RUSSELLIAN ACQUAINTANCE AND PHENOMENAL CONCEPTS

A DISSERTATION
SUBMITTED TO THE DEPARTMENT OF PHILOSOPHY
AND THE COMMITTEE ON GRADUATE STUDIES
OF STANFORD UNIVERSITY
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

BY
DONOVAN ERIC WISHON
MARCH, 2012
I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.

John Perry, Primary Adviser

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.

Krista Lawlor, Co-Adviser

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.

Mark Crimmins

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.

Dagfinn Foellesdal

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.

David Beisecker

Approved for the Stanford University Committee on Graduate Studies.

Patricia J. Gumport, Vice Provost Graduate Education

This signature page was generated electronically upon submission of this dissertation in electronic format. An original signed hard copy of the signature page is on file in University Archives.
ABSTRACT

This dissertation develops a novel acquaintance-based version of the phenomenal concepts strategy for defending neurobiological accounts of consciousness against recent dualist arguments, and in the course of doing so, it explores the role that acquaintance plays in singular reference, cognition, and perception.

Chapter 1 provides a careful exposition of Bertrand Russell’s crucial notion of “acquaintance”, or direct conscious awareness. It argues that acquaintance is an experiential relation that grounds singular reference and that is more fundamental than, and is an enabling condition for, our knowledge of truths about things. But contrary to widespread philosophical interpretation, Russell does not hold that acquaintance safeguards a subject from misidentifying the objects of his or her acquaintance. Nor does it reveal the essential nature of such objects or provide subjects with a foundation of certain or infallible knowledge on which all of our ordinary and scientific knowledge of the world rests. Chapter 1 also shows how these widespread misconceptions about acquaintance arise from standard misreading of Russell’s famous “On Denoting”.

Chapter 2 argues that if we view acquaintance relations as the right sort of direct information relations, then we can avoid problematic naïve realist and disjunctivist theories of perceptual experience. It also advances “the problem of phenomenal indiscriminability” for disjunctivism, which challenges it to provide an epistemologically and metaphysically adequate account of non-veridical perceptual experience. The apparent inability of disjunctivism to do so suggests that it cannot fully respect “the explanatory role of experience”. Chapter 2 also explains how the informational character of perceptual experience allows us to be direct realists about the content of perceptual
experience. Lastly, it maintains that perceptual experience has a “two-faced presentational character” which allows us to hold on to the idea that the subjective, qualitative character of perceptual experience is a feature of the experiential state itself.

Chapter 3 presents the five main recent dualist arguments against physicalism: Chalmers’ zombie argument, Jackson’s knowledge argument, Kripke’s modal argument, Levine’s explanatory gap argument, and Chalmers’ two-dimensional argument. It also explains how the phenomenal concepts strategy promises to show why there is a distinctive epistemic gap between consciousness and the physical, but no corresponding ontological gap. Lastly, it argues against popular demonstrative-recognitional and constitutional models of phenomenal concepts by showing that neither one adequately captures our epistemic situation with respect to consciousness.

Finally, Chapter 4 contends that phenomenal concepts are special concepts of neurobiological/experiential states with a distinctive cognitive role: their reference is fixed by an inner demonstration to states to which we stand in special causal-informational relations—acquaintance relations—where these special relations are tied to special epistemic methods and capacities for picking up information about the states to which we are related and that provide us with distinctive conceptions of them. But these special acquaintance relations do not have epistemic or semantic features that threaten neurobiological accounts of consciousness. They do, however, explain why dualist intuitions seem powerful, even though they are mistaken. The main upshot is that we can (at least in principle) reconcile our everyday conceptions of ourselves as conscious beings with our scientific understanding of ourselves as complex biological organisms (and ultimately as complex physical systems).
ACKNOWLEDGEMENTS

Bertrand Russell once said that “the good life is one inspired by love and guided by knowledge”. In writing this dissertation, I have done my best to lead such a good life. And while I cannot say for sure I have achieved any great success in attaining genuine knowledge about conscious experience, I can honestly say that I strove with all of my mental might to squeeze out what insights I could about such matters. In contrast, I have no doubt whatsoever that my time writing this dissertation was inspired by love. Thus, I dedicate this dissertation to my beautiful, wonderful wife Christy, whose love, patience, friendship, encouragement, and support gave me the strength and focus to see this project to its completion. Whatever I have accomplished in the pursuit of knowledge herein is as much her achievement as mine, though of course every shortcoming is mine alone.

I want to offer my deepest thanks to my advisors and mentors John Perry and Krista Lawlor. A graduate student is lucky to find even one advisor who makes an impact on their life in the way that both have made an impact on me. They showed me more patience, kindness, encouragement, and dedication than I had ever hoped to find in an advisor, let alone in two. I would not be the philosopher I am today without their insightful feedback, sage advice, sincere generosity, and (seemingly) limitless energy. I hope that I will one day be half the philosopher, teacher, mentor, and person that each of them is. I will be forever thankful for having the true fortune of calling both of them my friends. I also owe a great deal of gratitude to Dagfinn Føllesdal, Mark Crimmins, and especially David Beisecker. I have learned an immense amount from them, and my dissertation would be vastly weaker without their insightful comments, daunting challenges, and wise guidance—not to mention the often lengthy discussions I had with
each of them on an impressive range of philosophical topics. David gets special thanks for first introducing me to the deep mysteries of consciousness and for stoking my passion to get to the bottom of them. I honestly could not have hoped for a better dissertation committee.

In addition, I want to thank the following people for providing feedback about, comments on, discussions about, and/or encouragement regarding various parts of this dissertation: John Searle, John Campbell, Todd Jones, Ken Taylor, David Hills, Paul Skokowski, Thomas Ryckman, Anna-Sara Malmgren, Michael Friedman, Graciela De Pierris, Russell Wahl, Bernard Linsky, Ian Proops, Kepa Korta, Galen Strawson, Michelle Montague, David Papineau, Michael Tye, Bruce Glymour, Christopher Hom, David Chalmers, Eric Schwitzgebel, Dan Giberman, Luis Cheng-Guajardo, Matthew Darmalingum, Wes Holliday, Ralf Bader, Eliot Michaelson, and Dikran Karagueuzian. I also want to thank Stanford University and my many friends, peers, and students there. I would also like to thank the audiences of the 2011 Bertrand Russell Society Annual Meeting, my Fourth-Year Talk at Stanford, my talk for the Center for the Explanation of Consciousness and the “Interdisciplinary Approaches to Consciousness” Geballe Workshop at Stanford, and my talk for “Perry Fest” here at Stanford. On each occasion, I received invaluable feedback on my project.

I am also very grateful to my family for their constant love and support. In particular, I’d like to thank my mother Sharon, stepfather Ray, brother Jeremiah, mother-in-law Vicky, father-in-law Tom, cats Socrates and Osiris, ferret Loki, and especially my very silly son Dylan for helping to keep me (mostly) sane while in graduate school.

Lastly, I’d like to thank Bertrand Russell for inspiring me in so many ways…
# TABLE OF CONTENTS

**ABSTRACT** ........................................................................................................ iv

**ACKNOWLEDGEMENTS** .................................................................................... vi

**INTRODUCTION**

§1 ANTECEDENT PHYSICALISM .............................................................................. 1

§2 CONSCIOUSNESS AND THE REEMERGENCE OF DUALISM ...................... 4

§3 RUSSELLIAN ACQUAINTANCE ....................................................................... 10

§4 ACQUAINTANCE, REFERENCE, AND COGNITION .................................. 13

§5 PERCEPTUAL ACQUAINTANCE ..................................................................... 17

§6 RUSSELLIAN ACQUAINTANCE AND PHENOMENAL CONCEPTS.. 20

**CHAPTER 1: RUSSELLIAN ACQUAINTANCE WITHOUT DISCRIMINATING KNOWLEDGE**

§1 INTRODUCTION ............................................................................................... 24

§2 THE RECEIVED INTERPRETATION ............................................................... 29

§3 RUSSELL ON PHENOMENAL CONTINUA CASES ....................................... 36

§4 KNOWLEDGE OF THINGS AND KNOWLEDGE OF TRUTHS ............... 41

§5 RECONSIDERING “ON DENOTING” ............................................................... 55

§6 CONCLUSION ................................................................................................. 65
CHAPTER 4: RUSSELLIAN ACQUAINTANCE AND THE
PHENOMENAL CONCEPTS STRATEGY

§1  INTRODUCTION................................................................. 176
§2  THE ACQUAINTANCE MODEL OF PHENOMENAL CONCEPTS.... 179
§3  THE ZOMBIE ARGUMENT....................................................... 200
§4  THE KNOWLEDGE ARGUMENT.............................................. 207
§5  THE MODAL ARGUMENT......................................................... 218
§6  THE EXPLANATORY GAP....................................................... 231
§7  THE TWO-DIMENSIONAL ARGUMENT.................................... 243
§8  CONCLUSION....................................................................... 261

REFERENCES.......................................................................... 264
INTRODUCTION

“The move from mere common sense to antecedent physicalism consists, then, of supposing that the subjective characters of our experiences are physical states of the brain. This is a supposition of an identity between types or kinds of events. The first is a type of event we are aware of when we have an experience… The second is a physical type, of the sort that can be in theory physically observed…It is not a piece of dogma but an antecedently plausible hypothesis. Our job is to see if the neo-Dualist arguments give us reason to abandon it.” – John Perry, Knowledge, Possibility, and Consciousness

§1 ANTECEDENT PHYSICALISM

In recent decades, the scientific and philosophical study of conscious experience has (re)emerged as one of the most important theoretical endeavors of our time. At the same time, scientists and philosophers alike have become reminded of just how deep and perplexing many of the questions surrounding the nature and role of conscious experience are. Many regard the deepest of these problems to be the reconciliation of what our current best scientific theories tell us about the nature and function of conscious experience and how conscious experience manifestly seems to be from the subjective point-of-view of creatures like us who are fortunate enough to have it. John Searle (1984) has succinctly stated the central problem:

We have a certain commonsense picture of ourselves as human beings which is very hard to square with our overall ‘scientific’ conception of the physical world. We think of ourselves as conscious, free, mindful, rational agents in a world that science tells us consists entirely of mindless, meaningless physical particles. Now, how can we square the two conceptions? How, for example, can it be the case that the world contains nothing but unconscious physical particles, and yet that it also contains consciousness? How can a mechanical universe contain intentionalistic human beings – that is, human beings that can represent

1 Perry (2001a), pp. 64-5.
the world to themselves? How, in short, can an essentially meaningless world contain meanings?²

When the problem is put in such stark terms as these, it becomes very easy to see just how difficult of a problem we are dealing with.

In fact, many philosophers and scientists of various stripes have despaired of ever arriving at a satisfactory reconciliation of our scientific picture of ourselves and reality with our ordinary, commonsense picture of ourselves and reality.³ So they invariably choose to favor one picture or the other as getting at the real nature of things. In other words, they typically either embrace a purely scientific understanding of the world and reject our commonsense picture of ourselves and the world as wildly off-the-mark, or they resist the completeness of the scientific picture of the world and hold onto the idea that there is something special about us (and about our conscious mental lives in particular) that fundamentally sets us apart from all of the rest of the world. But whichever way they choose, they nevertheless tend to agree that one or the other conception of ourselves and the world simply cannot be right.

Still, while it is likely that both our current scientific and current commonsense pictures of ourselves and the world contain a good deal of falsity and incompleteness, I not only believe that the central cores of each are indeed compatible with one another, I think that there is good reason to think that they are both true (or at least no reason for thinking either one is utterly false). Like a number of other bona fide Identity Theorists about the nature of conscious experience (and the self), including John Perry (2001a), John Searle (1984, 2004), Galen Strawson (1994, 1998, 1999, 2008), Herbert Feigl

---

³ Following Sellars (1962), we might call these the “Scientific Image” and the “Manifest Image” of the world.
(1958/1967), Grover Maxwell (1978), and (most famously) the later Bertrand Russell (1927/1954, 1927/1979, 1948/1992, 1959), I believe that in approaching the scientific and philosophical study of conscious experience, we ought to adopt the standpoint of what Perry (2001a) calls “Antecedent Physicalism”.

