if the region occupied by the lump at that time is spherical. It seems incoherent, for example, to suppose that a lump of clay is spherical at a time *t* but the region it occupies at *t* is cubical.

One might object to Correspondence on the grounds that shapes as they apply to persisting objects seem to be dyadic properties—in particular, dyadic relations to times—whereas shapes as they apply to subregions of simultaneity hyperplanes seem to be monadic. One response to this objection is to deny that shapes as they apply to regions are monadic. Perhaps a spherical region is best described as spherical at the time of which it is a subregion. Another response is to deny that shapes as they apply to persisting objects are dyadic. Perhaps persisting objects and spatial regions can have the same

⁵² I use "world path" in place of the more familiar term "world line" because the latter carries the suggestion that the relevant paths are without spatial thickness. Note that the association of objects with world paths is officially neutral between endurantist and perdurantist views on the persistence of objects. An endurantist will say that a persisting object is exactly and wholly present at each time slice of the object's world path, whereas a perdurantist will say that an object is exactly and wholly present only at the world path itself, and only partially present at each time slice.

monadic property, but in different ways: where a region simply *has* sphericity (say), a persisting object *has-at-t* sphericity (cf. Johnston 1987). A third response (which I somewhat prefer) is to deny the objection's implicit assumption that a single property cannot be both monadic and dyadic. A sufficient condition on F's being a monadic property is that something has F *simpliciter*. A sufficient condition on F's being a dyadic property is that something has F relative to something. Why suppose that nothing could satisfy both these conditions? (I suspect that the assumption that no property could be both monadic and dyadic is an artifact of taking too seriously the "slot" picture of adicity, on which an n-place property is pictured as having exactly n "slots" to be filled by objects. Since it would be a contradiction for something to have both exactly one and exactly two slots, we are tempted to think that no property could be both monadic and dyadic. But we should not take the "slot" picture of acidity so seriously.)

Let us now consider questions (i)-(iii) from the standpoint of SR. Question (i) is somewhat more complicated within a relativistic setting. As is well known, Minkowski space-time (the space-time for SR) does not feature a relation of absolute simultaneity with which we might construct simultaneity hyperplanes. But the intrinsic geometry of Minkowski space-time does allow us to pick out a special class of hyper-surfaces known as "space-like hyperplanes."⁵³ A space-like hyperplane is equivalent to a maximal collection of events or space-time points that are simultaneous in some inertial frame.

⁵³ The situation is somewhat more complicated in the context of General Relativity, where there is no guarantee that a given event falls on any space-like hyperplanes. For the purposes of this essay I restrict my attention to Special Relativity.

Within Minkowski space-time, space-like hyperplanes are the most natural physical structure with which to identify moments of time.⁵⁴

Question (ii) is also more complicated within a relativistic setting. Even under the idealizing assumption that Albert's experience corresponds to a single "event" or space-time point, the task of specifying the time of Albert's experience is complicated by the fact that every event within Minkowski space-time falls along infinitely many mutually intersecting space-like hyperplanes. Despite these complications, a natural answer to (ii) suggests itself: the time of Albert's experience is that space-like hyperplane the events on which are all simultaneous with respect to Albert's own frame of reference.⁵⁵

When we turn to question (iii), however, the transition to Minkowski space-time from Galilean space-time does not affect matters at all. This is because the intrinsic geometrical structure of a time (space-like hyperplane) in Minkowski space-time is exactly the same as that of a time (simultaneity hyperplane) in Galilean space-time. Both have the geometrical structure of a three-dimensional Euclidean space. The metrical structure of a space-like hyperplane is described by what is called the *Invariant Relativistic Interval*, so-called because its magnitude between any pair of points is frame-

⁵⁴ Times understood as space-like hyperplanes will have some strange features. Because distinct space-like hyperplanes can intersect one another, any given event or space-time point will fall along infinitely many times, so understood. Another strange feature of times understood as space-like hyperplanes is that there will be no objectively distinguished linear ordering of times. This is a consequence of the fact that there is no privileged way of ordering a given pair of intersecting space-like hyperplanes with respect to "before" and "after." These are strange consequences of the identification of times with space-like hyperplanes, but they are derivative upon the strangeness of Minkowski space-time itself.

⁵⁵ For simplicity, I assume that Albert occupies an *inertial* frame.

invariant. When restricted to a space-like hyperplane, it is formally equivalent to a metric function on three-dimensional Euclidean space.⁵⁶

As in Galilean space-time, we can associate ordinary persisting objects in Minkowski space-time with four-dimensional world paths. As before, we will call the region that results from intersecting the world path of a persisting object x with a given time (space-like hyperplane) t “the region x occupies at t .” We can therefore answer question (iii) just as we did before, with our Correspondence principle. In other words, in Minkowski space-time, as in Galilean space-time, the shape of a persisting object at a time is the shape of the region it occupies at that time. Now, the world path of an object x is not a frame-relative matter. (Observers in different reference frames might disagree on how to slice up its world path into “simultaneity slices,” but they will not disagree on which space-time points belong to the world path.) Nor, given some space-like hyperplane t , is it a frame relative matter which region x occupies at t . And crucially, it is not a frame-relative matter what shape that region has, at least inasmuch as the shape of a region is part of its objective physical structure. As we’ve seen, the shape of a three- (or fewer-) dimensional subregion of a space-like hyperplane is furnished by the Interval, an objective, frame-invariant magnitude. It follows, given our Correspondence principle and the identification of times with space-like hyperplanes, that *within Minkowski space-time*,

⁵⁶ The interval between space-time points p and q is represented in any Lorentz coordinate system by the equation $I(p,q) = \sqrt{(X(p) - X(q))^2 + (Y(p) - Y(q))^2 + (Z(p) - Z(q))^2 - (T(p) - T(q))^2}$. Here X , Y , Z , and T are functions mapping space-time points to coordinate values. The points constituting a space-like hyperplane are those that have the same T value in some Lorentz coordinate system. Where p and q have the same T value in a given coordinate system, the equation for the Interval between p and q reduces to $\sqrt{(X(p) - X(q))^2 + (Y(p) - Y(q))^2 + (Z(p) - Z(q))^2}$, the form of a metric function on three-dimensional Euclidean space.

the shape an object has at a time is not a frame-relative matter. An object's shape at a time does not involve any relations to reference frames, for it is a property that may be possessed by objects in a Galilean space-time, where the notion of a frame-relative shape has no meaningful application. I conclude that P1 in the above argument against RES is false.

Then what of the familiar claim that, in SR, the length of an object is frame-relative, or its corollary, that the shape of an object is frame-relative? It turns out that the alleged frame-relativity of shape in SR amounts to nothing more than the truism that objects have different shapes at different times, something we knew to be possible well before Einstein. We may illustrate this point by reference to Chalmers's Lorentz- Earth scenario. Albert measures (and experiences) O_1 to be a square. Twin Albert, relative to whom O_1 is traveling at 0.87 the speed of light, measures O_1 to be a 2:1 rectangle. On the above proposal, the "time of Albert's experience" is a space-like hyperplane t_1 on which Albert's experience falls and whose constituent points are all simultaneous relative to Albert's frame of reference. In the scenario we've described, the world path of O_1 will carve out a square-shaped region from t_1 —a two-dimensional four-sided region with equal angles and sides, as measured by the Interval. Given our Correspondence principle, this means that O_1 is square at t_1 . But the world path of O_1 cuts through other times besides t_1 . One such time, which partially overlaps t_1 , is t_2 : the space-like hyperplane on which Twin Albert's experience falls and whose constituent points are all simultaneous relative to Twin Albert's frame of reference. The region that O_1 's world path carves out

from t_2 is a 2:1 rectangle rather than a square. Given our Correspondence principle, this means that O_2 is a 2:1 rectangle at t_2 . Here we simply have a case of one thing having different shapes at different (albeit overlapping) times, something which happens in Newtonian worlds as well (though of course distinct times never overlap in Newtonian worlds). The shape of an object is frame-relative in SR only in the sense that the class of times one cares about (the class of times relative to which both experience and standard measuring equipment represents the shapes of things) is determined by one's frame of reference.

It follows from what has been said that Albert's and Twin Albert's experiences are veridical. As I've suggested, Albert's experience, considered in terms of its Edenic content, represents O_1 as being (absolutely) square at the time of the experience. As we've seen, O_1 is (absolutely) square at the time of his experience; it is square at the time of experience in the same sense in which an object might be square at a time in a classical Newtonian world. So it appears that his experience, considered in terms of its Edenic content, is veridical. The same holds, *mutatis mutandis*, for Twin Albert's experience. I conclude that Edenic shapes are just as much at home in our own relativistic world as they are in a classical Newtonian world. SR provides no reason to reject realism about Edenic shape.

4.9 The Argument from Phenomenal Relationism

I conclude this chapter with one further argument for realism, which I call the *argument from phenomenal relationism*. Like the argument (template) in §4.3, the argument from phenomenal relationism applies to every major class of sensible qualities.

Many have observed that the sensible qualities seem to stand in an intimate, constitutive relation to phenomenal properties. In particular—restricting our attention to sensory phenomenology—having an experience with a given phenomenal character seems to be a matter of standing in a certain relation to a sensible quality or complex of sensible qualities, a relation which might variously be called “sensory awareness,” “sensory acquaintance,” or “sensory representation.” More precisely, we can say that for any sensory phenomenal property P, there is some sensible quality (or complex of sensible qualities) Q such that what it is to have P is to be sensorily aware of Q. If we consider a highly specific or detailed phenomenal property, such as the complex phenomenal property that characterizes your total current sensory phenomenology, then the relevant instance of Q will be a rich complex of sensible qualities.⁵⁷ On the other hand, for highly general phenomenal properties, such as “phenomenal reddishness,” the relevant instance of Q may be an individual sensible quality, such as redness.