Antecedent Physicalism is the view that we have good independent reasons for thinking that all of concrete reality, including mental reality, is ultimately physical in nature, so we should only give up physicalism about conscious experience if we have compelling reasons to do so. At the same time, Antecedent Physicalism demands that we take seriously the core of our commonsense picture of conscious experience. Combining these two aspects together, we can characterize Antecedent Physicalism as holding that the philosophical and scientific study of conscious experience ought to have a prima facie commitment to the truth of physicalism and to the compatibility of our commonsense conception of conscious experience with a scientific, theoretical account of our biological (and ultimately physical) natures.

One strong reason for taking physicalism as our initial starting point for the philosophical and scientific study of the conscious experience is that it is the currently dominant assumption of scientifically-informed commonsense. The history of philosophical and scientific study is one replete with instances in which previously mysterious phenomena have turned out to be physical in nature (and wholly explainable by means of scientific theory), so there is at least some plausible inductive grounds from which to infer that conscious experience will turn out to be physical as well. What’s more, physicalism promises to offer a simple and unified picture of reality and the laws which govern it, and so is quite attractive from the standpoint of “Ockham’s Razor”. Less
abstractly, physicalism also promises a simple and elegant explanation of the systematic correlations we take to hold between various kinds of mental phenomenal and neurobiological states or processes in the brain (and central nervous system). Together, such considerations render physicalism at least prima facie plausible.

We also have good (yet defeasible) grounds for taking physicalism as our default position for the philosophical and scientific study of conscious experience based on causal considerations. From the side of reflective common sense, we have overwhelming reason to think that conscious experience (like other mental phenomena) is *causally efficacious*; it has some physical occurrences as causes, and other physical occurrences as effects. From the side of scientific theory, we have good reason to think that physical reality is a *causally-closed system*; if a physical occurrence has a cause, then that cause is itself a physical occurrence. And when we combine these two views together with the quite plausible assumption that there is no systematic causal redundancy in the workings of the world, we are led straightforwardly to the conclusion that conscious experience *just is* a kind of physical occurrence. So we once again have some reason to think some version or other of physicalism about conscious experience is true.

§2 CONSCIOUSNESS AND THE REEMREGENCE OF DUALISM

On the other hand, for a good number of reasons, it is extremely hard to see how physicalism could be true, given the plain facts of conscious experience. One initial hurdle we must cross in order to make any significant progress in the philosophical and scientific investigation of conscious experience is to disentangle the various different phenomena that are often lumped together by our everyday, commonsense notions of
“consciousness” and “conscious experience”. Sometimes the term “consciousness” is used to distinguish those subjective states of mind a subject has when he or she is in some sense *awake* (or dreaming) from those in which the subject is completely unconscious. Other times the term “consciousness” is used to capture the difference between mental representations which are “broadcast for free use in reasoning and for direct ‘rational’ control of action (including reporting)” and those that are not. Similarly, the term “consciousness” is used to describe those internal states of a system that play the right kind of role in mediating informational inputs and behavioral outputs. Some (often of a more mystical persuasion) even use the term “consciousness” for a being’s special states of “higher-awareness”.

While there is certainly much of scientific and philosophical interest in each of these notions of “consciousness”, I believe that the deepest and most interesting theoretical problems about “consciousness” are those concerning the *sentient awareness* of certain complex biological systems. Furthermore, I think that we can usefully tease apart the two aspects of sentient awareness which pose the most interesting and challenging theoretical problems. First, there is the remarkable fact that the states of sentient awareness of certain complex biological systems have subjective, qualitative characters (or “phenomenal properties”) such that there is (figuratively) “something it’s like” for the system to be in them. And second, there is an almost equally remarkable fact that states with such subjective, qualitative characters play an important (and perhaps even fundamental) role in grounding the ability of such complex biological systems to

---

7 Nagel (1974)
pick up information about, act on, and communicate about the broader world in which they are situated. If physicalism is true, both of these remarkable aspects of sentient awareness are ultimately physical in nature, and so presumably our physical theories eventually ought to provide us with some insight as to how this could be.

However, in recent years, there has been a significant resurgence of philosophers who argue that there is no way of consistently reconciling these features of sentient awareness with the physicalist picture of reality.\footnote{Chalmers (1996, 2002, 2010).} While they grant that the natural sciences provide us with ever-increasing understanding of the biological structure, functional organization (for information-processing and motor-control), and ultimate physical constitution of the brain (and central nervous system), they argue that even if our physical theories were extended to their limits and filled-in with as much detail as one would like, they would still shed no real light on the subjective, qualitative characters of states of sentient awareness. At best, the physical sciences can, together with our introspective knowledge of our own conscious experiences, provide us with knowledge of the systematic correlations between such-and-such neurobiological states of the brain and our experiences with such-and-such subjective, qualitative characters. But the subjective, qualitative characters of our experiences will never be fully integrated into a single, unified, physicalist picture of reality. Rather, they will be properly seen as a completely distinct aspect of reality which, for one reason or another, happens to be systematically linked to (and perhaps even dependent on) the physical domain by means of contingent psycho-physical laws of nature.

This new wave of such Dualist philosophers has relied on a number of related arguments to make a remarkably seductive case that physicalism is falsified by its
inability to provide a satisfactory account of the subjective, qualitative characters of our conscious experience. For instance, Chalmers (1996, 2002, 2010) has argued that we can conceive of complex biological systems which are physically like us in every respect yet which lack conscious experience entirely, and that the right sort of conceivability of such circumstances shows that such beings are possible and so physicalism is false. Jackson (1982) has argued against physicalism on the grounds that one could evidently know all the physical facts about the world without being able to deduce a priori “what it’s like” to actually have such-and-such experiences. Kripke (1980) has argued that in some cases of empirically discovered identities, we have a lingering sense that the relation between the identified phenomena is contingent despite the fact that identity is a necessary relation; but in the case of psycho-physical identities, we cannot explain away this sense of contingency like we can do in standard cases of empirically discovered identities, or in any other way. Chalmers (2002, 2010) has used Levine’s (1983, 2001) contention that conscious experience seemingly eludes explanation on the basis of any of the explanatory models of the natural sciences to make the case that the very nature of physical explanation itself rules out the possibility of incorporating conscious experience into a physicalist picture of reality. Finally, Chalmers (1996, 2002, 2010) has argued that proper reflection on the “two-dimensional” semantic character of our concepts of the physical and of conscious experience show us why the conceivability of physical duplicates of us without our experience entails their possibility, and thus the falsity of physicalism.

Interestingly, all of these Dualist arguments have the same basic structure. They begin with an epistemic premise which purports that there is some gap in our understanding of the relation between conscious experience and the physical domain.
Next, they argue that the particular character of the gap in our understanding of the relevant psycho-physical relations is indicative of an underlying *ontological* gap between conscious experience and the physical domain. Finally, they straightforwardly conclude from the distinction in the fundamental nature of conscious experience and physical reality that physicalism is false.

Unsurprisingly, physicalists of every sort vehemently disagree with the conclusions of arguments of this kind and so object to one or the other (or occasionally both) of the premises that lead to it. And the great majority of them have (rightly) leveled complaints about the second premise in which the Dualists allege that the particular character of the epistemic premise has ontological implications. However, I believe that there is good reason for thinking that none of these physicalist responses to the recent Dualist arguments is entirely adequate: they simply haven’t located the core problem shared by all of the Dualist arguments.

In particular, I believe that insufficient attention has been directed towards the (often hidden) role that “knowledge by acquaintance” plays in the various arguments against physicalism. Roughly speaking, the notion of “knowledge by acquaintance” is that of an especially intimate, direct epistemic relation between a conscious subject and objects or features of the world. It was first introduced as a philosophical term-of-art by John Grote (1865), but it is most often associated with Bertrand Russell, who developed the notion in great detail and deployed it towards solving a good number of problems in

---

9 McGinn (2003) is one notable exception. He (rightly) argues that our special acquaintance with the subjective, qualitative characters of our conscious experiences lies at the very heart of mind-body problem. However, he has radically mistaken views about both the nature of acquaintance itself and Russell’s understanding of it.
the philosophy of mind, epistemology, the philosophy of language, the philosophy of mathematics, logic, and even moral psychology and philosophy.10

Historically, Russell’s notion of knowledge by acquaintance has been (wrongly) associated with a number of extremely robust epistemic theses. For instance, many have taken acquaintance to reveal the essential nature of its objects to a subject. Others take acquaintance to disclose every part or aspect of its objects to a subject. Still others take acquaintance to provide subjects with infallible knowledge about its objects which can serve as the secure foundation for all of our knowledge about the world. Most often, acquaintance is seen as providing a subject with especially secure knowledge of the reference of our thought and talk, knowledge which safeguards a subject from the possibility of misidentifying the objects of his or her thought and talk, which are most often taken to be items in the subject’s own conscious mental life.

I believe that the fundamental reason why Dualists take the distinctive epistemic gap between our understanding of conscious experience and our scientific understanding of the physical domain to have ontological import is that they believe that our acquaintance with our own conscious experiences has some (or even all) of these robust epistemic features. Certainly, if one were to think our acquaintance with our own conscious mental lives had any of these epistemic features, it would be quite easy to conclude that our conscious mental lives are wholly distinct from anything we might reasonably expect the physical sciences to tell us about the world. And in fact, many Dualists are quite explicit both about their acceptance of a notion of knowledge by

acquaintance along these lines and about the central role in plays in their arguments for dualism.

Nevertheless, Dualists radically misunderstand both Russell’s notion of acquaintance and our own epistemic situation with respect to conscious experience. Hence, the main aim of this dissertation is to provide the Antecedent Physicalist with a proper understanding of the nature of our acquaintance with our own conscious experiences (and the rest of the physical world) so that he or she can steadfastly resist the seductive allure of the recent arguments for dualism. In fact, I argue that once we have a proper understanding of the nature of acquaintance and of its role in our thought and talk, we can see that a Russellian notion of acquaintance provides the Antecedent Physicalist with a powerful tool for defending physicalism against the threat of dualism, contrary to the received wisdom of many.

§3 RUSSELLIAN ACQUAINTANCE

One reason for why Russell’s notion of knowledge by acquaintance has typically been wrongly seen as helpful for defenses of dualism rather than physicalism is that no one has carried out an adequate exposition of what Russell actually says about the nature and role of acquaintance through the complete corpus of his writings on it. In Chapter 1 of this dissertation, I try to correct this scholarly shortcoming by presenting a careful and detailed examination of Russell’s own views about acquaintance.

I begin Chapter 1 by arguing that many important theorists, including Gareth Evans (1983), Donald Davidson (1989/2001), John Campbell (2009), and Ian Proops (2011), mistakenly attribute to Russell what I call “the received view of acquaintance”: 
the view that acquaintance safeguards us from misidentifying the objects of our acquaintance. Most often, the received view of acquaintance is attributed to Russell in order to make sense of how his famous Theory of Descriptions (and Descriptive Theory of Ordinary Names) might possibly serve as a generalized solution to “Frege’s Puzzle” and related substitution problems involving co-designating concepts or expressions which he at least seems to be concerned with answering.

Next, I draw on long-underappreciated remarks of Russell’s from his 1914 *Our Knowledge of the External World* (and elsewhere) concerning *phenomenal continua* cases which I believe decisively show that he does not accept the received view of acquaintance, or any epistemic thesis which purports that acquaintance provides us with infallible or revelatory knowledge about its objects. Phenomenal continua are cases in which one quality transitions into another quality such that no difference can be discriminated between two distinct points A and B in the transition, nor between two distinct points B and C, and yet some difference can be discriminated between points A and C. What they show is that our capacities for introspective discrimination are outrun by changes in the items or qualities that we are experiencing such that we might erroneously conclude that there are no differences or boundaries between them. But then even in the case of the subjective, qualitative characters of our own conscious experiences, acquaintance does not secure us against the possibility of misidentifying the objects of our thought and talk. Nor does it provide us with revelatory knowledge of the essential nature of our conscious mental lives.

Moreover, I also show that the possibility of misidentifying the objects of acquaintance follows from underappreciated aspects of Russell’s overall theory of
knowledge and acquaintance. At its core, Russell’s notion of acquaintance is of an
experiential relation that grounds singular reference and that is more fundamental than,
and is an enabling condition for, our knowledge of truths about things. However, it does
not all by itself provide us with any knowledge of truths about things, and nothing
ensures that we will end up with genuine knowledge of the nature or identity of things
from our more basic acquaintance with them.

More specifically, Russell’s basic picture of cognition and reference is this: There
is some special class of objects (concrete and abstract alike) and their features with which
we have experiential contact. On the basis of this experiential contact, we are in a position
to direct conscious attention to them. Our conscious attention to them in turn grounds our
ability to make demonstrative, singular reference to them and puts us in a position to
acquire knowledge of such demonstrative, singular reference. Furthermore, our conscious
attention to experienced objects and features also puts us in a position to pick up
information about, and thereby form some conception of, them. Finally, on the basis of
our conceptions of these objects and features and our more general conception of how the
world works, we can designate objects and features with which we lack experiential
contact. We do so by employing representations that encode identifying conditions that
an object or feature must uniquely satisfy in order to be designated by them, or by
employing representations that are purely quantificational in character.

Finally, I close Chapter 1 by considering the radical impact that Russell’s actual
views on acquaintance have for our understanding of his well-known George IV case in
“On Denoting”. In particular, I argue that Russell’s treatment of the George IV case is not
a one-size-fits-all solution to “Frege’s Puzzle” and provides no support for the received
view of acquaintance. One thing that this shows is many philosophers have misunderstood what Russell was up to in “On Denoting”.

§ 4 ACQUAINTANCE, REFERENCE, AND COGNITION

Though I accept much (and perhaps even most) of what Russell actually says about the nature and role of acquaintance, I also think that the Antecedent Physicalist must go crucially beyond Russell’s own views. For one thing, if I am correct, then in “On Denoting” Russell is not in fact trying to provide a solution to “Frege’s Puzzle”. But any adequate theory of reference and cognition must have some way of addressing it (and similar puzzles such as Kripke’s “Pierre” and “Paderewski” cases).11

For another, Russell’s own solution to the problem of reference-failure is to restrict acquaintance to special items (such as sense-data, abstract objects, and perhaps the self) for which we have the highest assurance of their existence (or “subsistence” in the case of abstract objects). But it is far from clear that an Antecedent Physicalist should take such a strategy. On the contrary, there is good reason for thinking that we are in fact acquainted with many parts to the external physical world, and Russell’s own strategy does little to help with reference-failure regarding it. Fortunately, recent theoretical advancements provide us with many alternative ways of explaining the semantics of our “empty” thought and talk.

What’s more, Russell’s own account of the role of acquaintance in reference and cognition ignores the very real social dimension to our thought and talk. Hence, the Antecedent Physicalist will have to incorporate many elements of recent theoretical work on the semantic and pragmatic dimensions of interpersonal communicative acts. Finally,

11 Kripke (1979).
if the Antecedent Physicalist is to make legitimate use of the notion of acquaintance, he or she must clearly provide some naturalistic account of it, one which is consistent with his or her ontological commitments.

Given its actual epistemic features and role in reference and cognition, I believe that the most promising way forward is for the Antecedent Physicalist to see acquaintance as the right sort of direct informational relations between a phenomenally conscious signal (i.e. a representational vehicle in a subject’s overall representational system) and some corresponding aspect of the world. In other words, a conscious subject is, on my view, acquainted with a thing if and only if he or she has a phenomenally conscious state the occurrence of which puts him or her in direct original (in roughly John Searle’s sense of “original”) informational contact with that thing.\(^\text{12}\)

What’s more, I think that the best account of what it is to be in direct original informational contact with a thing is provided by Perry’s (2002) naturalistic theory of information and informational content. According to Perry, information is “what one part or aspect of the universe (the signal) shows about some other part or aspect [of the universe] (the subject matter)” given the particular circumstances of the signal’s occurrence and given the way the world actually works.\(^\text{13}\) Informational content, in contrast, is what one part or aspect of the universe (the signal) shows about some part or aspect of the universe (the subject matter), if any, given the particular circumstances of the signal’s occurrence and assuming some (possibly false) constraints on the way the world works. Original informational content is what one part or aspect of the universe (the signal) shows about some part of aspect of the universe, if any, given the particular

\(^{12}\) Searle (1983).

circumstances of the signal’s occurrence, given some (possibly false) constraints on how
the world works, and given the natural job of the signal, where the natural jobs of signals
are typically determined by their selective, ontogenetic, and social histories, together with
their subjective, qualitative characters, if any.

Modifying Perry’s framework ever-so-slightly, my proposal is that we are in the
right sort of direct original informational contact with a thing, and thereby acquainted
with it, if and only if we have a phenomenally conscious signal bearing original
informational content directly about that thing. More schematically, I hold that:

We are acquainted with something \( P \) on the basis of having a phenomenal
conscious signal \( S \) if and only if given (possibly false) constraints \( C \),
circumstances \( F \), and \( S \)’s natural job \( N \), \( S \) occurs only if, and because, \( P \).

Moreover, I basically hold that our direct singular thought about, and direct singular
reference to, things is enabled by our having experiences whose original informational
content is directly about them.

On my view, the distinctive feature of a singular concept is that its semantic
contribution at the level of “what is thought” or “what is said” to the contents, or
satisfaction conditions, of our thought and talk is constituted by the very objects or
features, if any, that the device is about rather than by some way of getting at or
identifying them. Typically, singular concepts are also rigid designators: their semantic
contributions to the representational content of our thought and talk are invariant across
modal contexts. Many rigidly-designating singular concepts are also directly referential.
My view is that directly referential concepts are ones whose contributions to the
representational contents of our thought and talk are determined by mechanisms other
than the representational fit of that which is designated to some grasped conception of, or way of getting at, it. What’s more, the subject who deploys a directly referential concept in thought or talk needn’t have a reflective grasp of the (often external) mechanisms which determine its semantic contribution to the contents of our thought and talk. Rather, the subject need only be attuned to such reference-determining mechanisms.

However, I also think that the social nature of communication allows subjects to pass along linguistic (and other communicative) devices which allow others, who lack direct acquaintance with the thing for which the device has the job of standing, to likewise make singular or directly singular reference with them in their thought and talk. In this way, social communication extends the domain of the objects to which we can make singular and directly singular reference beyond the limits of our own personal acquaintance, at least so long as someone’s direct acquaintance with the referent was involved in the initial reference-fixing of the device at the far end of our “communicative chains”. In addition, I am inclined to believe that we can use our store of singular and directly singular referential concepts more-or-less at will to construct denoting or quantificational complexes whereby we can designate things in a less direct and/or non-singular manner (and we can seemingly rigidify them or use them to introduce “descriptive names” for things as well). ¹⁴

Finally, I think that once we understand acquaintance as a matter of having a phenomenally conscious representational vehicle with the right sort of direct original informational content with aspects of the world, we draw on recent work by Perry

¹⁴ Put in terms of the recent debates in Jeshion (2010), I am more-or-less an “acquaintance theorist” with respect to direct singular thought and talk, but a “semantic instrumentalist” with respect to indirect singular thought and talk. But one crucial insight we can draw from Russell’s views is that our direct singular thought in more fundamental than, and an enabling condition for, our derivative indirect thought and talk about things, including our indirect singular thought and talk about them.
Recanati (2010), and many others to provide adequate accounts of what happens in cases
of reference-failure and misidentification.

§5 PERCEPTUAL ACQUAINTANCE
I also think that many longstanding problems in the philosophy of perception can be dealt
with once we view acquaintance relations as the right sort of direct informational
relations between a phenomenally conscious signal and aspects of the world. And so I
turn to these issues in Chapter 2 of this dissertation.

I begin Chapter 2 by noting that recent work on the nature of perceptual
acquaintance, such as that of Campbell (2002a, 2002b, 2009) and Tye (2009), has tended
to treat naïve realism and indirect realism as an exhaustive disjunction of possible views.
Roughly speaking, naïve realism in the view that what we are directly aware of in
veridical perceptual experience are the objects and features of an external perceptual
scene, together with the view that its qualitative character inheres in the external
perceptual scene itself. Indirect realism, on the other hand, is the view that what we are
directly aware of in perceptual experience are the subjective, qualitative characters of our
own inner sensations, together with the view that our “perception” of objects and features
in an external perceptual scene is a matter of implicit or explicit inference.

Because they are well-aware of the deep problems raised by indirect realism about
perception, Campbell and Tye have developed their accounts of perceptual acquaintance
along the lines of naïve realism. On the other hand, they are equally well-aware of the
challenges naïve realism faces from illusions, hallucinations, and the currently dominate views about perceptual experience which tend to arise from the scientific study of perception. For this reason, Campbell and Tye both adopt their own versions of Disjunctivism about perception, a family of currently-popular views according to which there is no “common factor” between veridical perceptual experiences and non-veridical perception-like experiences.

After briefly outlining the various key features of Campbell’s and Tye’s Disjunctivist views about perceptual acquaintance and perception-like experience, I contend that no such theory can provide an adequate account of our perceptual experience of things. More specifically, I argue that there is a problem of what I call “phenomenal indiscriminability” for Disjunctivism to which there is no satisfactory reply. The basic problem is that if we cannot introspectively tell the difference between our veridical perceptual experiences of things from entirely distinct non-veridical perception-like experiences, there must be some reason why we cannot tell them apart. And the most obvious explanation for why we can’t tell them apart is that the qualitative character of each is extremely similar. But the Disjunctivist has no plausible account of where in the world to locate the qualitative characters of our non-veridical perception-like experiences, or of how we become aware of them.

What’s more, their inability to provide a metaphysically and epistemologically satisfactory account of non-veridical perception-like experience puts pressure on what Campbell (2002a, 2002b) calls “the explanatory role of experience”. On one hand, an account of perceptual experience must explain how it is that experience provides us with our knowledge of demonstrative reference and our conceptions of objects and features as
mind-independent and *as* categorical in nature (as opposed to as a mere cluster of “affordances”). But as I argue, it must also provide a satisfactory explanation of why we form false beliefs about demonstrative reference and mistaken conceptions of things when all does not go well with perceptual experience. However, it is far from clear how Disjunctivism can respect this demand.

In the second half of Chapter 2, I turn to my own information-based account of perceptual experience, which has the resources to help us avoid both indirect realism and naïve realism (including its Disjunctivist versions). I argue that perceptual experience has a “two-faced presentational character” which allows us to be *direct realists* about the *content* of perceptual experience while holding onto the idea that the *subjective, qualitative characters* of perceptual experiences are features of our own inner perceptual states. We are “phenomenally” aware of the “self-presenting” subjective, qualitative characters of our perceptual experiences simply by having them. And we are “intentionally” aware of objects and features of the external perceptual scene by being attuned to what the occurrence of our perceptual experiences tells us about the external perceptual scene (i.e. to their direct informational content). Consequently, we are all at once acquainted with (and non-inferentially aware of) both the subjective, qualitative characters of our own inner perceptual states and the objects and features of the external perceptual scene, albeit in very different ways. Moreover, I believe that only this kind of “two-faced” view of perceptual acquaintance can adequately respect the full explanatory role of experience.
Finally, in Chapters 3 and 4, I take on the recent Dualist arguments against Antecedent Physicalism by means of my own novel acquaintance-based account of our special first-personal “phenomenal concepts” of the subjective, qualitative characters of our conscious experiences. I begin Chapter 3 by presenting in greater detail the five main recent Dualist arguments against physicalism: Chalmers’ zombie argument, Jackson’s knowledge argument, Kripke’s modal argument, Levine’s explanatory gap argument, and Chalmers’ two-dimensional argument. In doing so, I try to point out the role that an overly robust notion of knowledge by acquaintance is playing in each one.