The view espoused here, that (sensory) phenomenology consists in a relation of phenomenal awareness to (something in the vicinity of) sensible qualities or sensible-

⁵⁷ Some will prefer an alternative on which these determinate phenomenal properties consist in a relation to structured propositional contents. So long as these propositional contents of experience are taken to involve sensible qualities, then such philosophers will agree that phenomenal properties constitutively involve relations to sensible qualities, which is the important point for my purposes.

quality complexes is sometimes called *relationism* about sensory phenomenology, and has a wide following among contemporary philosophers of mind (Campbell 1993, Johnston 2004, Chalmers 2006, Pautz 2009, Byrne 2009, Tye 2014). In my view, the relation in question is an intentional relation. By this I mean that when a subject stands in this relation to a quality, the subject, or her experience, may be said to (sensorily) “attribute” or “ascribe” that quality to the object of her experience—to (sensorily) represent the object as having that quality. The subject, or her experience, may then be described as *veridically* representing that object just in case the object has the quality in question (cf. Chalmers 2006, Pautz 2009, Tye 2014). For convenience I will presuppose this intentional conception of phenomenal awareness in what follows, but what I say should be adaptable to other relationist theories, e.g. versions of direct realism that construe the relation of phenomenal awareness in non-intentional terms.

Loosely speaking, relationism about sensory phenomenology (hereafter “phenomenal relationism”) is the converse of the relationist theories of sensible qualities discussed above. A relationist about color might hold, for example, that what it is to be red is to bear a certain relation to phenomenal reddishness, (e.g. the dyadic relation expressed by the open sentence “x is disposed to cause normal perceivers to [have experiences which] instantiate y in normal viewing conditions”), where phenomenal reddishness may be construed as an intrinsic phenomenal property of subjects (or their experiences). By contrast, a phenomenal relationist will hold that what it is for a subject to undergo phenomenal reddishness is to bear a certain relation to redness. Intuitively, it is phenomenal relationism that has things the right way around. Color (shape, loudness,

pitch...) seems to be intrinsic and non-relational, whereas color (shape, loudness, pitch . . .) *experience* seems essentially to relate us to objects and qualities beyond ourselves.

Many irrealists about the secondary qualities embrace phenomenal relationism (Chalmers 2006; Pautz 2011, ms.). But phenomenal relationism creates formidable difficulties for irrealism. For it apparently follows from irrealism that we do not bear any ordinary physical or functional relations to the secondary qualities. For example, we do not bear any causal or informational relations to them; our internal states do not “track” or causally co-vary with exemplifications of secondary qualities. So it is not clear where the irrealist about the secondary qualities would even begin in trying to give an explanation, whether reductive or non-reductive, of how we come to sensorily represent these qualities. By contrast, if realism is true and the secondary qualities are instantiated in the world more or less as they seem to be, then we bear various interesting physical, causal, and informational relations to them. We therefore have the materials to at least begin to venture explanations of how we come to sensorily represent a given secondary quality. For example, it is natural to suggest that we come to be aware of certain sensible qualities in virtue of standing in broadly informational, or informational-cum-teleological, relations to them (Millikan 1984, Fodor 1987, Dretske 1988, Tye 1995). Some simplistic suggestions, which are surely overly simplistic but are plausibly on the right track, include the relation expressed by the open sentence “x is undergoing some state or other which causally co-varies under normal conditions with exemplifications of

y” or “x is undergoing some state or other which has the biological function of indicating that y is exemplified.”

Although proposals of this kind are typically associated with reductive theories of sensory awareness, the points above hold whether one is aiming for a reductive explanation of our phenomenal awareness of sensible qualities or a non-reductive explanation. On a reductive view, we might simply identify the relation of phenomenal awareness with (some refinement of) one of the broadly informational relations mentioned above. On a non-reductive view, we might say that phenomenal awareness is grounded in, but not identical to, some such relation. Alternatively, dualist phenomenal relationists like Chalmers and Pautz might use some such relation R to formulate the psychophysical laws governing our phenomenal awareness of sensible qualities. It might be suggested, for example, that the fundamental psychophysical law governing sensory phenomenology takes something like the following form: for all individuals x and qualities q, x is phenomenally aware of q iff xRq .

Again, the above suggestions are obviously simplistic and in need of various refinements. The point is that, given realism about the sensible qualities, we stand in various natural relations to sensible qualities in terms of which sensory phenomenology might be explained, either reductively or non-reductively. We can therefore at least begin to develop proposals and subject them to progressive refinements. On the other hand, with an irrealist starting point, one has no idea even where to begin in offering an explanation, whether reductive or non-reductive, of how we come to be phenomenally acquainted with the various sensible qualities.

Kant (1783/2004) famously derided the realist view that sensory experience presents us with objects as they are in themselves, saying, “[I]t is incomprehensible how the intuition of a thing that is present should allow me to cognize it the way it is in itself, since its properties cannot migrate over into my power of representation” (p. 34). I hope the above discussion makes clear that this objection gets matters backwards. We can at least begin to comprehend how properties that feature in our environment, to which we therefore stand in various natural relations, might come to be sensorily represented. What is incomprehensible is rather how qualities to which we bear no natural relations at all should find their way into our “power of representation.”

Chapter 5: Russellian Monism about the Secondary Qualities

5.1 The Non-Reductive Realist's Dilemma

In the last chapter, I mentioned, and set aside, two objections to a non-reductive realist view of the secondary qualities. The first was the causal-exclusion objection, which runs roughly as follows: if realism about the secondary qualities is true, then the secondary qualities have causal relevance. If the secondary qualities have causal relevance, then they must be identical to physical properties, in which case reductionism is true. Therefore, if realism is true, then reductionism is false. Equivalently, non-reductive realism is false.

The second objection was the objection from simplicity: the non-reductive realist is committed to a significantly more metaphysically complex world than the reductionist or the irrealist. Considerations of parsimony therefore favor reductionism or irrealism over non-reductive realism.

As I mentioned in §4.6, it is not possible to give an adequate reply to these objections without answering what I'll call the *question of integration*: how do the secondary qualities relate to, or integrate into, physical reality? It is the primary task of this section develop and defend a particular answer to this question.

One initially attractive answer to the question of integration is given by the following thesis:

Qualitative Grounding: Instantiations of secondary qualities are grounded in instantiations of physical properties.

“Grounded in” here expresses the converse of the relation of metaphysical grounding, familiar from the writings of Kit Fine (2001), Jonathan Schaffer (2009), Gideon Rosen (2010), and others (cf. §3.5.2). Officially, I’ll take grounding to be a relation between a collection of facts (the facts doing the grounding) and an individual fact (the grounded fact). For definiteness, I will take facts to be true propositions, but other conceptions of facts (e.g. one on which facts are more like concrete states of affairs) would serve our purposes just as well. An “instantiation of” a property F should be understood as a fact consisting in something’s having F. In order to keep my sentences readable and down to a reasonable length, I’ll sometimes speak in terms of one *property* grounding another. But this should always be taken as an abbreviated way of saying that an *instantiation* of the one property grounds an *instantiation* of the other.

For our purposes, there are two important features of the grounding relation. First, grounding entails metaphysical necessitation. That is, if facts $A_1 \dots A_n$ ground a fact B, then it’s metaphysically necessary that if $A_1 \dots A_n$ obtain, then B obtains. It is therefore a consequence of Qualitative Grounding that the secondary qualities supervene with metaphysical necessity on physical properties. The second important fact about grounding is that grounding claims entail a corresponding “because” claim. That is, if $A_1 \dots A_n$ ground B, it follows that B obtains *because* $A_1 \dots A_n$ obtain. The relevant notion of grounding here is what’s sometimes called “full” or “total” grounding. It’s going to be convenient later on to have at our disposal a derivative notion of *partial* grounding. Partial grounding may be defined in terms of full grounding in the expected way, namely:

a fact *A* *partially* grounds a fact *B* iff *A* belongs to some collection of facts which fully ground *B*.

As others have noted, the objections from causal exclusion and simplicity do not have much bite against a version of non-reductive realism that endorses Qualitative Grounding. Qualitative Grounding defangs the objection from causal exclusion because, as many have observed (especially in response to analogous objections to non-reductive physicalist theories of mental properties), supervenient properties aren't generally excluded from causal relevance by subvening properties (Campbell 1993; Yablo 1992, 1995, Tye 2000; Byrne and Hilbert 2007). Consider the following case (identical in form but not in content to an example from Yablo 1995, pp. 486-7): John's body is so constituted that he always perspires when the room is 85 degrees or hotter—to abbreviate, when the room is "hot." The current temperature of the room is just over 85 degrees—to abbreviate, it is "barely hot." The room's being barely hot is nomically sufficient in the circumstances for John's perspiring. But the room's bare hotness does not exclude its hotness from causal relevance. Indeed, although we could explain John's perspiring by citing either property, if we had to pick one, we should cite the supervenient property of being hot, not the subvening property of being barely hot. For the former is, as Yablo says, "commensurate with the effect, in the sense of including what the effect needed with a minimum of irrelevant extras (p. 487). John would still have perspired if the room hadn't been barely hot, so long as it had still been hot. Likewise, if secondary qualities supervene on physical properties, then it would appear that the former, far from being "screened off" by the latter, are often in a better position to explain why objects

appear the way they do. As Campbell (1993) puts the point, although the microphysical properties of an object may be nomically sufficient in a given circumstance for its looking red, “the explanation in terms of redness adds modal data to a description of the physical sequence. It says that in nearby worlds in which the physical character of the thing was varied but its redness maintained, an experience of redness was still the upshot.” (p. 263, cf. Watkins 2005).

Qualitative Grounding also defangs the simplicity objection because, plausibly, supervenient properties do not increase the metaphysical complexity of a world in any sense of “metaphysical complexity” relevant to theory choice. Consider just about any property expressed by a predicate of English, e.g. the property of being a cat, the property of being a paint can, or the property of being clumsy. It is highly unlikely that we can give reductive analyses of any of these properties in the vocabulary of the physical sciences. A reductive definition of a property minimally specifies some non-trivial necessary and sufficient conditions for the instantiation of that property. And it should not come as news to philosophers of this generation that it is very difficult to give non-trivial necessary and sufficient conditions for the instantiation of just about any property, much less necessary and sufficient conditions that satisfy whatever other requirements there may be on reductive definitions. But of course, even if it turns out that there are no reductive definitions of such properties, realism about cats, paint cans, and clumsy individuals is not on that account metaphysically inflationary, at least not in any objectionable sense. Plausibly this is because these properties, while not reducible to physical properties, nonetheless are grounded in or supervene on physical properties.