Next, I provide a general characterization of the family of physicalist responses, those falling under the heading of the “phenomenal concepts strategy”, which maintain that the distinctive epistemic gap between our understanding of conscious experience and our scientific understanding of the physical domain is the result of our having two very different kinds of concepts of them: phenomenal concepts and physical-theoretical concepts. Nevertheless, many physicalists (rightly) argue, both kinds of concepts rigidly-designate one and the same aspects of reality, so there is no ontological gap corresponding to the distinctive epistemic gap.

In the rest of Chapter 3, I present and evaluate the two leading accounts of our special phenomenal concepts in the current philosophical literature. Some, like Loar (1990), Perry (2001a), and Levin (2007), think that phenomenal concepts behave in crucial respects like demonstrative or recognitional concepts. I call this the “demonstrative-recognitional model” of phenomenal concepts. Others, like Papineau (2002), Block (2007), Chalmers (2003, 2010), and Balog (Forthcoming), maintain that
they are special concepts which are partly constituted by the very phenomenal properties for which they stand. I call this the “constitutional model” of phenomenal concepts.

While I grant that there is much to be said in favor of each model, I argue that neither is entirely satisfactory. On one hand, the demonstrative-recognitional model is too thin to explain adequately our epistemic situation with respect to conscious experience. On the other hand, the constitutional model makes the relation between conscious experience and our thought about it too intimate.

In Chapter 4, I advance my own acquaintance-based account of phenomenal concepts which avoids the shortcomings of its competitors. I begin the chapter by pointing out that several well-known recent attempts to defend physicalism by means of a notion of knowledge by acquaintance, such as Churchland (1985, 1989), Conee (1994) and Tye (2009), fail because they do not develop a corresponding account of phenomenal concepts. Next, I take stock of the many important lessons about acquaintance, reference and cognition, perceptual experience, and phenomenal concepts that we learned in the previous three chapters. Bringing these crucial insights together, I present my own acquaintance-based account on the nature of phenomenal concepts (and of physical-theoretical concepts). Finally, I use it to systematically contest in new, powerful ways each of the five leading Dualist arguments against physicalism, emphasizing throughout how widespread mistaken conceptions of the nature and role of acquaintance have lent them false credibility.

My principal claim in Chapter 4 is that our phenomenal concepts are concepts of neurobiological/experiential states with a distinctive cognitive role; their reference is fixed by an inner demonstration to states to which we stand in special causal-
informational relations (i.e. acquaintance relations), where these special relations are tied to special epistemic methods and capacities for picking up information about the states to which we are related and that provide us with distinctive conceptions of them. Nevertheless, the special acquaintance relations that we bear to our own states of conscious awareness (those that ultimately ground our special phenomenal concepts) do not have epistemic or semantic features that threaten neurobiological accounts of consciousness. They do, however, explain why there is a distinctive epistemic gap between our conceptions of conscious experience and physical reality (and so why Dualist intuitions seem powerful) even though there is no reason for supposing that there is a corresponding ontological gap between them. Hence, we can (at least in principle) reconcile our everyday conceptions of ourselves as conscious beings with our scientific understanding of ourselves as complex biological organisms (and ultimately as complex physical systems). In this way, the potent resources of my acquaintance-based account of phenomenal concepts show that the Antecedent Physicalist has nothing to fear from the recent seductive, but ultimately unconvincing, Dualist challenges to it.

There are, however, a few things that my dissertation will not do. First, and most importantly, it does not attempt to solve the hard problem of consciousness itself. That is, it does not answer the deep question of how, and why, some of our neurobiological states have subjective, qualitative characters such that we are “phenomenally” aware of their subjective, qualitative characters simply by having them. The hard problem is far too difficult to hope to solve in the span of a dissertation, and it is unclear whether we will ever have a satisfactory scientific and philosophical solution to it. On the other hand, my
dissertation shows why we needn’t give up Antecedent Physicalism in the face of such daunting explanatory impediments.

Second, although the dissertation advances a substantive account of the fundamental role that conscious experience plays in our cognition of, and reference to, the world (including about conscious experience itself), it does not attempt provide a detailed cognitive or linguistic theory. The cognitive and linguistic sciences are amassing much information about the nature and inner-workings of human (and animal) cognitive and communicative capacities, but we are still a long way away from possessing anything approaching an adequate general theory of cognition or language use.\(^{15}\)

Finally, it will not directly challenge the truth of Dualism, Neutral Monism, Idealism, Panpsychism, Eliminative Materialism, or any other alternative metaphysical view of the metaphysics of conscious experience to my own preferred Antecedent Physicalist Identity Theory. While I find each of these views problematic in many respects, I am not convinced that we can conclusively rule out any of them (perhaps with the exception of Eliminative Materialism) as being true of conscious experience. On the other hand, my dissertation does show that the existence of a distinctive epistemic gap between the physical and the experiential gives us no cause for adopting such radical and/or mysterious metaphysical alternatives.

\(^{15}\) On a related note, I will also not offer a substantive theory of “knowledge of truths” or of how we get from knowledge of things to knowledge of truths. Nor will I take a stand on the conditions which make “knowledge of truths” count as bona fide \textit{knowledge} rather than mere \textit{belief}. 
CHAPTER 1:
RUSSELLIAN ACQUAINTANCE WITHOUT DISCRIMINATING KNOWLEDGE

“[Acquaintance is] the simplest and most pervading aspect of experience... All cognitive relations—attention, sensation, memory, imagination, believing, disbelieving, etc.—presuppose acquaintance.” – Bertrand Russell, *Theory of Knowledge*\(^{16}\)

“Knowledge of things, when it is of the kind we call knowledge by acquaintance, is essentially simpler than any knowledge of truths, and logically independent of knowledge of truths, though it would be rash to assume that human beings ever, in fact, have acquaintance with things without at the same time knowing some truth about them.” – Bertrand Russell, *The Problems of Philosophy*\(^{17}\)

“The faculty of being acquainted with things other than itself is the main characteristic of a mind. Acquaintance with objects essentially consists in a relation between the mind and something other than the mind; it is this that constitutes the mind’s power of knowing things.” – Bertrand Russell, *The Problems of Philosophy*\(^{18}\)

§1 INTRODUCTION

Russell’s Theory of Descriptions is at best a partial solution to “Frege’s Puzzle”.\(^{19}\)

Roughly speaking, Frege’s Puzzle is the problem of explaining how it is that distinct co-referential expressions (and their analogues in thought) can make different contributions to the cognitive significance of the utterances (or thoughts) in which they occur.\(^{20}\)

Following Frege, Frege’s Puzzle is most often discussed in the context of identity

---

\(^{16}\) Russell (1992), p. 5.

\(^{17}\) Russell (1912/1997), p. 46.


\(^{19}\) For an excellent, but not unflawed, recent discussion of this issue, see Proops (2011).

\(^{20}\) In saying here and elsewhere in this paper that linguistic expressions have “analogues in thought”, I mean to be neutral on the issue of whether there is a genuine, full-blown “language of thought” as well as on the issue of whether Russell thinks that there is one.
statements: an identity sign flanked by two referring expressions. In these paradigmatic cases, the problem is to explain how thoughts or utterances of the form ‘a = a’ and ‘a = b’ can differ in cognitive significance if ‘a’ and ‘b’ co-refer. In such cases, the first is analytic and a priori while the second need not be. But it is important to keep in mind that similar problems arise whenever we use (rather than mention) distinct expressions or concepts whose semantic contributions at the level of reference or denotation is the same and yet whose contributions to cognitive significance differ. For instance, neither ‘Hesperus is visible in the evening sky’ nor ‘Phosphorus is visible in the evening sky’ seem analytic or a priori, and yet there is a clear difference in cognitive significance between them.\(^2\)

Russell’s Theory of Descriptions (when combined with his well-known Descriptive Theory of Ordinary Names) provides two powerful resources for solving Frege’s Puzzle. First, it draws a distinction between the surface grammatical form of an utterance and its underlying logical form, allowing Russell to claim that most thoughts or utterances that appear to have the logical form ‘a = b’, for example, do not really have it. And second, it maintains that what most apparently referring expressions contribute to the actual logical form of the thoughts or utterances in which they occur is a denoting or quantificational complex. Thus, thoughts and utterances of the grammatical form ‘a = a’ either assert the self-identity of a particular object a or that there is a unique satisfier of such and such description which is the unique satisfier of that very description, both of which are trivial (aside from their existential import) . In contrast, most thoughts and utterances of the grammatical form ‘a = b’ either assert that a particular object a uniquely

\(^2\) Also consider familiar substitution problems in attitude ascriptions, as well as cases such as ‘Hesperus is visible in the evening sky only if Hesperus is’ and ‘Hesperus is visible in the evening sky only if Phosphorus is’.
satisfies such and such description or that there is a unique satisfier of one description which is also the unique satisfier of another description, both of which are often enough substantive (in addition to their existential import).

In this way, Russell’s Theory of Descriptions (combined with the Descriptive Theory of Ordinary Names) promises to solve Frege’s Puzzle for a significant class of thoughts or utterances. But it cannot by itself solve Frege’s Puzzle for thoughts or utterances of all logical forms. It cannot explain the difference in cognitive significance, if any, between thoughts or utterances really of the logical form ‘a = a’ and the logical form ‘a = b’ where ‘a’ and ‘b’ are co-referential genuinely singular referential devices. Nor can it explain the difference in cognitive significance, if any, between thoughts or utterances of the logical form ‘the unique F is the unique F’ and the logical form ‘the unique F is the unique G’ where ‘the unique F’ and ‘the unique G’ contribute the same denoting or quantificational complexes to thoughts and utterances in which they occur (think of ‘the capital of this country’ and ‘the capital of that country’ where the ‘this’ and ‘that’ co-refer).

On the received view, Russell has a simple, elegant solution to both lingering problems. The first part of this solution is to place a substantive constraint, which is often called “Russell’s Principle”, on what it takes to think or talk significantly about things. Roughly speaking, “Russell’s Principle” demands that we know which things we are thinking or talking about if we are to think or speak with understanding. Russell enunciates “Russell’s Principle” clearly in The Problems of Philosophy when he claims that:

It is scarcely conceivable that we can make a judgement or entertain a supposition without knowing what it is that we are judging or supposing about.
We must attach *some* meaning to the words we use, if we are to speak significantly and not utter mere noise; and the meaning we attach to our words must be something with which we are acquainted. However, it is important to recognize that “Russell’s Principle” does not all by itself solve the lingering versions of Frege’s Puzzle. After all, one can presumably know that ‘Cicero’ refers to the particular individual Cicero with whom we have ordinary “acquaintance” and that ‘Tully’ refers to the particular individual Tully with whom we have ordinary “acquaintance” without knowing that Cicero is Tully.

For this reason, the second part of Russell’s supposed solution to Frege’s Puzzle is to impose a particularly demanding version of “Russell’s Principle” on our thought and talk about things: we can only think or talk about things with understanding if we have an especially intimate acquaintance with them, or if we think of them as the unique satisfier of some description composed entirely out of elements with which we have especially intimate acquaintance. On the received interpretation, Russell’s technical notion of this especially intimate acquaintance, and his restrictions on objects we can be acquainted with, *preclude* cases in which $a$ and $b$ are the same, we are acquainted with the thing they both are, and yet we don’t know the right answer when we consider whether $a$ is $b$.