Given that Qualitative Grounding defangs the objections from causal exclusion and simplicity, it is no surprise that non-reductive realists tend to endorse Qualitative Grounding (Campbell 1993, McGinn 1996, Watkins 2005, Yablo 1995). But there is a problem with Qualitative Grounding which non-reductive realists have not squarely faced: it seems to be false. As many have observed, there is a close analogy between secondary qualities and phenomenal properties on this point (cf. Johnston 1996, Byrne 2006, Shoemaker 2003, Campbell 1993). Secondary qualities resist integration into the “scientific image” of the world much as phenomenal consciousness does. The connection between the physical properties of an object and its secondary qualities has at least an *appearance* of contingency, much like the connection between the physical properties of a conscious organism and its phenomenal or experiential properties. Just as a complete physical description of a human perceiver, on an occasion of viewing a ripe tomato, seems to be consistent with a range of alternative hypotheses about her phenomenology (e.g. that she’s having a phenomenally reddish experience, that she’s having a phenomenally greenish experience, or that she’s having no experience at all), so too the complete physical description of a ripe tomato seems to be consistent with a range of alternative hypotheses about its color (e.g. that it’s red, that it’s green, or that it has no color at all).

Of course, the non-reductive realist can declare this appearance of contingency an *illusion* of contingency. But this blunt response is inadvisable, for at least two reasons. First, the principal advantage of non-reductive realism over its reductionist and irrealist rivals lies in its respect for appearances. Such a declaration would undermine the non-

reductive realist's theoretical advantage, especially relative to the irrealist, who after all can readily accept these apparent modal truths. Second, the relevant modal appearances seem to belong to the same broad class as the second-order appearances leveraged against the reductionist in chapters 2 and 3. Like the latter appearances, the former appearances pertain to the nature of the sensible qualities in some broad sense.⁵⁸ If these modal appearances are systematically deceptive, the broader class of appearances to which they belong are called into doubt, thus vitiating the case against reductionism.

Hence, the non-reductive realist faces a dilemma: accept Qualitative Grounding and systematically flout modal appearances, or deny Qualitative Grounding and fall prey to the objections from causal exclusion and simplicity. In the following section, I develop and defend a model of the relation between secondary qualities and the physical inspired by Russellian Monist views on the mind-body problem which resolves this dilemma for non-reductive realism.

5.2 Secondary Quality Russellian Monism

In this section, I begin by characterizing Russellian Monism, a view about the relationship between phenomenal consciousness and the physical inspired by ideas in Russell's *The Analysis of Matter* and other works. My characterization will not be overly concerned with the details of Russell's own view(s), but with articulating a distinctive position on the phenomenal-physical relation on which there is rough agreement among philosophers

⁵⁸ Indeed, the relevant class of apparent modal truths is probably best considered a subclass of the apparent second-order truths about the secondary qualities, as the former are equivalent to predications of modal properties to secondary qualities, e.g. "the color yellow is possibly uninstantiated by something with such-and-such physical property." (I.e. $\lambda F \diamond \exists y [y \text{ has such-and-such physical property and } y \text{ does not have } F] \text{Yellowness.}$)

in the Russellian Monist tradition (Eddington 1928, Maxwell 1979, Lockwood 1989, Stoljar 2001, Strawson 2006, Montero 2010, Chalmers forthcoming-b). After characterizing the Russellian Monist view of the phenomenal-physical relation, I introduce an analogous view of the relation between secondary qualities and the physical. Although the Russellian Monist framework was originally designed to account for the former relation, I argue that it is much better suited to account for the latter relation. Finally, I show how my proposed Russellian-Monist-inspired view resolves the dilemma described above for non-reductive realism.

At the foundation of Russellian Monism is the observation, due to Russell (1927), Eddington (1928), and others, that physics characterizes matter in terms of its relational structure, but does not reveal its intrinsic nature. In other words, physics characterizes the fundamental physical entities in terms of their relations to other things, including their spatiotemporal and causal relations to other things, but physics doesn't tell us how the fundamental physical entities are *intrinsically*. It's true that physics tells us that the fundamental physical entities have certain properties like mass and charge. But such attributions just amount to more relational information—in particular, information about their *causal* relations to other things. After all, it seems that the cash value of the claim that something has mass is just that it resists acceleration, attracts other massive things, and so forth. But this is just information about how it is disposed to *affect* and *be affected* by other things. We might put the point by saying that physics tells us about the *causal dispositions* of the fundamental physical entities, but does not tell us which intrinsic qualities serve as the *categorical bases* for these causal dispositions. Or, to put the point

within a somewhat different theoretical framework, we might say that physics tells us about the *second-order role properties* of the fundamental physical entities—for instance, that a given particle has *some property or other* that plays the “mass role,” a role which is characterized in great detail and mathematical precision—but it gives us no insight into which first-order properties *occupy* these causal roles.⁵⁹

Let’s use the term “quiddity” for these unknown qualities of matter. That is, the “quiddities” are those properties which serve as the categorical bases for the basic causal dispositions characterized by physics, or (alternatively) the properties that occupy the causal roles associated with the basic theoretical terms of physics. For now, let’s take for granted that there are quiddities. (Later we’ll return to the question of whether this is a warranted assumption.) If there are quiddities, the question arises: are quiddities *physical* properties? On the one hand, physics is supposed to be silent about them, which suggests that we should not count them as physical properties. On the other hand, they are properties of the fundamental physical entities and serve as the grounds for the dispositional/relational properties that physics attributes to the fundamental physical entities, which suggests that perhaps we should count them as physical properties.

The question seems to be verbal, not substantive. We can resolve it by simply distinguishing two senses of the term “physical property.” Following Chalmers’s (forthcoming-b) helpful terminology, we’ll say that physical properties in the narrow

⁵⁹ This foundational idea of Russellian Monism is closely related to what Frank Jackson (1998, p. 23) calls “Kantian physicalism”—“a large part (possibly all) of the intrinsic nature of our world is irretrievably beyond our reach”—and what David Lewis (2009) calls “Ramseyan humility”: “We are irremediably ignorant about the identities of the fundamental properties that figure in the actual realization of the true final theory” (p. 214).

sense, or *narrowly physical properties*, exclude quiddities, and include only the relational/structural properties in terms of which physics characterizes matter, including spatiotemporal properties/relations and the causal dispositions (or second-order role properties) associated with mass, charge, etc. And we'll say that physical properties in the broad sense, or *broadly physical properties*, include quiddities, as well as all the narrowly physical properties.

How is the idea that we are ignorant of the intrinsic, qualitative nature of matter relevant to the mind-body problem? Well, it has often been assumed that physical reality has an intrinsic nature different from and incongruous with the nature of consciousness and that the latter therefore cannot be a physical phenomenon. The Russellian Monist contends, to the contrary, that given the abstract and structural character of our knowledge of physical reality, this assumption is unfounded. As Russell (1949) says, "The physical world is only known as regards certain abstract features of its space-time structure—features which, because of their abstractness, do not suffice to show whether the physical world is, or is not, different in intrinsic character from the world of mind" (p. 240). Because physics doesn't tell us what the quiddities are like, it is tempting to suppose that the qualitative character of experience is somehow continuous with or grounded in the quiddities. As Chalmers (1996) says, echoing Russell (1927, pp. 382-93) and Eddington (1928, pp. 258-60):

There is only one class of intrinsic, non-relational property with which we have any direct familiarity, and that is the class of phenomenal properties. It is natural to speculate that there may be some relation or even overlap between the

uncharacterized intrinsic properties of physical reality, and the familiar properties of experience (pp. 153-4).

In this way, our ignorance of the nature of the quiddities has been taken to suggest the core claim of Russellian Monism: that the quiddities of which physics leaves us ignorant play an important role in the explanation of phenomenal consciousness.⁶⁰

In order to give a more precise definition of Russellian Monism, it will be helpful to introduce two versions of the “physicalist” thesis that phenomenal consciousness is grounded in the physical:

Narrow Phenomenal Grounding: Instantiations of phenomenal properties are grounded in instantiations of narrowly physical properties.

Broad Phenomenal Grounding: Instantiations of phenomenal properties are grounded in instantiations of broadly physical properties.

Note that because the narrowly physical properties are a subset of the broadly physical properties, Narrow Phenomenal Grounding entails Broad Phenomenal Grounding but not *vice versa*.

From here, we can define Russellian Monism (about phenomenal properties) as the conjunction of Broad Phenomenal Grounding with the denial of Narrow Phenomenal Grounding. That is, the Russellian Monist holds that instantiations of phenomenal properties are grounded in instantiations of broadly physical properties, but not (wholly)

⁶⁰ Cf. Pereboom’s (2011) characterization of Russellian Monism: “Russellian Monism is any view that combines [...] the claim that physics, or at least current physics, leaves us ignorant of certain categorical bases of physical dispositional properties, with [...] the proposal that these categorical properties have a significant role in explaining consciousness or experience” (p. 89).

in instantiations of narrowly physical properties. It is therefore a consequence of Russellian Monism that the quiddities play an essential role in the explanation of phenomenal consciousness.

Although it won't be included in the official definition of Russellian Monism, it is common for Russellian Monists to endorse something like the following:

Broad Phenomenal Scrutability: There is an intelligible or *a priori* connection between information about the distribution of broadly physical properties and information about the distribution of phenomenal properties.

It is a familiar point that an intelligible or *a priori* connection is wanting between the scientific description of a conscious organism and its phenomenal properties. The Russellian Monist has a natural diagnosis of this explanatory gap: the physical sciences only characterize things in terms of their narrowly physical features; as there is a genuine “ontological” gap between the the narrowly physical features of a conscious organism and its phenomenal properties, an explanatory gap here is to be expected. But there is no ontological gap between the *broadly* physical features of a conscious organism and its phenomenal properties, so we may reasonably suppose that there is an intelligible or *a priori* connection between them.

Hereafter I shall call the Russellian Monist position on the phenomenal-physical relation *Phenomenal Russellian Monism* (PRM), to distinguish it from the analogous position on the relation between secondary qualities and the physical, which I shall call *Secondary Quality Russellian Monism* (SQRM). In order to define SQRM, we'll need to

define counterparts to Narrow Phenomenal Grounding and Broad Phenomenal Grounding for the secondary qualities:

Narrow Qualitative Grounding: Instantiations of secondary qualities are grounded in instantiations of narrowly physical properties.

Broad Qualitative Grounding: Instantiations of secondary qualities are grounded in instantiations of broadly physical properties.

SQRM can then be defined as the conjunction of Broad Qualitative Grounding with the denial of Narrow Qualitative Grounding.

Again, although it won't be included in the official definition of SQRM, it is natural for the proponent of SQRM to endorse something like the following:

Broad Qualitative Scrutability: There is an intelligible or *a priori* connection between information about the distribution of broadly physical properties and information about the distribution of secondary qualities.

The apparent contingency discussed in §5.1 between an object's scientific description and its secondary qualities is closely related to the absence of an intelligible or *a priori* connection from the former to the latter. As before, the proponent of SQRM has a ready diagnosis of this explanatory gap: the physical sciences only reveal the narrowly physical features of objects. Between these and the secondary qualities there is a genuine contingency, an "ontological gap," so an apparent contingency or explanatory gap is to be expected. But there is no ontological gap between an object's *broadly* physical properties and its secondary qualities. We may therefore reasonably suppose that there is an intelligible or *a priori* connection between the broadly physical features of an object and

its secondary qualities. More precisely, since *a priori entailment* is not, strictly speaking, a relation between properties but a relation between properties-under-concepts (or modes of presentation), we should say that there is an *a priori* entailment between an object's broadly physical properties conceived under revelatory or nature-revealing concepts and its secondary qualities conceived in the standard way.⁶¹ Of course, *we* are not in possession of revelatory concepts of the quiddities. So we aren't in a position to apprehend the *a priori* connection that exists between the broadly physical features of an object and its secondary qualities, just as a creature lacking adequate geometric concepts could not apprehend the *a priori* connection that exists between the fine-grained geometric properties of a slab of marble and its smoothness. But if Broad Qualitative Scrutability is true, it would be clear to a Laplacian intelligence who had insight into the natures of the quiddities that any object with such-and-such broadly physical properties *must* have so-and-so secondary qualities.

5.3 The Non-reductive Realist's Dilemma Revisited

In §5.1 the non-reductive realist was presented with a dilemma: accept Qualitative Grounding and contradict modal appearances, or deny Qualitative Grounding and fall prey to the objections from causal exclusion and simplicity. SQRM resolves this dilemma by accepting Qualitative Grounding with respect to broadly physical properties (Broad

⁶¹ Chalmers (2005) characterizes a revelatory concept as, intuitively, "a property-concept such that possessing the concept puts one in a position to know (through a priori reflection) what the property is." Within the framework of epistemic two-dimensionalism, revelatory concepts will be *epistemically rigid* property-concepts, concepts whose primary intensions pick out the same property in each scenario, in contrast with epistemically non-rigid concepts such as *water*, whose primary intensions pick out different properties in different scenarios.

Qualitative Grounding) and denying Qualitative Grounding with respect to narrowly physical properties (Narrow Qualitative Grounding). That is, SQRM maintains that instantiations of secondary qualities are grounded in the distribution of broadly physical properties, but not (wholly) grounded in the distribution of narrowly physical properties. Plausibly, whatever apparent contingency exists between the physical properties of an object and its secondary qualities concerns its *narrowly* physical properties. As we have no insight into the natures of broadly physical properties beyond the horizon of the narrowly physical, we cannot say that broadly physical properties either seem to be, or seem not to be, modally independent of the secondary qualities. It is therefore only *Narrow Qualitative Grounding*, not *Broad Qualitative Grounding*, that conflicts with modal appearances. By denying Narrow Grounding, SQRM affirms the apparent modal truths.

Moreover, by accepting Broad Qualitative Grounding, SQRM overcomes the objections from causal exclusion and simplicity. As we've seen, properties that supervene on other causally efficacious properties are generally not excluded from causal relevance. Quiddities, if they exist, are causally efficacious *par excellence*. SQRM's commitment to Broad Qualitative Grounding entails that secondary qualities supervene on broadly physical properties, including quiddities, so epiphenomenalism about the secondary qualities does not threaten SQRM.⁶² SQRM also answers the simplicity objection: no

⁶² As many have observed, PRM avoids worries about the causal exclusion of phenomenal properties for the same sorts of reasons (Chalmers 1996, forthcoming-b; Stoljar 2001; Alter and Nagasawa 2012).

view that endorses Broad Qualitative Grounding may reasonably be accused of endorsing a metaphysically inflationary view of the secondary qualities.

Accepting SQRM therefore enables the non-reductive realist to resolve the dilemma raised in §5.1. Inasmuch as we have independent reason to accept non-reductive realism, as I have argued we do, this provides a reason to accept SQRM. But beyond its connections to non-reductive realism, SQRM is an interesting thesis in its own right. I shall therefore conclude this chapter by setting out several further more-or-less independent considerations in support of SQRM. PRM is one of the major contenders in the contemporary debate about consciousness. It is therefore somewhat surprising, given the tight parallels between phenomenal properties and the secondary qualities, that SQRM, the analogue to PRM, seems to have no proponents (other than myself) within the contemporary debate about color and other secondary qualities. In §§5.4-5.7, I hope to show not only that SQRM is a viable position, but that the sorts of considerations normally taken to support PRM actually do a better job supporting SQRM.

First, though, I want to return briefly to the question of whether there *are* quiddities. Quiddities are supposed to be the properties which serve as the categorical bases for the causal dispositions characterized by fundamental physics. But it's not clear that there are quiddities, so understood. After all, it's not obviously incoherent to suppose that the basic causal dispositions characterized by physics are "bare dispositions," dispositions without any categorical bases (Mumford 2006). Now, some have argued against the possibility of bare dispositions on general metaphysical grounds (Prior et al. 1982). For my part, I'm not sure if these attempts are successful. For this reason, I will

not *presuppose* that there are quiddities in what follows. To be sure, the conclusion for which I'll be arguing—SQRM—*entails* that there are quiddities. So we'll get our quiddities in the end, but not by begging the question against the opponent of quiddities. In other words, I will not invoke any premise in what follows such that our justification to believe it depends on our having *antecedently* accepted the existence of quiddities.

5.4 Considerations Favoring PRM Provide More Support for SQRM

It seems to me that the sorts of motivations that have led philosophers to accept PRM actually provide much stronger support for SQRM. As mentioned above, proponents of PRM are often motivated by something like the following thought: conscious experience has a mysterious “qualitative character” that is famously difficult to integrate into physical descriptions of reality. It is therefore natural to suppose that the intrinsic qualitative character of physical reality left unspecified by physics is somehow related to the qualitative character of experience.

However, the qualities which give a sensory experience its mysterious “qualitative character” are simply the sensible qualities. As argued in §2.2.3, these are not qualities *of* the experience (as the Qualia View supposes), but rather qualities the experience is an experience of—qualities *represented by* the experience. As Campbell (1993) puts the point: “the qualitative character of a color-experience is inherited from the qualitative character of the color” (p. 268). Similarly, Byrne (2006) writes, “If we like, we can say experiences of blue have a ‘qualitative character,’ but that is simply because they represent that objects have a ‘qualitative’ property—namely, blueness.” These remarks are closely related to the oft-made observation that experience is *transparent*: when one

goes looking for the supposed intrinsic qualities of one's experience of blue, for example, "one cannot help but see right through [the experience] so that what one actually ends up attending to is the real color blue" (Tye 1992, p. 160; cf. Harman 1990). If we are aware of no intrinsic qualities of experience, then *a fortiori*, we are aware of no intrinsic qualities of experience whose integration into physical reality poses a problem. If there are any qualitative properties in the vicinity of sensory experience whose integration into physical reality poses a problem, it is the sensible qualities, especially the secondary qualities.⁶³ So if the intrinsic qualitative character of physical reality left unspecified by physics is somehow related to the qualitative character of experience, this can only be because the former is somehow related to the sensible qualities. And the latter suggestion is most naturally implemented with SQRM.

In the passage quoted in §5.2, Chalmers motivates PRM with the claim that phenomenal properties are the *only* "class of intrinsic, non-relational property with which we have any direct familiarity." But this claim is doubly mistaken. First, the transparency observation suggests that the phenomenal properties of experience are *not* "intrinsic, non-relational properties." They seem rather to be relational properties. In particular, as suggested in §4.9, having a (sensory) experience with a given phenomenal character seems to be a matter of standing in a certain relation to a sensible quality or complex of sensible qualities, a relation which might variously be called "sensory awareness,"

⁶³ I do not mean to suggest that once we have shown how to integrate the sensible qualities into physical reality, the work of integrating phenomenal consciousness into physical reality will be finished. For there would remain the question of how we come to represent the sensible qualities in experience. But it is not clear how the Russellian Monist framework would help with this problem.

“sensory acquaintance,” or “sensory representation” (Campbell 1993, Johnston 2004, Chalmers 2006, Pautz 2009, Byrne 2009, Tye 2014). Second, there is another class of properties with which we have “direct familiarity,” which really do seem to be “intrinsic, non-relational properties”—namely, (at least some of) the sensible qualities. (Notably, in later writings, Chalmers comes around to endorse both of these claims [Chalmers 2006, 2012].)

Nonetheless, Chalmers and others who have voiced similar sentiments seem to be correct in thinking that the Russellian Monist framework is well suited to account for the intrinsic qualitative properties of macrophysical objects. Since phenomenal properties are not properties of this sort, and (many of) the secondary qualities are, it would appear that this framework is best suited to provide an account of the latter. It sometimes happens that the function a thing is destined to serve is not the function for which it was designed. So it is with the Russellian Monist framework, originally designed to handle the mind-body problem, but ultimately destined to handle the problem of the secondary qualities.

5.5 The Happy-Synthesis Argument

In this section, I argue that a fairly standard argument for PRM, which I call the *happy-synthesis argument*, can be straightforwardly adapted to support SQRM as well. I shall then proceed to argue in §5.5.1 that this sort of argument actually provides *more* support for SQRM than for PRM.