On this reading of Russell, cases such as that of Cicero and Tully above simply cannot arise when bona fide Russellian acquaintance is involved. Nor can cases of the form ‘the unique $F$’ and ‘the unique $G$’ where we are acquainted in Russell’s technical sense with all of the elements composing the complexes they contribute to our thought and talk. When combined with Russell’s Theory of Descriptions (and the Descriptive

---

23 From now on I will use the term ‘acquaintance’ in Russell’s technical sense, rather than in the ordinary sense.
Theory of Ordinary Names), the received view of Russellian acquaintance and its role in thought and talk constitutes a tidy and alluring possible solution to Frege’s Puzzle.

Nevertheless, there is scant evidence that Russell himself *ever* holds the received view of acquaintance, and none of it is compelling. On the contrary, Russell recognizes and discusses cases throughout his work in which “Russell’s Principle”, under the strong interpretations required by the received view, is false. Most importantly, Russell notes on several occasions that there are phenomenal continua cases in which we are acquainted with both *a* and *b* without thereby knowing whether we are acquainted with two distinct things, or one thing twice-over. In other words, *a* and *b* can be the same thing, we can be acquainted with the thing they both are, and yet *not* know whether *a* is *b*. But if this is so, then Russell does not solve Frege’s Puzzle by supplementing his Theory of Descriptions (and Descriptive Theory of Ordinary Names) with “Russell’s Principle” strongly construed. In my view, he never intended acquaintance to help solve Frege’s Puzzle in the first place.

In this chapter, I proceed as follows. First, I show that important theorists, such as Gareth Evans, Donald Davidson, John Campbell, and Ian Proops, have indeed attributed what I call “the received view of acquaintance” to Russell. Second, I present Russell’s discussions of phenomenal continua cases, which show that he recognizes that “Russell’s Principle”, strongly interpreted, is false. Third, I show that the possibility of misidentifying the objects of acquaintance follows from underappreciated aspects of Russell’s theory of knowledge and acquaintance. Finally, I address his discussion of George IV in “On Denoting”, which has (wrongly) convinced many interpreters that
Russell *does* intend the combination of acquaintance and the Theory of Descriptions to solve Frege’s Puzzle wholesale.

§2 THE RECEIVED INTERPRETATION

Many previous interpreters of Russell have attributed to him some version or other of what I call “the received view of acquaintance”, at least during some period of his thought. In its most general form, the received view of acquaintance holds that when a subject is acquainted with \(a\) and \(b\), designates \(a\) and \(b\) with distinct genuine proper names, and uses these genuine proper names to consider whether \(a\) is \(b\), he or she cannot fail to determine the correct answer as to whether \(a\) is \(b\) if the question arises. But it also encompasses the view that when a subject is acquainted with \(a\) and \(b\), he or she *thereby* knows whether or not \(a\) is \(b\). In either case, the central idea is that acquaintance is supposed to safeguard us from misidentifying the objects of our acquaintance.

One of the most influential of these interpretations is Gareth Evans’ in his *The Varieties of Reference* (1983). According to Evans, Russell holds the received view of acquaintance partly because of his deep-rooted Cartesian sentiments and partly because he thinks it is the only satisfactory way of complying with “Russell’s Principle”.\(^{24}\) Indeed, Evans claims that Russell requires that a subject have what he calls a “discriminating conception” of a thing in order to be able to think or talk about it with understanding, where a *discriminating conception* is a conception of a thing that enables the subject to distinguish it from all other things.\(^{25}\) Evans maintains that Russell sees only two ways of attaining such a demanding discriminating conception of a thing: either by

\[^{24}\text{Evans (1983), p. 45.}\]
\[^{25}\text{Evans (1983), p. 65.}\]
being acquainted with it or by thinking of it as the unique satisfier of a description
composed from elements all of which the subject is acquainted with. As Evans puts it,

[Russell] held a theory of thought which incorporated the principle that it is not
possible for a person to have a thought about something unless he knows which
particular individual in the world he is thinking about. Russell also held that
there are only two ways of discharging this requirement: one is to be, or to have
a memory of being, ‘acquainted’ with the object concerned, and the other is to
think of the object as the unique satisfier of some description.26

Regarding acquaintance itself, Evans correctly notes that Russell sees it as a two-
place relation that either holds between a subject and a corresponding object, or not at all.
However, Evans also recognizes that even if acquaintance is such a two-place relation
between a subject and an object, this alone would not secure a discriminating conception
of the object and thereby preclude the possibility of misidentifying it. For a subject could
conceivably stand in this two-place relation to one and the same thing many times over in
different circumstances or on different occasions without realizing that it is the self-same
object. Evans argues that this is the principal reason why Russell radically restricts our
acquaintance to sense-data and other “Russellian objects”.27 For these are the sorts of
objects, one might plausibly think, that a subject presumably could not fail to identify or
re-identify in different circumstances. Indeed, he says,

Russell himself had an excuse for his espousal (in effect) of the ordered-couple
conception of monadic Russellian thoughts, in that he restricted the objects of
such thinking mainly to items which were conceived to be so fleeting and
insubstantial that it seemed unintelligible to suppose a person might identify the
same one twice without knowing it was the same. But this excuse is not
available to those who extend the ordered-couple conception of Russellian
thoughts outside the sphere of thoughts about objects of Russellian
acquaintance.28

26 Evans (1983), p. 44.
27 The term “Russellian objects” comes from Proops (2011).
28 Evans (1983), p. 82.
Hence, to summarize Evans’ take on Russell, it is partly due to the relational character of acquaintance, and partly due to the special nature of the objects of acquaintance, that it provides a subject with a discriminating conception of an object sufficiently demanding to satisfy “Russell’s Principle”.

Evans is not the only influential philosopher to interpret Russell along these lines. In his “What is Present to the Mind?” (1989), Donald Davidson attributes a very similar view about the nature of acquaintance, and of “Russellian objects”, to Russell. Davidson disputes the general view that whenever we think about a thing, we have special objects before our mind (whether they are sense-data or propositions) which are impossible to misidentify and which fix the contents of our thought. Furthermore, Davidson aims to show that there is no special epistemic relation, such as Russellian acquaintance, which a subject bears to objects that enables him or her to entertain genuine de re thoughts about them. Rather, he contends, “any property of an object may, under suitable conditions, be considered the relevant identifier” which allows the subject to think de re thoughts about it.

Like Evans, the problem Davidson sees with Russell’s view is a commitment to a thoroughly Cartesian conception of thought, one according to which we cannot be wrong about what it is that we are thinking. Indeed, he argues,

What lies behind some of these attempts to characterize the special relation between the mind and its objects is, of course, the Cartesian drive to identify a sort of knowledge which is guaranteed against failure. If this search is combined with the assumption that all knowledge consists in the mind being in psychological contact with an object, then objects must be found about which

error is impossible—objects that must be what they seem and seem what they are.\(^{31}\)

Moreover, Davidson (like Evans) takes it that Russell’s solution consists not only in
taking thought to be ultimately grounded in a special epistemic relation subjects bear to
things, but also in restricting the things we can think about directly to special “Russellian
objects”. But, Davidson argues, not only is there no such thing as a single, special kind of
epistemic relation that allows us to think \emph{de re} thoughts about things,

there simply are no such [“Russellian objects”]. Not even appearances are
everything we think they are. Nor can the ‘aspects’ of sense data, if they really
are objects, be protected from one or another sort of misidentification.\(^{32}\)

Consequently, he remarks,

\begin{quote}
If the mind can think only by getting into the right relation to some object which
it can for certain distinguish from all others, then thought is impossible. If a
mind can know what it thinks only by flawlessly identifying the objects before
it, then we must very often not know what we think.\(^ {33}\)
\end{quote}

Incidentally, we will see that Russell himself would very much agree with Davidson on
this point. Still, Davidson clearly \emph{takes} Russell to be committed to the received view of
acquaintance.

More recently, in his 2009 “Consciousness and Reference”, John Campbell too
attributes the received view of acquaintance to Russell. On Campbell’s view, one of the
central issues motivating Russell’s use of the notion of acquaintance is the important
connection he sees between our consciousness of things and our knowledge of singular,
demonstrative reference. Campbell recognizes that for Russell, acquaintance is a non-
propositional species of epistemic contact that a subject can have with things that both

grounds the reference of simple referring terms and that provides him or her with knowledge of the reference of such terms (or at least positions the subject to acquire such knowledge of reference). Moreover, Campbell rightly sees Russellian acquaintance with things as psychologically and epistemologically more primitive than any propositional knowledge of truths about them, or descriptive designation of them by means of such truths.

In trying to make sense of such primitive, non-propositional knowledge by acquaintance, Campbell appeals to G. E. Moore’s (1903) notion of the “transparency of experience”. The basic idea here is that experience is a generic two-place relation between a subject and the objects of experience such that by specifying the objects experienced, we have thereby fully specified the character of the experience. The problem with this view, Campbell acknowledges, is that it makes it difficult to account for manifest sameness of reference across changes in the character of experience. According to Campbell, Russell was sensitive to this problem and responded to it by simply restricting the objects of experience to those that can only be experienced in one way. In other words, Campbell argues, Russell’s strategy was:

[to] keep the idea that we should characterize consciousness in Moore’s terms, as the holding of a generic relation between the self and an object. But we should restrict the range of the relation. We should keep it restricted to objects of which the subject automatically has comprehensive knowledge. (emphasis added)

Campbell warns, however, that we not take Russell’s view here to be that acquaintance exhaustively reveals the essential nature of the objects of experience, despite Russell’s

---

34 Campbell (2009), p. 650.  
36 Campbell (2009), p. 656. Campbell’s own solution is to construe acquaintance as a three-place relation between a subject, a standpoint, and an object.
(1912) well-known remarks that a subject knows a color “perfectly and completely” when he or she is acquainted with it. For as Campbell notes, this suggests that acquaintance is a species of knowledge of truths about essences rather than a species of knowledge more primitive than, and wholly distinct from, any knowledge of truths. Indeed, he contends,

Russell’s remark about completeness of knowledge should not be read as relating to propositional knowledge of essences; the remark is, rather, his response to the problem of partial awareness. The idea is that the colors are such that there is nothing partial about our awareness of them; so we can characterize acquaintance with them fully merely by saying which colors are being encountered. There are not, on Russell’s view, different ways of being acquainted with one and the same color. The point is rather that the knowledge of the thing is complete; there is no further, non-propositional knowledge of the thing to be had, once you have encountered it in experience. When put generally, this leads to the doctrine of special objects of awareness, which can be encountered in only one way.37

Finally, Proops (2011) attributes to the early Russell (prior to 1914) roughly the same view about the nature of acquaintance with objects as does Campbell. According to Proops, reflection on the case of George IV in “On Denoting” made Russell very much aware that his Theory of Descriptions is at best a partial solution to Frege’s Puzzle, and to substitution problems in general.38 His solution prior to 1914, Proops argues, was to supplement the Theory of Descriptions with the view that acquaintance fully discloses to a subject the objects of his or her acquaintance.39 Furthermore, Proops glosses this notion of “full disclosure” with the following epistemological principle:

---

37 Campbell (2009), p. 661.
38 Proops never actually claims that Russell is concerned with Frege’s Puzzle. Instead, Proops more plausibly takes Russell to want to defend what Proops calls the “Saving (truth-)Value Principle” (or “SV”), which he formulates as follows: “The substitution of identical propositional constituents within a proposition preserves that proposition’s truth-value”. Still, Proops does see Russell as concerned with substitutions of propositional constituents in attitude ascriptions and is to that extent concerned with Frege’s Puzzle. Proops (2011), p. 156.
Full Disclosure: Whenever a subject, $S$, is acquainted (in Russell’s technical sense of that term) with an object, $x$, $S$ is acquainted with every part of $x$.\footnote{Proops (2011), p. 152.}

Proops takes it that this is precisely the sense in which Russell takes our knowledge of a thing to be “perfect” and “complete” when we are acquainted with it: that we are acquainted with every part of it.\footnote{Proops bases his interpretation on a passage, which we will see in section §4, in which Russell maintains that acquaintance does not come in degrees, but rather is “either complete or non-existent”. One prima facie problem with Proops’ interpretation is that in the same passage Russell claims that we cannot literally become ‘better acquainted’ with any object but only “with more parts of a certain whole”. But if acquaintance provides Full Disclosure, then there is no real sense to be made of becoming ‘better acquainted’ with a thing since we are already acquainted with every part of it. Proops does not find this a problem “since Full Disclosure is a substantive thesis, [and] the idea of becoming better acquainted with something is not unintelligible.” In §4, I will offer a more natural reading of Russell’s remarks. Russell (1914/1960), p. 115 and Proops (2011), p. 170.} What this ensures is that when a subject is acquainted with a thing, it has no hidden parts and hence there is no possibility of the subject having different, partial “takes” on it.\footnote{Proops (2011), p. 171.} Consequently, Proops suggests, the subject’s acquaintance with a thing would preclude the possibility of misidentifying it, thereby blocking the lingering threat of Frege’s Puzzle.