What I’m calling the happy-synthesis argument can be found in several recent papers on PRM, e.g. Chalmers (forthcoming-b), Stoljar (2001), and Alter and Nagasawa (2012). The basic idea is that PRM is supposed to be a happy-synthesis of standard forms

of physicalism and standard forms of dualism, retaining the strengths of each while also avoiding the weaknesses of each. To put it another way: it's said that the standard arguments in favor of physicalism, such as the causal exclusion argument, don't specifically support physicalism in its standard *narrow* form (what I earlier called Narrow Phenomenal Grounding), but rather support physicalism in its weaker *broad* form (what I earlier called Broad Phenomenal Grounding). On the other hand, it's said that the standard dualist arguments *against* physicalism, such as the conceivability argument or the knowledge argument, only tell against *narrow* physicalism; they don't tell against *broad* physicalism. That means we can respect all the standard considerations on each side simply by accepting physicalism in its broad form and denying physicalism in its narrow form, which gives us PRM as a happy synthesis.

To give you a feel for how this goes: the main argument for physicalism about phenomenal properties is the causal exclusion argument. We begin with the idea that phenomenal properties are causally relevant to what happens in the world, for example to how we behave. We then invoke some sort of causal-completeness-of-the-physical principle to motivate the claim that phenomenal properties can have causal relevance only if they are grounded in physical properties. The idea here is that physical properties are in some sense the *causally basic* properties, and that other causally relevant properties inherit causal relevance only by being grounded in physical properties. Finally, it's concluded that phenomenal properties are grounded in physical properties.

Here the Russellian Monist points out, rightly it seems, that these causal considerations don't get us the stronger thesis of *Narrow* Phenomenal Grounding, but

only the weaker thesis of *Broad* Phenomenal Grounding. As mentioned in §5.3, if there are quiddities, then quiddities are certainly among the “causally basic” properties. After all, quiddities are supposed to be the properties that play the fundamental causal roles characterized by physics. So properties grounded in quiddities ought to be able to inherit causal relevance just as well as properties grounded in narrowly physical properties. So it looks like considerations of causal relevance at best can establish physicalism in its weaker, broad form, not physicalism in its stronger, narrow form.

On the other side, consider some standard dualist argument against physicalism, such as the conceivability argument (Chalmers 1996). The conceivability argument runs roughly as follows: we begin with the claim that it’s coherently conceivable that there should be a physical duplicate of me who differs from me with respect to his phenomenal or experiential properties. It’s then said that *if* this is coherently conceivable, then it’s possible, and if it’s possible, it follows that my physical properties don’t necessitate, and therefore don’t ground, my phenomenal properties.

Here again, the Russellian Monist claims that these considerations plausibly only tell against *Narrow* Qualitative Grounding. After all, when we imaginatively hold fixed the physical properties of a human subject, presumably we’re holding fixed properties of the sort that the physical sciences reveal. But that does not include the quiddities. Since we don’t know what the quiddities are like, it’s arguable that we have no grounds for thinking that a *broadly* physical duplicate of me without my actual phenomenal properties is coherently imaginable.

Now it seems clear that these considerations can be adapted very straightforwardly to support SQRM as well. To begin with, as we saw in §4.6, there are causal exclusion arguments that apply to the secondary qualities running more-or-less exactly parallel to causal exclusion arguments in the philosophy of mind, and which (given the assumption of realism) can be used to support something like Qualitative Grounding. And as we've seen, for the reasons given above, it's plausible that these considerations at best establish the weaker thesis of Broad Qualitative Grounding, not the stronger thesis of Narrow Qualitative Grounding.

On the other side, as we saw in §5.1, there will be conceivability arguments against Qualitative Grounding that run perfectly parallel to conceivability arguments in the philosophy of mind. And for the reasons given in §5.3, it can be argued that these considerations only tell against Narrow Qualitative Grounding, but not against Broad Qualitative Grounding.

As with the happy-synthesis argument for PRM, we can respect the considerations on each side by accepting Broad Qualitative Grounding and rejecting Narrow Qualitative Grounding, which gives us SQRM as a happy synthesis.

5.5.1 Advantages of Happy-Synthesis Argument for SQRM

In fact, as I shall now argue, the happy-synthesis style of argument actually does a better job motivating SQRM than it does in motivating PRM. This is for two reasons: first, it seems that the causal considerations that support Broad Phenomenal Grounding provide even more support for Broad Qualitative Grounding. This is a consequence of the following rough and intuitive principle:

Principle of Increasing Confidence: The lower something is on the Great Chain of Being, the more confident we should be that its causally relevant properties are grounded in broadly physical properties.

Without taking the the idea of the Great Chain of Being with too much metaphysical seriousness, we can all nonetheless recognize a kind of hierarchy among beings—a (very) partial ordering where, generally speaking, animate beings outrank inanimate beings, sentient beings outrank non-sentient beings, rational beings outrank non-rational beings, and so on. The idea behind the Principle of Increasing Confidence principle is that, however confident we are that causally relevant properties of sentient beings, such as humans, are grounded in broadly physical properties, we should be even more confident that causally relevant properties of inanimate physical objects, like rocks, are grounded in broadly physical properties. After all, there’s a lot about humans and other sentient things that remains pretty mysterious. Even if we’re fairly confident that there’s nothing about humans that transcends the merely physical, we should be even more confident that there’s nothing about *rocks* that transcends the merely physical. Now, phenomenal properties seem to be peculiar to beings fairly high up on the Great Chain of Being. By contrast, if our senses can be trusted, the secondary qualities belong to things very low on the Great Chain of Being. We should therefore be even more confident that the secondary qualities are grounded in the broadly physical than we are about phenomenal properties.

The second reason why I think the considerations above provide more support for SQRM than for PRM is that the familiar arguments against Narrow Phenomenal

Grounding, such as the conceivability argument, threaten to undermine Broad Phenomenal Grounding as well, and thereby to undermine PRM. By contrast, the corresponding arguments against Narrow *Qualitative* Grounding aren't apt to undermine Broad Qualitative Grounding, and therefore aren't apt to undermine SQRM in the same way.

Suppose a proponent of PRM rejects Narrow Phenomenal Grounding on the basis of the conceivability argument. In other words, she holds that it's coherently conceivable that the narrowly physical facts be just as they actually are while the phenomenal facts differ, and on this basis she concludes that the phenomenal facts are not necessitated by, and thus not grounded in, the narrowly physical facts. Now whatever this conceivability claim ultimately comes to, at a minimum it involves *rejecting* the following thesis.

Narrow Phenomenal Scrutability: There is an intelligible or *a priori* connection between information about the distribution of narrowly physical properties and information about the distribution of phenomenal properties.

But if the proponent of PRM is going to rule out Narrow Phenomenal Grounding on the basis of such conceivability considerations, then she will want to hold that similar considerations do *not* tell against Broad Phenomenal Grounding. So while she rejects Narrow Phenomenal Scrutability, there will be pressure to *accept* the corresponding thesis about *broadly* physical properties introduced in §5.2, namely:

Broad Phenomenal Scrutability: There is an intelligible or *a priori* connection between information about the distribution of broadly physical properties and information about the distribution of phenomenal properties.

Of course, for similar reasons, the proponent of SQRM, at least if she wants to use conceivability arguments to rule out Narrow Qualitative Grounding, will want to make analogous claims. In particular, she'll want to reject the following thesis:

Narrow Qualitative Scrutability: There is an intelligible or *a priori* connection between information about the distribution of narrowly physical properties and information about the distribution of secondary qualities.

And for the reasons just given, she'll want to *accept* the corresponding thesis about broadly physical properties, which was also introduced in §5.2, namely:

Broad Qualitative Scrutability: There is an intelligible or *a priori* connection between information about the distribution of broadly physical properties and information about the distribution of secondary qualities.

For both kinds of Russellian Monists, the hope is that, even though we have reason to reject the relevant *narrow* scrutability theses, we don't have any reason to reject the relevant *broad* scrutability theses because physics doesn't tell us what the quiddities are like. But it seems to me that, although we don't have good reason to reject Broad *Qualitative* Scrutability, we do have principled reason to reject Broad *Phenomenal* Scrutability. That's because we know enough about phenomenal properties—or, better, our concepts of phenomenal properties—to rule it out. In particular, it seems to be a perfectly general feature of our concepts of phenomenal properties that, roughly speaking, information about the phenomenal properties of an individual *x* is never *a priori* entailed by information about things *distinct* from *x*, even if those things happen to compose *x*. Put more carefully, we can formulate the principle as follows:

Phenomenal Distinctness Principle: If x is distinct from each of the y s, then information about the phenomenal properties of x is not *a priori* entailed by information about the intrinsic qualities and structural organization of the y s (even given the information that the y s compose x).⁶⁴

Given phenomenal distinctness, it's going to follow that Broad Phenomenal Scrutability is false. That's because if Broad Phenomenal Scrutability is true, then it should turn out that information about the intrinsic qualities (quiddities) and structural organization (narrowly physical properties/relations) of the fundamental physical entities *a priori* entails that things *other* than the fundamental physical entities—macroscopic things like humans—have certain phenomenal properties. But that would be a violation of the Phenomenal Distinctness Principle. So it looks like we have principled reason to reject Broad Phenomenal Scrutability.

It's important to note that similar principles don't hold for *all* properties. For example, nothing like this is true for *geometric* properties. We can often work out the shape of a thing *a priori* given sufficient information about the shapes and structural organization of its parts. Given the information that x is composed of two squares of the same size lying on a single plane and pressed up against each other edge to edge, I can infer *a priori* that x is a 2:1 rectangle. The same plausibly holds for secondary qualities as well. Given the further information that those two square parts of x are red, I can

⁶⁴ Every more precisely: where $\Phi(x_1, x_2, \dots, x_n)$ is any open sentence expressing the concept of some n -place structural relation, $\Psi(y)$ is any open sentence expressing an ordinary introspective concept for a phenomenal property, and $M(x_1, x_2, \dots, x_n, y)$ is any open sentence expressing a purely mereological concept (e.g. *composition*), the following is not *a priori*: $\forall x_1, x_2, \dots, x_n, y ((\Phi(x_1, x_2, \dots, x_n) \& M(x_1, x_2, \dots, x_n, y)) \rightarrow \Psi(y))$

plausibly infer *a priori* that x is red. So this means that the corresponding principle about secondary qualities is plausibly false:

Secondary Quality Distinctness Principle: If x is distinct from each of the y s, then information about the *secondary qualities* of x is not *a priori* entailed by information about the intrinsic qualities and structural organization of the y s (even given the information that the y s compose x).

If the foregoing is correct, then it looks like the familiar arguments against Narrow Phenomenal Grounding threaten to undermine Russellian Monism about phenomenal properties by undermining Broad Phenomenal Grounding as well. On the other hand, the corresponding arguments against Narrow Qualitative Grounding aren't apt to undermine SQRM.

5.6 The Argument from Qualitativeness

The next argument for SQRM, which is loosely inspired by an argument from Chalmers (1996, pp. 118-20) known as the “structure-and-dynamics argument,” begins with the premise that the secondary qualities are *qualitative* properties. The relevant notion of a qualitative property is to be understood in contradistinction to the notion of a (purely) *structural* property. As we saw above, the Russellian Monist holds that the physical sciences tell us a great deal about the abstract structural features of physical reality, but do not reveal its intrinsic *qualitative* features. Let us try to make these notions a bit more precise. We'll say that a property is purely structural iff it admits of real definition in purely structural vocabulary. Following Chalmers (2012, forthcoming-b), we'll take “structural vocabulary” to include mathematical and logical vocabulary (including

devices for first- and second-order quantification), causal or nomic vocabulary, and also spatiotemporal vocabulary. (Narrowly physical properties are therefore purely structural.)

We can then say that a property is *qualitative* iff it is not purely structural.⁶⁵

The premise that the secondary qualities are qualitative properties has a great deal of intuitive support. Any analysis in purely structural terms of a secondary quality like sea-foam green, sweetness, or middle C would seem to leave something out. What it would leave out, one wants to say, is its intrinsic qualitative nature—what Chalmers (2006, p. 66) calls its “distinctive sensuous nature.”

Another route to the same conclusion comes from familiar considerations about “qualia.” Many have claimed that qualia are, or at least seem to be, qualitative in something like the sense at issue here (Jackson 1982, Block 1978, Chalmers 1996, Stoljar 2001). Qualia are supposed to be directly consciously accessible qualities that determine the phenomenal character of an experience. But as argued in §2.2.3, the only qualities associated with sensory experience that satisfy this description are *the sensible qualities represented by the experience*. The only quality in view—the only quality “before the mind”—when we make intuitive pronouncements about the quale associated with an experience of red (for example) is *the color red*. So any intuitive support there is for the

⁶⁵ There is another notion of a “qualitative property” in currency within metaphysics which is opposed primarily to the notion of an *object-involving* property (e.g. being next to Napoleon) and which includes structural properties. It is perhaps an odd result of my defining “qualitative” in contradistinction to “structural” that object-involving properties classify as qualitative. If one finds this result objectionable, nothing in what follows should be affected by amending the definition to require that qualitative properties (in the sense relevant here) be “qualitative properties” (in the sense opposed to object-involving properties).

view that color qualia are qualitative is really support for the view that *colors* are qualitative. The same holds true for other secondary qualities as well.

The next premise in our argument is the intuitively plausible principle that, when an object possesses a qualitative property, it never does so wholly in virtue of its purely structural properties. In slogan form: *no quality from structure*. It follows from this principle, together with our initial premise that the secondary qualities are qualitative, that instantiations of secondary qualities are *not* wholly grounded in instantiations of narrowly physical properties alone, since narrowly physical properties are purely structural. In other words, Narrow Qualitative Grounding is false. Given the reasonable assumption that the secondary qualities *are* wholly grounded at least in instantiations of *broadly* physical properties—Broad Qualitative Grounding—it follows that SQRM is true.

5.7 The Argument from Categoricity

The final argument for SQRM, which I call the *argument from categoricity*, runs as follows:

C1. If Narrow Qualitative Grounding is true, then instantiations of colors are at least partially grounded in instantiations of microphysical dispositions (i.e. the basic causal dispositions characterized by fundamental physics). (Premise)

C2. Colors are categorical properties. (Premise)

C3. No instantiation of a categorical property is grounded, even partially, in an instantiation of a disposition. (Premise)

C4. Therefore, Narrow Qualitative Grounding is false (C1-C3).

C5. Broad Qualitative Grounding is true. (Premise)

C6. Therefore, SQRM is true. (C3, C5)

Let us consider the premises of the argument, beginning with C1. C1 is very plausible. If Narrow Qualitative Grounding is true, it follows that instantiations of colors are grounded in instantiations of narrowly physical properties, which include spatiotemporal properties/relations together with microphysical dispositions. Assuming the proponent of Narrow Qualitative Grounding doesn't just want to ground color in spatiotemporal structure, she's going to have to say that certain microphysical dispositions, such as those associated with electromagnetic charge, for example, are going to play at least a *partial* role in grounding instantiations of colors.

Turning to C2: this premise claims that colors are *categorical properties*. As mentioned in §2.2, the notion of a categorical property is, intuitively, the notion of a wholly *non-modal* property, or alternatively, a property the instantiation of which does not constitutively depend on what goes on in other possible worlds. As urged in §2.2, colors seem to be categorical. Being red, like being round and unlike being fragile, is in some intuitive sense a situation wholly contained within actuality.

In order to make things a bit rigorous, it will be helpful to give a more precise analysis of the notion of a categorical property. To a first approximation, we want to say that F is a categorical property iff instantiations of F aren't grounded, even in part, in *modal facts*. Here we can identify the modal facts as those facts whose linguistic expression involves modal vocabulary, where modal vocabulary minimally includes the familiar unary modal operators "possibly" and "necessarily," as well as familiar binary

modal operators like the “would” counterfactual and the “might” counterfactual.

Officially, our analysis will be as follows:

F is a categorical property iff necessarily, instantiations of F are not weakly partially grounded in modal facts.

Here *weak* partial grounding is just the reflexive closure of partial grounding. The reason we want *weak* partial grounding rather than partial grounding is that, in the case where instantiations of F *are* modal facts, we’d want F not to count as a categorical property.

With this analysis in hand, we’re ready to move on to the third premise, C3. C3 says that no instantiation of a categorical property is grounded, even partially, in an instantiation of a disposition. Why think this? Well, it’s plausible that dispositional facts are *themselves* modal facts. Arguably, predicates for dispositions should be regarded as (paradigmatic!) modal expressions, in which case instantiations of dispositions will trivially count as modal facts. Another route to the conclusion that dispositional facts are modal facts is through the idea that dispositions can ultimately be analyzed in terms of subjunctive conditionals (Ryle 1949, Goodman 1954, Quine 1960). If dispositions can be analyzed in terms of subjunctive conditionals, it plausibly follows that dispositional facts are subjunctive facts. Since subjunctive facts are (paradigmatic) modal facts, it therefore follows that dispositional facts are modal facts.

Now, if dispositional facts are themselves modal facts, it will follow straightaway from our analysis of a categorical property that C3 is true. But perhaps you’re not sure whether dispositional facts should count as modal facts. Here’s an argument that C3 is true either way:

D1. If dispositional facts are modal facts, then C3 is true.

D2. If dispositional facts are not modal facts, then dispositional facts are grounded in modal facts.

D3. If dispositional facts are grounded in modal facts, then C3 is true.

D4. Therefore, (whether or not dispositional facts are modal facts) C3 is true.

As just mentioned, D1 follows from our analysis of categoricity. Why think D2 is true?

Well, if it weren't true, then given our analysis of categoricity, we'd have the result that at least some dispositions count as categorical properties. But I take it that one of the basic desiderata of a theory of categorical properties and dispositions is that it should turn out that categorical properties and dispositions form disjoint classes. To ensure that, we'll need to accept D2. Finally, D3 follows from our analysis of categoricity, together with the reasonable assumption that partial grounding is transitive (Fine 2001, Rosen 2010, Raven 2013). For in that case, if an instantiation of a categorical property were partially grounded in an instantiation of a disposition, then given the transitivity of partial grounding, it would follow that the categorical fact is grounded in subjunctive facts. And since subjunctive facts are modal facts, that again would violate our analysis of a categorical property. Therefore, whether or not we count dispositional facts as modal facts, we get the truth of C3.

C1-C3 entail the falsity of Narrow Qualitative Grounding. The final premise is C5, which asserts the truth of Broad Qualitative Grounding. This argument has nothing *new* to say in support of Broad Qualitative Grounding. But we can appeal to the considerations mentioned earlier in support of Broad Qualitative Grounding, e.g. causal

considerations or considerations of parsimony. Given Broad Qualitative Grounding, together with the falsity of Narrow Qualitative Grounding, SQRM follows.

5.8 Conclusion

I argued in chapters 2-4 that if we respect the first- and second-order appearances that constitute that fragment of the Manifest Image which pertains to the sensible qualities, we will accept realism about all the sensible qualities and non-reductive realism about the secondary qualities. But there is a *prima facie* tension between the Scientific Image and the non-reductive realist view of the secondary qualities embodied in the Manifest Image, a tension which finds its most forceful expression in the dilemma described in §5.1. As we've seen, SQRM, an independently motivated view about the relation between the secondary qualities and physical reality, resolves this dilemma. In doing so, SQRM resolves the apparent conflict between the Manifest Image and the Scientific Image without denying or diluting either. Through SQRM, we may hold the two images together in one stereoscopic view.

References

- Alter, T. & Y. Nagasawa (2012) "What is Russellian Monism?" *Journal of Consciousness Studies* 19 (9-10):67-95.
- Armstrong, D. (1968a) *A Materialist Theory of Mind*. London: Routledge.
- (1968b) "The Headless Woman Illusion and the Defense of Materialism." *Analysis* 29 (2): 48-9.
- (1978) *Universals and Scientific Realism: A Theory of Universals Vol. II*. Cambridge: Cambridge University Press.
- (1999) *The Mind-Body Problem: An Opinionated Introduction*. Boulder, CO: Westview Press.
- Asher, N. and M. Morreau (1995). "What Some Generic Sentences Mean." In G. Carlson and F. J. Pelletier. *The Generic Book*. Chicago: Chicago University Press, 300–339.
- Batty, C. (2010) "A Representational Account of Olfactory Experience." *Canadian Journal of Philosophy* 40 (4):511-538.
- Berkeley, G. (1710/1982). *A Treatise Concerning the Principles of Human Knowledge*. K. Winkler (ed.). Indianapolis: Hackett.
- (1713/1979). *Three Dialogues Between Hylas and Philonous*. R.M. Adams (ed.). Indianapolis: Hackett.
- Block, N. (1978) "Troubles with Functionalism." *Minnesota Studies in the Philosophy of Science* 9:261-325.
- (1990). "Inverted Earth." *Philosophical Perspectives* 4:53-79.