Unlike many interpreters of Russell, however, Proops recognizes that Russell could not have held this received view of acquaintance throughout his career given his remarks about phenomenal continua cases in *Our Knowledge of the External World* (1914).\footnote{Proops (2011), p. 169.} Still, Proops is convinced that the earlier Russell, at least, is committed to the received view, one important consequence of which, Proops surmises, is Russell’s eventual abandonment of propositions as the objects of psychological attitudes in favor of his Multiple-Relation Theory of Judgment.\footnote{Proops (2011), p. 152.}
It is doubtful that Russell ever holds the received view of acquaintance, however. Though Russell clearly is interested in some substitution problems when discussing George IV in “On Denoting”, it is doubtful that his treatment of that case is meant to be a general solution to substitution problems, and in any case it offers scant support for “Full Disclosure” or for the received view of acquaintance. In point of fact, Russell’s actual views concerning acquaintance are primarily motivated by very different concerns than answering Frege’s Puzzle, contrary to prevailing Russell interpretation. But before I advance my own interpretation of his views on the nature of acquaintance, I will show that Russell indeed recognizes that “Russell’s Principle”, strongly interpreted, is false.45

§3 RUSSELL ON PHENOMENAL CONTINUA CASES

There are a number of aspects of Russell’s overall theory of knowledge that, to a greater or lesser extent, suggest that he did not in fact endorse the received view of acquaintance. However, the most decisive evidence in favor of the view that Russell rejected “Russell’s Principle” strongly construed comes from his frequent discussions of so-called “phenomenal continua”.46 Roughly speaking, phenomenal continua are cases in which one item or quality transitions into another item or quality such that no difference can be discriminated between two distinct points A and B in the transition, nor between two distinct points B and C, and yet some difference can be discriminated between points A and C. Such phenomenal continua cases conflict with the received view of acquaintance accept “Full Disclosure” at that time even while holding the Multiple-Relation Theory of Judgment. I would like to thank Russell Wahl for drawing attention to this second point. Also see Clark (1981) for similar remarks.


in that they generate circumstances in which a subject can evidently be acquainted with \( A \) and \( B \), designate \( A \) and \( B \) with distinct genuine proper names, consider whether or not \( A \) is \( B \), and yet fail to know what the right answer is.

Russell’s earliest sustained discussion of phenomenal continua cases appears in his 1914 Lowell Lectures, which were published the same year under the title *Our Knowledge of the External World*. In these lectures, Russell’s principal aim is to defend the view that we can usefully deploy logical-mathematical methodologies for solving traditional philosophical problems. For instance, Russell argues, we can usefully deploy logical-mathematical models exhibiting compactness to solve Zeno’s Paradox.

However, this is not Russell’s only aim. He also wants to answer what he takes to be the most powerful objection to his use of logical-mathematical methodologies in philosophical analysis: namely, that we cannot do so because experience shows us that the world is a single unity which does not have the parts or discrete units necessary for logical-mathematical modeling.\(^ {47} \) Russell voices this monistic challenge, which we might call “the Thesis of Interpenetration”, to philosophical analysis as follows:

> It is often urged that, as a matter of immediate experience, the sensible flux is devoid of divisions, and is falsified by the dissections of the intellect. Now I have no wish to argue that this view is contrary to immediate experience: I wish only to maintain that it is essentially incapable of being proved by immediate experience.\(^ {48} \)

The basic argument against philosophical analysis here evidently is this: Firstly, our experiential field does not present itself to us as being of a manifold of discrete sense-data. Secondly, the best, and only, explanation for this fact is that our experiential field is not of a manifold of discrete sense-data. But, thirdly, Russell’s ‘analytic’ methodology is

---

\(^ {47} \) Russell (1914/1960), pp. 103-21.

only adequate for analyzing discrete units (“logical atoms”). Therefore, Russell’s conception of philosophical analysis cannot be used to solve traditional philosophical problems.

In response to this challenge, Russell’s strategy is to let the first and third premises pass but to reject the second one. The problem with the second premise, he argues, is this: although the view that the world as we experience it is “devoid of divisions” is empirically adequate, so is his own sense-data view.\(^{49}\) For, he points out, if there are phenomenal continua cases, as we have good reason to believe, then they would present themselves to us in way that is indistinguishable for us from an experiential field “devoid of divisions”. But then, the above argument’s conclusion wouldn’t follow and we wouldn’t be led to give up his view of philosophical analysis.

In making his case, Russell relies on two different kinds of phenomenal continua case: a diachronic case of our experience of weight and a synchronic case of our experience of shades of color. A diachronic phenomenal continua case is one in which \(A\) at time \(t_1\) is indistinguishable from \(B\) at time \(t_2\) and \(B\) at time \(t_2\) is indistinguishable from \(C\) at time \(t_3\) but \(A\) at time \(t_1\) is distinguishable from \(C\) at time \(t_3\). A synchronic case is one in which \(A, B,\) and \(C\) are all co-present, where \(A\) and \(B\) or \(B\) and \(C\) cannot be distinguished pair-wise, but \(A\) and \(C\) can be. In describing the first, diachronic case, Russell says:

When we are considering the actual data of sensation in this connection, it is important to realise that two sense-data may be, and must sometime be, really different when we cannot perceive any difference between them. An old but

\(^{49}\) Indeed, in Russell (1914/1960), p. 113, he contends that “The question whether sense-data are composed of mutually external units is not one which can be decided by empirical evidence.” These remarks also show that Russell did not conceive of acquaintance as revealing the nature of its objects, as some interpreters have maintained. For if the Thesis of Revelation were correct, a subject could presumably tell on the basis of experience alone whether or not his or her experiential field has divisions.
conclusive reason for believing this was emphasised by Poincaré. In all cases of sense-data capable of gradual change, we may find one sense-datum indistinguishable from another, and that other indistinguishable from a third, while yet the first and third are quite easily distinguishable. Suppose, for example, a person with his eyes shut is holding a weight in his hand, and someone noiselessly adds a small extra weight. If the extra weight is small enough, no difference will be perceived in the sensation. After a time, another small extra weight may be added, and still no change would be quite easily perceptible; but if both weights had been added at once, it may be that the change would be quite easily perceptible.\footnote{Russell (1914/1960), p. 113.}

And, in describing the second, synchronic case, he remarks:

> Or, again, take shades of colour. It would be easy to find three stuffs of such closely similar shades that no difference could be perceived between the first and second, nor yet between the second and third, while yet the first and third would be distinguishable. In such a case, the second shade cannot be the same as the first, or it would be distinguishable from the third; nor the same as the third, or it would be distinguishable from the first. It must, therefore, though indistinguishable from both, be really intermediate between them.\footnote{Russell (1914/1960), p. 113.}

What Russell sees as important in each of these cases is that our capacities for discrimination are outrun by changes in the items or qualities that we are experiencing such that we might erroneously conclude that there \emph{are} no differences or boundaries between them. Consequently, the Thesis of Interpenetration is not the only empirically adequate explanation for why we do not experience the world as being a manifold of discrete sense-data, and hence his opponent’s monistic argument doesn’t go through.

However, the significance of phenomenal continua cases is not exhausted by their usefulness in defending Russell’s view of logical analysis against monistic challenges. For they also demonstrate that Russell recognized that \emph{even in the case of our own sense-data}, acquaintance does not secure a subject against the possibility of misidentifying the objects of his or her thought and talk.
According to Russell, the problem with the Thesis of Interpenetration is that it ultimately rests on the supposition that "the colours, being immediate data, must appear different if they are different". But, he urges, what phenomenal continua cases clearly demonstrate is that we should abandon this faulty supposition. In fact, Russell thinks that the prima facie attractiveness of this principle ultimately rests on a mistake: it is a result of the common conflation between a subject’s acquaintance with an object and the subject’s possession of knowledge of truths about that object. Indeed, he remarks:

It is unconsciously assumed, as a premise for reductio ad absurdum of the analytic view, that, if A and B are immediate data, and A differs from B, then the fact that they differ must also be an immediate datum. It is difficult to say how this assumption arose, but I think it is too be connected with the confusion between ‘acquaintance’ and ‘knowledge about.’ Acquaintance, which is what we derive from sense, does not, theoretically at least, imply even the smallest ‘knowledge about,’ i.e. it does not imply knowledge of any proposition concerning the object with which we are acquainted.

Moreover, Russell argues, the conflation of acquaintance and knowledge of truths can lead to other related misconceptions about the nature of acquaintance. For instance, he says:

It is a mistake to speak as if acquaintance had degrees: there is merely acquaintance and non-acquaintance. When we speak of becoming ‘better acquainted,’ as for instance with a person, what we must mean is, becoming acquainted with more parts of a certain whole; but the acquaintance with each part is either complete or non-existent. Thus it is a mistake to say that if we were perfectly acquainted with an object we should know all about it. ‘Knowledge about’ is knowledge of propositions, which is not involved necessarily in acquaintance with the constituents of the propositions.

Driving home the crucial significance of his remarks, Russell concludes thusly:

---

53 I will explain Russell’s distinction between acquaintance and knowledge of truths in greater detail in §4.
To know that two shades of colour are different is knowledge about them; hence acquaintance with the two shades does not in any way necessitate the knowledge that they are different.\textsuperscript{56}

What these passages clearly demonstrate is that Russell in fact \textit{explicitly rejects} the received view of acquaintance. Indeed, what these remarks show is that on Russell’s view, we can be acquainted with both $A$ and $B$, but not thereby be in a position to ‘trade on the identity’ of $A$ and $B$.\textsuperscript{57} For Russell, knowledge of the numerical identity or diversity of $A$ and $B$ is \textit{knowledge of truths} about $A$ and $B$, which he holds to be logically distinct from our more fundamental acquaintance with them. Hence, we cannot know, simply on the basis of being acquainted with $A$ and $B$, whether $A$ and $B$ are identical or distinct—we cannot know whether we are acquainted with two distinct things or one thing twice over. Moreover, even if we designate $A$ and $B$ with genuine names and consider whether $A$ is $B$, we might not know what the right answer is. That is why Russell’s view of logical analysis is on equal footing with the monistic view, empirically speaking. But given Russell views here, he clearly did \textit{not} think acquaintance guards a subject from problems of misidentification. So, Evans, Davidson, Campbell, and Proops are mistaken in thinking that Russell intends for acquaintance to help solve Frege’s Puzzle.\textsuperscript{58}

\section*{§4 KNOWLEDGE OF THINGS AND KNOWLEDGE OF TRUTHS}

Despite the clear evidence presented above, the received view of acquaintance is so familiar and established in philosophical lore that it might strike many philosophers as

\textsuperscript{57} Campbell (2009), p. 649.
\textsuperscript{58} In fairness to Proops, I still must show that Russell likely \textit{never} held the received view. I will do so in §5-6.
preposterous to suggest that Russell never held it. Nevertheless, if we pay careful attention to what Russell actually says about acquaintance and its role in reference and cognition it should come as no surprise that he allows for the possibility of misidentifying the objects of acquaintance. A careful reading of Russell’s works shows that his views on acquaintance are primarily motivated by concerns other than explaining how distinct co-referential expressions can make different contributions to the cognitive significance of thoughts or utterances in which they occur. So let us consider what Russell actually says about acquaintance.