- Boghossian, P. & D. Velleman (1989) "Color as a Secondary Quality." *Mind* 98 (January): 81-103.
- BonJour, L. (1998). *In Defense of Pure Reason*. Cambridge: Cambridge University Press.
- Bradley, P. and M. Tye (2001) "Of Colors, Kestrels, Caterpillars, and Leaves." *Journal of Philosophy* 98 (9):469-487.
- Byrne, A. (2003) "Color and Similarity." *Philosophical and Phenomenological Research* 66 (3):641-65.
- (2006) "Color and the Mind Body Problem." *Dialectica* 60 (2):223-44.
- (2009) "Experience and Content." *Philosophical Quarterly* 59 (236): 429-451.
- Byrne, A. & D. Hilbert (1997) "Colors and Reflectances." In *Readings on Color, Volume 1: The Philosophy of Color*, eds. A. Byrne and D. R. Hilbert. Cambridge: MIT Press.
- (2003a) "Color Realism and Color Science." *Behavioral and Brain Sciences* 26 (1): 3-21.
- (2003b) "Color Realism Revisited." *Behavioral and Brain Sciences* 26 (6): 791-3.
- (2007a) "Color Primitivism." *Erkenntnis* 66: 77-105.
- (2007b) "Truest Blue." *Analysis* 67 (293): 87-92.
- Byrne, Alex & Michael Tye (2006). "Qualia Ain't in the Head." *Noûs* 40 (2):241-255.
- Campbell, J. (1993) "A Simple View of Colour." In J. Haldane & C. Wright (eds.), *Reality: Representation and Projection*, 257-68. Oxford: Oxford University Press.
- Chalmers, D. (1996) *The Conscious Mind: In Search of a Fundamental Theory*. Oxford: Oxford University Press.

- (2002). “Does Conceivability Entail Possibility?” In Tamar S. Gendler & John Hawthorne (eds.), *Conceivability and Possibility*. Oxford: Oxford University Press. 145-200.
 - (2005) “Revelation, Humility, and the Structure of the World.” <http://consc.net/powerpoint.html>.
 - (2006) “Perception and the Fall from Eden.” In T. Gendler and J. Hawthorne (eds.), *Perceptual Experience*. Oxford: Oxford University Press, 49-125.
 - (2012) *Constructing the World*. Oxford: Oxford University Press.
 - (forthcoming-a) “Three Puzzles about Spatial Experience.” In A. Pautz and D. Stoljar (eds.) *Themes from Block*. Cambridge, MA: MIT Press.
 - (forthcoming-b) “Panpsychism and Panprotopsychism.” In T. Alter and Y. Nagasawa (eds.), *Consciousness in the Physical World: Essays on Russellian Monism*. Oxford: Oxford University Press.
- Cohen, J. (2003) “On the Structural Properties of the Colours,” *Australasian Journal of Philosophy* 81: 78–95.
- (2009) *The Red and the Real: An Essay on Color Ontology*. Oxford: Oxford University Press.
- Cohen, S. (1984) “Justification and Truth.” *Philosophical Studies*. 46: 279–295.
- Correia, Fabrice (2010). “Grounding and Truth-Functions.” *Logique et Analyse/ 53* (211)251-79.
- Cutter, B. and M. Tye (2011). “Tracking Representationalism and the Painfulness of Pain.” *Philosophical Issues* 21 (1):90-109.

- Dancy, J. & C. Hookway (1986) "Two Conceptions of Moral Realism." *Aristotelian Society Supplementary Volume* 60: 167-205.
- Davidson, D. (1973). "Radical Interpretation." *Dialectica* 27 (1):314-328.
- Davies, M. (1993) "Aims and Claims of Externalist Arguments." *Philosophical Issues* 4: 227-49.
- Dennett, D. (1987) *The Intentional Stance*. Cambridge, MA: MIT Press.
- Descartes, René (1641/1998). *Discourse on Method and Meditations on First Philosophy*. (trans.) Donald Cress. Indianapolis: Hackett.
- Dorr, C. (2002). *The Simplicity of Everything*. Ph.D. thesis, Princeton University.
- Dretske, F. (1988) *Explaining Behavior: Reasons in a World of Causes*. Cambridge, MA: MIT Press.
- (1995) *Naturalizing the Mind*. Cambridge, MA: MIT Press.
- Eddington, A. (1928) *The Nature of the Physical World*. New York: Macmillan.
- Fodor, J. (1987) *Psychosemantics: The Problem of Meaning in the Philosophy of Mind*. Cambridge, MA: MIT Press.
- (1990) *A Theory of Content and Other Essays*. Cambridge, MA: Bradford Books/MIT Press.
- Fine, K. (2001) "The Question of Realism." *Philosopher's Imprint* 1 (2):1-30.
- Gettier, E. (1963) "Is Justified True Belief Knowledge?" *Analysis* 23 (6):121-123.
- Goldman, Alvin (1979). "What Is Justified Belief?" In G. Pappas (ed.), *Justification and Knowledge*. Dordrecht: Reidel.

- Goodman, N. (1954) *Fact, Fiction, and Forecast*. Cambridge: Harvard University Press.
- Gupta, A. (2006) *Empiricism and Experience*. Cambridge, MA: Harvard University Press.
- Hardin, C. L. (1988). *Color for Philosophers: Unweaving the Rainbow*. Indianapolis: Hackett.
- (1997). “Reinverting the Spectrum.” In *Readings on Color, Volume 1: The Philosophy of Color*, eds. A. Byrne and D. R. Hilbert. Cambridge: MIT Press.
- Harman, Gilbert (1990). “The Intrinsic Quality of Experience.” *Philosophical Perspectives* 4:31-52.
- Hartmann, W. (1997) *Signals, Sound, and Sensation*. New York: Springer.
- Hawthorne, J. & K. Kovakovich (2006). “Disjunctivism.” *Aristotelian Society Supplementary Volume* 80 (1):145-83.
- Hilbert, David (1987). *Color and color perception: A study in anthropocentric realism*. Center for the Study of Language and Information. CSLI.
- Hill, C. (2009) *Consciousness*. Cambridge: Cambridge University Press.
- Hospers, J. (1997). *An Introduction to Philosophical Analysis*, 4th Edition. London: Routledge.
- Huemer, M. (2001) *Skepticism and the Veil of Perception*. Lanham: Rowman & Littlefield.
- Hume, D. (1739/1978) *A Treatise of Human Nature*. L.A. Selby-Bigge and P.H. Nidditch (eds.). Oxford: Oxford University Press.

- Hurvich, L. and D. Jameson (1957). "An opponent-process theory of color vision."
Psychological Review 64 (6, Part I): 384-404.
- Husserl, E. (1900/2001). *Logical Investigations*, vol. 2. J.N. Findlay (trans.). New York: Routledge.
- Huxley, T.H. (1888) *Lay Sermons, Addresses, and Reviews*. New York: Appleton & Co.
- Jackson, F. (1977) *Perception: A Representative Theory*. Cambridge: Cambridge University Press.
- (1982) "Epiphenomenal Qualia." *Philosophical Quarterly* 32: 127-36.
- (1998) *From Metaphysics to Ethics*. Oxford: Oxford University Press.
- Jackson, F. & R. Pargetter (1987) "An Objectivist's Guide to Subjectivism about Colour."
Revue Internationale de Philosophie 41:127-41.
- Johnston, M. (1987) "Is There a Problem About Persistence?" *Proceedings of the Aristotelian Society* 61:107-135.
- (1992) "How to Speak of the Colors." *Philosophical Studies* 68 (3):221-63.
- (1996) "A Mind-Body Problem at the Surface of Objects." *Philosophical Issues* 7: 219-29.
- (2001) "The Authority of Affect." *Philosophy and Phenomenological Research* 63 (1): 181-214.
- (2004) "The Obscure Object of Hallucination." *Philosophical Studies* 120: 113-83.
- Kalderon, M. (2007) "Color Pluralism." *Philosophical Review* 116 (4):563-601.
- Kant, I. (1781/1999) *The Critique of Pure Reason*. P. Guyer and A. Wood (trans.). Cambridge: Cambridge University Press.

- (1783/2004) *Prolegomena to any Future Metaphysics*. G. Hatfield (trans.). Cambridge: Cambridge University Press.
- Kaplan, D. (1989) “Demonstratives.” In J. Perry, J. Perry, and H. Wettstein, *Themes from Kaplan*. Oxford: Oxford University Press, pp. 481–563.
- Kelly, T. (2008) “Common Sense as Evidence: Against Revisionary Ontology and Skepticism.” *Midwest Studies in Philosophy* 32 (1):53-78.
- (2010). “Peer Disagreement and Higher Order Evidence. In R. Feldman and T. Warfield (eds.) *Disagreement*. Oxford: Oxford University Press.
- Kim, J. (1993) *Supervenience and Mind*. Cambridge: Cambridge University Press.
- Kripke, S. (1980). *Naming and Necessity*. Oxford: Blackwell.
- Kroedel, T. (2012). “Counterfactuals and the Epistemology of Modality.” *Philosophers Imprint* 12 (12).
- Leslie, S.J. (2008) “Generics: Cognition and Acquisition.” *Philosophical Review* 117 (1): 1-47.
- Lewis, D. (1974). “Radical Interpretation.” *Synthese* 27 (July-August):331-344.
- (1983). “New Work for a Theory of Universals.” *Australasian Journal of Philosophy* 61: 343-77.
- Lewis (1986). *On the Plurality of Worlds*. Oxford: Blackwell.
- (1997) “Naming the Colours.” *Australasian Journal of Philosophy* 75 (3):325-42.
- (2001) “Redefining ‘Intrinsic’.” *Philosophy and Phenomenological Research* 63 (2): 381-3.
- (2002) “Tensing the Copula. *Mind* 111 (441):1-14.