On Russell’s picture of reference and cognition, there is a fundamental distinction, one that is often marked in natural language, between two kinds of knowledge we can have about individuals and their features. First of all, there is our familiar notion of knowing that such-and-such is the case. This sort of knowledge paradigmatically involves conceptually articulated beliefs and is propositional (or multiply-relational) in character; it involves cognitive attitudes that can be evaluated in terms of their truth or falsity. When we have this kind of knowledge about an object, we know that the object is thus-and-so. However, Russell points out that there is another way in which we use the term ‘knowledge’ in ordinary language: we talk of knowing individuals in the sense that we have directly encountered them in experience and therefore have some familiarity with them. When we have this kind of knowledge about an object or feature, we simply know it.

Correspondingly, Russell argues, any adequate theory of our knowledge of the world must be careful to distinguish our knowledge of things from our knowledge of...
truths about those things. Indeed, in his *The Problems of Philosophy* (1912/1997), Russell remarks:

The word ‘know’ is... used in two different senses. (1) In its first use it is applicable to the sort of knowledge which is opposed to error, the sense in which what we know is true, the sense which applies to our beliefs and convictions, i.e. to what are called judgements. In this sense of the word we know that something is the case. This sort of knowledge many be described as knowledge of truths. (2) In the second use of the word ‘know’ above, the word applies to our knowledge of things, which we may call acquaintance.  

Knowledge of things, for Russell, is the sort of knowledge a subject has that is constituted by his or her basic experiential contact, or “acquaintance”, with things and features in the world. Knowledge of truths, on the other hand, is knowledge of facts or propositions involving things and features in the world.

Roughly speaking, Russell conceives of acquaintance as a *fundamental experiential relation* we bear to objects and features, where the knowledge of things that it constitutes is logically independent of, but serves as the ultimate enabling grounds for, our knowledge of truths about, and conceptions of, those objects. For Russell, acquaintance is relational in the straightforward sense that it is a dyadic relation that obtains between a subject and the objects or features with which the subject is acquainted; both the subject and the objects or features are constituents of the acquaintance relation (or, more properly, of the relational fact). As Russell puts it:

The faculty of being acquainted with things other than itself is the main characteristic of a mind. Acquaintance with objects essentially consists in a

---

60 P. 44.
61 Note that Russell is not entirely consistent in his uses of the terms “acquaintance” and “knowledge of things”, even within the same works. Sometimes he identifies them, but other times he treats acquaintance as one kind of knowledge of things, in which case he allows for the possibility of knowledge of things by description. Given that knowledge by description puts one in a position to think and talk about things, but doesn’t all by itself assert that something is thus-and-so, I prefer the latter practice. However, for ease of exposition, I here use “knowledge of things” as equivalent to “knowledge by acquaintance” throughout.
relation between the mind and something other than the mind; it is this that constitutes the mind’s power of knowing things.\(^{62}\)

Moreover, Russell holds that acquaintance is an experiential relation between a conscious subject and objects or their features. Indeed, Russell is quite explicit in identifying acquaintance with a kind of direct consciousness awareness. In Theory of Knowledge (1992), Russell remarks that:

Now, since we have decided that experience is constituted by a relation, it will be better to employ a less neutral word we shall employ synonymously the two words ‘acquaintance’ and ‘awareness’, generally the former. Thus when \(A\) experiences an object \(O\), we shall say that \(A\) is acquainted with \(O\).\(^{63}\)

In addition, Russell holds that acquaintance is a direct conscious awareness, or “presentation”, of things in the sense that our awareness of them is not inferentially or descriptively mediated. He says:

We shall say that we have acquaintance with anything of which we are directly aware, without the intermediary of any process of inference or any knowledge of truths. Thus in the presence of my table I am acquainted with the sense-data that make up the appearance of my table—its colour, shape, hardness, smoothness, etc.; all these are things of which I am immediately conscious when I am seeing and touching my table.\(^{64}\)

Most importantly, acquaintance is an epistemically fundamental cognitive relation, one which constitutes our most basic form of intentionality, or object-directedness. He contends:

All our knowledge, both knowledge of things and knowledge of truths, rests upon acquaintance as its foundation.\(^{65}\)

In claiming this, Russell is not here endorsing any form of epistemic Foundationalism.\(^{66}\)

Rather, what Russell has in mind is that the knowledge we have of things when we are


\(^{64}\) Russell (1912/1997), p. 46. See also, Russell (1910/1929), pp. 209-10.

acquainted with them plays a fundamental role in our cognition of them. In particular, Russell’s view is that our conscious awareness of things and their features is what puts us in a position to select and attend to them, to direct our information-gathering methods and capacities at them, and to thereby form a conception of them that, if all goes well, constitutes knowledge of truths about them.

Nevertheless, acquaintance, all by itself, is not a sufficient condition for our having any conception of, or knowledge of truths about, the things with which we are acquainted. In fact, Russell makes very clear the logical independence of our knowledge of things from any knowledge of truths about them, when he contends that:

Knowledge of things, when it is of the kind we call knowledge by acquaintance, is essentially simpler than any knowledge of truths, and logically independent of knowledge of truths, though it would be rash to assume that human beings ever, in fact, have acquaintance with things without at the same time knowing some truth about them. Knowledge of things by description, on the contrary, always involves...some knowledge of truths as its source and grounds.  

Consequently, acquaintance cannot be equated with, or analyzed in terms of, our possession of any conception of, or propositional knowledge about, the object of our acquaintance.

Few have recognized just how radically different knowledge of things and knowledge of truths are on Russell’s view. One reason for this is that many have been misled by the following oft-cited passage from The Problems of Philosophy:

---

66 In fact, Russell’s views on the structure of knowledge are Foundationalist only in this regard: our self-evident (but not infallible or absolutely certain) intuitive knowledge of truths, including intuitive knowledge of logical truths, is the epistemic foundation for our derivative knowledge of truths. Russell (1912/1997), pp. 109-10.
68 Wahl (2007), Campbell (2009), and Proops (2011) are among the few exceptions who recognize the basic distinction, but even they do not characterize Russell’s views on knowledge of things by acquaintance adequately.
The particular shade of colour that I am seeing may have many things said about it—I may say that it is brown, that it is rather dark, and so on. But such statements, though they make me know truths about the colour, do not make me know the colour itself any better than I did before: so far as concerns knowledge of the colour itself, as opposed to knowledge of truths about it, I know the colour perfectly and completely when I see it, and no further knowledge of it itself is even theoretically possible. Thus the sense-data which make up the appearance of my table are things with which I have acquaintance, things immediately known to me just as they are.69

In particular, Russell’s description here of knowing colors we are acquainted with “perfectly and completely” and “just as they are” has suggested to many that the crucial difference between knowledge of truths and knowledge of things is simply that the former is partial and fallible whereas the latter is in some important sense not. Granted, the specific views concerning how knowledge of things differs from knowledge of truths varies from views that the former is fully disclosing to views that it is “revelatory”, or even “infallible”.70 Still, the general prevailing view is that knowledge of things in some way or other provides us with especially intimate and secure knowledge about the objects of our acquaintance.

However, such readings of Russell’s remarks deeply misconstrue what he means by claiming that our knowledge of things does not allow for error. Indeed, Russell’s actual view becomes clearer later on in The Problems of Philosophy (among other places) when he remarks:

Our knowledge of truths, unlike our knowledge of things, has an opposite, namely error. So far as things are concerned, we may know them or not know them, but there is no positive state of mind which can be described as erroneous knowledge of things, so long, at any rate, as we confine ourselves to knowledge

---

70 For a few examples, Proops (2011) holds that acquaintance is fully disclosing, Johnston (1992) claims it is “revelatory”, and Bigelow and Pargetter (2006) see it as “incorrigible” (but it is clear that they mean “infallible”).
by acquaintance. Whatever we are acquainted with must be something; we may draw wrong inferences from our acquaintance, but acquaintance itself cannot be deceptive. Thus there is no dualism as regards acquaintance. But as regards knowledge of truths, there is a dualism. We may believe what is false as well as what is true.  

The crucial claim in this passage is that when it comes to things rather than truths “we might know them or not know them, but there is no positive state of mind which can be described as erroneous knowledge of things.” On Russell’s view, there simply is no sense to be made of acquaintance being erroneous because it isn’t the sort of epistemic relation that provides a ‘take’ on things. Indeed, unlike our knowledge of truths, it doesn’t purport that things are thus-and-so at all, and that is why it can’t purport them to be otherwise than they are. Rather, acquaintance is simply the direct experiential contact with things that puts us in a position to take them as being thus-and-so.

We might usefully make Russell’s point as follows: In the case of our knowledge of truths, we can not only have or not have the relevant cognitive states, when we do have them, they can be veridical (or satisfied) or not. But in the case of acquaintance, we can only have it or not have it; there simply is no further issue of veridicality (or satisfaction). In fact, this is precisely what Russell really seems to be getting at when, in a passage cited in the previous section of this paper, he says: “it is a mistake to speak as if acquaintance had degrees: there is merely acquaintance and non-acquaintance” and “acquaintance with each part [of a thing] is either complete or non-existent.”

---

71 P. 119. For a similar remark, see Russell (1992), p. 49.  
72 Of course, we achieve experiential contact with things by means of epistemic faculties that purport things to be thus-and-so. Still, Russell thinks there is a clear logical distinction to be made: acquaintance with an object is one thing, and the knowledge of truths it positions us to acquire is another.  
73 Indeed, contra Proops (2011), there is scant reason to read Russell here (or elsewhere) as endorsing anything like Full-Disclosure. Russell (1914/1960), p. 115. Also see page 16 above.
Russell introduces his notion of knowledge of things for a purpose quite different from trying to solve “Frege’s Puzzle” concerning differences in the cognitive significance of co-designating expressions: his real aim is to explain the fundamental role that conscious experience plays in reference and cognition. In particular, Russell wants to explain how our direct conscious awareness of things puts us in a position to make genuine singular reference to them in thought and talk, as well as how we can derivatively come to think and talk significantly about things with which we lack direct experiential contact.

Roughly speaking, Russell holds the view that our thoughts, and the statements that express them, are associated with conditions such that the thoughts and statements are true if the conditions obtain, and false if they don’t. Furthermore, he thinks that the objects and features that genuinely singular referential devices refer to are part of these conditions. With singular denoting devices, in contrast, the objects or features denoted are picked out by these conditions, and required to exist if the conditions obtain. But they are not themselves integral parts of the satisfaction conditions of our thought and talk.

We can also put the point as follows. On Russell’s view, the distinctive feature of a genuinely singular referential device is that its meaning (that is, its contribution to the satisfaction conditions of thoughts or utterances in which it occurs) is the very worldly object or feature, if any, for which it stands. Furthermore, unlike in cases of singular denoting devices, such as definite descriptions and their analogues in thought, in the case

— 74 —

We should note that Russell (1992), pp. 39-40, recognizes that “egocentric” expressions (such demonstratives and indexicals) have something like Kaplan’s (1978/2008) notion of a “character” or Perry’s (2001) notion of a “reflexive content” in addition to their referential meanings. Even still, he thinks that the demonstratively or indexically designated individual or feature is what figures directly into the satisfaction conditions of our thought and talk. Also see Russell (1940), chapter 7, Russell (1948), part II chapter 4, and Russell (1957).
of genuinely singular referential devices, we needn’t have any conception of their referents or the mechanisms whereby their reference is determined. The basic distinction Russell draws here is between what we might call “reference-guided conceptions of things” and “conception-guided reference to things”; and, on his view, the former is the more fundamental.