- (2009) “Ramseyan Humility.” In D. Braddon-Mitchell & R. Nola (eds.), *Conceptual Analysis and Philosophical Naturalism*. Cambridge, MA: MIT/Bradford, 203-22.
- Liebesman, D. (2011). “Simple Generics.” *Noûs* 45 (3):409-442.
- Locke, J. (1689/1996) *An Essay Concerning Human Understanding*. K. Winkler (ed.). Indianapolis: Hackett.
- Lockwood, M. (1989) *Mind, Brain, and the Quantum*. Oxford: Oxford University Press.
- Lycan, W. (2000) “The slighting of smell,” in N. Bhushan and S. Rosenfeld (eds.), *Of Minds and Molecules: New Philosophical Perspectives on Chemistry..* Oxford: Oxford University Press, pp. 273-89.
- (2001a). “The Case for Phenomenal Externalism.” *Philosophical Perspectives* (15): 17-35.
- (2001b) “Moore Against the New Skeptics.” *Philosophical Studies* 103 (1):35-53.
- Maudlin, T. (2012) *Philosophy of Physics: Space and Time*. Princeton: Princeton University Press.
- Maxwell, G. (1979) “Rigid Designators and Mind-Brain Identity.” *Minnesota Studies in the Philosophy of Science* 9: 365-404.
- Mohan, M. (1988) “Biological Functions and Perceptual Content.” *Journal of Philosophy* 85: 5-27.
- Maud, B. (1995) *Colours: Their Nature and Representation*. New York: Cambridge University Press.
- Maxwell, G (1979). “Rigid Designators and Mind-Brain Identity.” *Minnesota Studies in the Philosophy of Science* 9: 365-404.

- Montero, B. (2010) "A Russellian Response to the Structural Argument Against Physicalism." *Journal of Consciousness Studies* 17 (3-4):70-83.
- McGinn, C. (1983) *The Subjective View: Secondary Qualities and Indexical Thoughts*. Oxford: Clarendon Press.
- (1996). "Another Look at Color." *Journal of Philosophy* 93 (11):537- 53.
- McLaughlin, B. (2003) "The Place of Color in Nature." in *Colour: Connecting the Mind to the Physical World*. R. Mausfeld & D. Heyer (eds.). Oxford: Oxford University Press, pp. 475-505.
- Millikan, R. (1984) *Language, Thought, and Other Biological Categories*. Cambridge, MA: MIT Press.
- (1989) "Biosemantics." *The Journal of Philosophy*. 86: 281-97.
- Mumford, S. (2006) "The Ungrounded Argument." *Synthese* 149 (3): 471-89.
- Noë, A. (2004) *Action in Perception*. Cambridge, MA: MIT Press.
- O'Callaghan, C. (2007) *Sounds: A Philosophical Theory*. Oxford: Oxford University Press.
- O'Shaughnessy, B. (2002) *Consciousness and the World*. Oxford: Oxford University Press.
- Papineau, D. (1993). *Philosophical Naturalism*. Oxford: Blackwell.
- Pasnau, R. (1999) "What is Sound?" *Philosophical Quarterly* 49: 309–324.
- Pautz, A. (2006a) "Sensory Awareness is not a Wide Physical Relation: An Empirical Argument Against Externalist Intentionalism." *Noûs* 40 (2): 205-40.

- (2006b) “Can the Physicalist Explain Colour Structure in Terms of Colour Experience?” *Australasian Journal of Philosophy* 84 (4):535-64.
 - (2009) “A Simple View of Consciousness.” In R. Koons & G. Bealer (eds.), *The Waning of Materialism: New Essays*. Oxford: Oxford University Press, 25-66.
 - (2010a) “Do Theories of Consciousness Rest on a Mistake?” *Philosophical Issues* 20 (1):333-67.
 - (2010b) “Review of Jonathan Cohen, *The Red and the Real: An Essay on Color Ontology*.” *Notre Dame Philosophical Reviews* 2010 (3).
 - (2011) “Can Disjunctivists Explain Our Access to the Sensible World?” *Philosophical Issues* 21 (1):384-433.
 - (ms.) “Color Eliminativism,” <https://webpace.utexas.edu/arp424/www/elim.html>.
- Peacocke, C. (1983) *Sense and Content*. Oxford: Oxford University Press.
- (1984) “Colour Concepts and Colour Experience.” *Synthese* 58 (March):365-82.
- Plato (2002). *Five Dialogues*. G.M.A. Grube (trans.). Hackett Publishing Company: Cambridge.
- Pereboom, D. (2011) *Consciousness and the Prospects of Physicalism*. Oxford: Oxford University Press.
- Plantinga, A. (1993) *Warrant and Proper Function*. Oxford: Oxford University Press.
- Pollock, J. (1984) “Reliability and Justified Belief.” *Canadian Journal of Philosophy*. 14: 103-114.
- Prior, E., R. Pargetter & F. Jackson (1982) “Three Theses About Dispositions.” *American Philosophical Quarterly* 19 (3):251-7.

- Pryor, J. (2000) "The Skeptic and the Dogmatist." *Noûs* 34 (4):517–549.
- Quine, W.V.O. (1960) *Word and Object*. Cambridge, MA: MIT Press.
- Reid, T. (1764/1983) *Inquiry and Essays*. R. Bleanblossom & K. Lehrer (eds.).
Indianapolis: Hackett.
- Rosen, G. (2010) "Metaphysical Dependence: Grounding and Reduction." In B. Hale and
A. Hoffman (eds.), *Modality: Metaphysics, Logic, and Epistemology*. Oxford: Oxford
University Press, 109-36.
- Russell, B. (1927) *The Analysis of Matter*. London: Kegan Paul.
- (1948) *Human Knowledge: Its Scope and Limits*. London: Routledge.
- Ryle, G. (1949) *The Concept of Mind*. London: Penguin.
- Salmon, N. (1986). *Frege's Puzzle*. Cambridge, MA: MIT Press/Bradford Books.
- Schaffer, J. (2009) "On What Grounds What." In David Manley, David Chalmers & Ryan
Wasserman (eds.), *Metametaphysics: New Essays on the Foundations of Ontology*.
Oxford: Oxford University Press, 347-83.
- Schwitzgebel, E. (2006) "Do Things Look Flat?" *Philosophy and Phenomenological
Research* 72 (3):589-599.
- Sellars, W. (1963) "Philosophy and the Scientific Image of Man." In *Science, Perception
and Reality*. London: Routledge and Kegan Paul Ltd.
- Shoemaker, S. (1994) "Phenomenal Character", *Noûs*, 28: 21–38.
- (2003) "Content, Character, and Color." *Philosophical Issues* 13 (1):253-78.

- Sider, T. (2003) "Reductive Theories of Modality." In M.J. Loux & D.W. Zimmerman (eds.), *The Oxford Handbook of Metaphysics*. Oxford: Oxford University Press. 180-208.
- (2013) "Against Parthood." *Oxford Studies in Metaphysics* 8:237–293.
- Smart, J.J.C. (1975) "On Some Criticisms of a Physicalist Theory of Colors." In C. Cheng (ed.), *Philosophical Aspects of the Mind-Body Problem*. University Press of Hawaii.
- Smith, A.D. (1990). "Of Primary and Secondary Qualities." *Philosophical Review* 99 (2) 221-254.
- Soames, Scott (1987). "Direct Reference, Propositional Attitudes, and Semantic Content." *Philosophical Topics* 15 (1):47-87.
- Sober, E. (1982) "Why Logically Equivalent Predicates May Pick Out Different Properties." *American Philosophical Quarterly* 19 (2):183-189.
- Stalnaker, R. (1984) *Inquiry*. Cambridge: Cambridge University Press.
- Stoljar, D. (2001) "Two Conceptions of the Physical." *Philosophy and Phenomenological Research* 62 (2):253-81.
- Strawson, G. (2006) "Realistic Monism: Why Physicalism Entails Panpsychism." *Journal of Consciousness Studies* 13 (10-11):3-31.
- Thompson, B. (2010) "The Spatial Content of Experience." *Philosophy and Phenomenological Research* 81 (1):146-184.
- Travis, C. (2004) "The Silence of the Senses." *Mind* 113 (449):57-94.

- Tye, M. (1992) "Visual Qualia and Visual Content." In T. Crane (ed.), *The Contents of Experience*. Cambridge: Cambridge University Press.
- (1995) *Ten Problems of Consciousness: A Representational Theory of the Phenomenal Mind*. Cambridge, MA: MIT Press.
- (2000) *Consciousness, Color, and Content*. Cambridge, MA: MIT/Bradford.
- (2009) *Consciousness Revisited: Materialism without Phenomenal Concepts*. Cambridge: MIT Press.
- (2012) "Cohen and Color Relationism." *Analytic Philosophy* 53 (3):297-305.
- (2014) "Transparency, Qualia Realism, and Representationalism." *Philosophical Studies* 170 (1):39-57.
- Watkins, M. (2005) "Seeing Red: The Metaphysics of Colours without the Physics." *Australasian Journal of Philosophy* 83 (1):33-52.
- White, R. (2006) "Problems for Dogmatism." *Philosophical Studies* 131 (3):525--57.
- Wiggins, D. (1987) *A Sensible Subjectivism?*. Oxford: Oxford University Press.
- Williamson, T. (2000) *Knowledge and its Limits*. Oxford: Oxford University Press.
- (2008) *The Philosophy of Philosophy*. Malden, MA: Blackwell.
- Wright, C. (1988). "Moral Values, Projection and Secondary Qualities." *Aristotelian Society Supplementary Volume* 62:1 - 26.
- (2004) "Warrant for Nothing (and Foundations for Free)?" *Aristotelian Society Supplementary Volume* 78 (1):167–212.
- Yablo, S. (1992) "Mental Causation." *Philosophical Review* 101 (2):245-80.

- (1993). “Is Conceivability a Guide to Possibility?” *Philosophy and Phenomenological Research* 53 (1):1-42.
- (1995) “Singling Out Properties.” *Philosophical Perspectives* 9: 477-502.