But what exactly does it take to have genuinely singular thoughts about objects? That is, what does it take to have thoughts whose meanings are constituted by the very objects and features, if any, that the thoughts are about rather than by some way of getting at or identifying them? Russell’s view is that we can only make genuine singular reference to a thing if we are acquainted with it. In fact, Russell holds that all cognition and linguistic designation ultimately rests more or less directly on our fundamental epistemic capacity to be acquainted with, or consciously aware of, objects and their features.

On Russell’s view, acquaintance puts us in a position both to make genuine singular reference to objects and features in the world and to pick up information about, and thereby form conceptions of, them. In turn, these conceptions enable us to denote, and thereby think and talk about, things with which we lack experiential contact. In other words, our genuinely singular knowledge of things explains how it is that we are able to acquire the descriptive materials, both about particular things and the general way the world works, required for conception-guided designation in the first place.\(^{75}\)

\(^{75}\) In Russell (1912/1997), p. 45, he says, “What happens, in cases where I have true judgement without acquaintance, is that the thing is known to me by description, and that, in virtue of some general principle, the existence of a thing answering to this description can be inferred from the existence of something with which I am acquainted.”
Basically, Russell’s picture of cognition and reference is this: We begin by having direct experiential contact with objects (concrete and abstract alike) and their features. On the basis of this experiential contact, we are in a position to direct our conscious attention to those objects and features. Our conscious attention to these experienced objects and features in turn grounds our ability to make demonstrative, singular reference to them and puts us in a position to acquire knowledge that we are referring to them. Furthermore, our conscious attention to these objects and features puts us in a position to pick up information about them, and thereby form some conception of them. Finally, on the basis of our conceptions of these objects and features and our more general conception of how the world works, we can designate objects and features with which we lack experiential contact. We do so by employing representations that encode identifying conditions that an object or feature must uniquely satisfy in order to be designated by it, or by employing representations that are purely quantificational in character.

However, Russell is acutely aware of a long-recognized problem confronting any view that identifies the referent of a linguistic expression (or its analogue in thought) with its meaning: namely, that there are many apparently meaningful genuinely singular referential devices in natural language and thought which lack corresponding referents in the world. For instance, we can have perfectly meaningful thoughts or make perfectly meaningful utterances using expressions such as ‘Santa Claus’, ‘Vulcan’, or even ‘Roundsquare (the largest round square)’ even though there is no actual Santa Claus, planet Vulcan, or round square in the world. Moreover, many of our thoughts and utterances involving such “empty” genuinely singular referential devices seemingly admit of truth or falsity. If nothing else, we can apparently judge or assert meaningfully

---

and truthfully that Santa Claus, Vulcan, and Roundsquare do not exist.\textsuperscript{77} Hence, Russell evidently needs to explain how our thought and talk involving “empty” genuinely singular referential devices can be meaningful and (at least sometimes) truth-evaluable.

At the same time, Russell wants to avoid two problematic strategies for addressing the problem of reference-failure. First, he wants to avoid views according to which the meanings of putative genuinely singular referential devices are “ideas”, “concepts”, “contents”, or “Fregean senses”. One reason for this, we’ve seen, is that Russell sees such conception-guided designation as derivative from our more basic reference-guided thought and talk.\textsuperscript{78} But he also thinks there are a number of reasons to reject such views that don’t stem from epistemological considerations. For Russell, many of the shortcomings of these views have to do with their application to logical problems, and particularly with their inability to help provide a satisfactory solution to his “Contradiction” involving the class of all classes that are not members of themselves.\textsuperscript{79}

However, Russell also thinks that views that take the meanings of putative genuinely singular referential devices to be ideas or Fregean senses simply assign the wrong satisfaction conditions to our thought and talk. For one thing, he sees such views, if taken literally, as having implausible consequences for predication. Indeed, he says, “If we say ‘the King of England [in 1905] is bald’, that is, it would seem, not a statement about the complex meaning ‘the King of England’, but about the actual man denoted by

\textsuperscript{77} It’s worth noting that for Russell (1918, 1919), existence is not something we can intelligibly predicate of objects. Instead, he takes existence to apply solely to propositional functions.

\textsuperscript{78} In fact, Russell makes this point at least as early as his unpublished “On Fundamentals” (1905) when he says, “But we can only know an object as denoted if we are acquainted with the denoting concept; thus immediate acquaintance with the constituents of the denoting concept is presupposed in what we may call \textit{denotative} [or “descriptive”] knowledge.”

\textsuperscript{79} Urquhart (1994) provides a nice overview of these issues. Also see Klement (2004).
the meaning”.  On the other hand, if we instead interpret such views as holding that what we make predications of is the denoted object, if any, of the idea or Fregean sense, then they seemingly commit us to holding that some thoughts or utterances we would intuitively take as having truth-values actually lack them. For instance, he argues, “[On such views], ‘the King of France [in 1905] is bald’ ought to be nonsense [that is, truth-valueless]; but it is not nonsense, since it is plainly false”. Likewise, he says,

The King in “The Tempest” might say, “If Ferdinand is not drowned, Ferdinand is my only son.” Now “my only son” is a denoting phrase, which, on the face of it, has a denotation when, and only when, I have exactly one son. But the above statement would nevertheless be true if Ferdinand had been in fact drowned. Given that such views do not assign the intuitively correct truth-values to thoughts or utterances involving “empty” genuinely singular referential devices, Russell concludes, we ought to avoid them—even if they could perhaps account for the meaningfulness of “empty” thought and talk.

In addition, Russell adamantly wishes to avoid the view (which he associates most frequently with the Austrian philosopher Meinong) that “empty” genuinely singular referential devices pick out objects or features that don’t exist in the world, but nevertheless “subsist in some shadowy Platonic world of being”. Russell rejects this view for a host of reasons, not the least of which is that he finds the notion of ‘non-existent being’ simply incredible when applied to objects (and likely even incoherent since it seems to “infringe the law of contradiction”).

---

80 Russell (1905/2000), p. 214. Of course, ‘the King of England’ is a denoting phrase rather than a genuinely singular referential device, but the same issues raised here by the former apply to the latter.
82 Russell (1905/2000), p. 215. Russell goes on to consider Frege’s view that “empty” expressions denote the null-class but rejects it on the grounds that “though it may not lead to actual logical error, it is plainly artificial, and does not give an exact analysis of the matter.”
83 Russell (1959), p. 84.
Russell’s own solution to the problem of reference-failure is straightforward: he simply restricts all genuinely singular reference to objects of our acquaintance. To see how this helps, we should recall that for Russell, it falls out of the very nature of acquaintance as a dyadic relation holding between a conscious subject and an object or feature that we cannot fail but to be acquainted with *something*, if we stand in the acquaintance relation. There simply is no sense to be made of standing in the acquaintance relation to nothing at all. As Russell puts it:

> If an object is given in acquaintance, then that object has a certain relation to the subject which is acquainted with it. But this would be meaningless if there were no such object. An object of acquaintance is an object to which it is possible to give a proper name, as opposed to a description; it may become the ‘this’ of attention. Of such an object, it is meaningless to suppose it unreal. A proper name which names nothing is not a proper name, but a meaningless noise. An acquaintance which is acquainted with nothing is not an acquaintance, but a mere absurdity.\(^{85}\)

Basically, Russell’s idea here is that if all reference is restricted to objects of acquaintance, and the latter requires both a subject and an object, then all reference will have some object or other. So reference-failure would be straightforwardly ruled-out.

Of course, reference itself is evidently a dyadic relation which also requires both a subject and an object, and so one might wonder if there is a need to restrict reference to the objects of our acquaintance. But on Russell’s view, there is a crucial difference between the relation of acquaintance and the relation of reference: while the former puts us in a position to form intuitive knowledge of truths to the effect that we stand in that very relation to some object, standing in the reference relation does not by itself similarly put us in a position to form any intuitive beliefs whatsoever.\(^{86}\) Hence, by restricting

---


reference to the objects of our acquaintance, we are at least in a position to know (with a fair degree of confidence) whether putative acts of reference do in fact refer to anything.

We are also now in a position to see why Russell somewhat notoriously restricts our acquaintance to sense-data, abstracta, and (perhaps) the self. In the case of most kinds of objects, there is a legitimate and meaningful question concerning whether or not they exist, a question that Russell takes to be “meaningless” when applied to objects of acquaintance. But in the case of sense-data, abstracta, and (perhaps) the self, it is, though possible, in some sense pathological to doubt their existence. Indeed, when it comes to the existence of sense-data, abstracta, and (perhaps) the self, we might not have absolute certainty, but we nevertheless have the highest degree of assurance that we can achieve. Hence, if acquaintance is restricted to such objects, and if our genuinely singular thought and talk is restricted to the objects of our acquaintance, then we have the highest assurance that our thought and talk is by-and-large meaningful—without any need to posit non-existent Meinongian objects or intermediary Fregean senses. For all of our meaningful thought and talk would either consist in, or be derived from, genuinely singular thought and talk about objects of whose existence we can be quite sure.

Of course, we must be careful here. Knowing that sense-data, abstracta, and (perhaps) the self exist is knowledge of truths, not knowledge of things. Indeed, while our

---

87 In the case of sense-data, Russell is also convinced by the Argument from Illusion and the Argument from Science that Naïve Realism is false. Russell (1912/1997), Chapters 1-3.

88 On the other hand, it is emphatically not pathological, on Russell’s view, to doubt any particular substantive account of sense-data, abstracta, or the self. In fact, Russell’s hesitancy in saying we are acquainted with the self stems from his deep uncertainty that there is a persisting self, or one over-and-above the collection (or “bundle”) of our present thoughts, experiences, and physical features (if any).

89 Russell is quite clear in saying that we can never attain complete certainty with respect to our knowledge. Indeed, he says, “all our knowledge of truths is infected with some degree of doubt, and a theory which ignored this fact would be plainly wrong.” Russell (1912/1997), p. 135.

90 In fact, this is one reason why Russell adopted the Multiple-Relation Theory of Judgment. For if all thinking is a matter of a subject standing in a multiple relation to objects all of which he or she is acquainted with, then we can straightforwardly account for its meaningfulness without the need for Fregean senses or Meinongian objects.
acquaintance with things ensures that those things exist, it does not all by itself provide us
with any knowledge that they exist. Acquaintance simply puts us in a position to acquire
such knowledge of existence by presenting us with objects and features in experience.
Likewise, acquaintance puts us in a position to acquire knowledge of the identities of
objects and features by presenting them to us in experience. But nothing Russell actually
says about acquaintance requires, or even supports, the received view of acquaintance or
“Russell’s Principle” strongly construed. Hence, we have no reason to be particularly
surprised that Russell rejected both.

§5 RECONSIDERING “ON DENOTING”

But why then, we might wonder, have so many taken Russell to hold the received view of
acquaintance? One evident source of confusion, which we saw above, is the manner in
which Russell himself occasionally describes knowledge by acquaintance: as knowing
things “perfectly and completely”, “just as they are”, and without the possibility of error.
Another possible source of misunderstanding is independent and antecedent conceptions
about what acquaintance and/or special “Russellian objects” would likely be like.91
However, I suspect that many of the misconceptions about Russell’s views about
acquaintance actually stem from widespread, but questionable, readings of Russell’s
well-known George IV case in “On Denoting” (1905).

91 Some of Russell’s own remarks about “fleeting” sense-data might seem to support common
misconceptions about one-sided “Russellian objects”. For instance, in Russell (1918, 1919), he says,
“[sense-data] last for a finite time, a matter of seconds or minutes or whatever it may happen to be”. However, it is doubtful that his remarks about fleeting sense-data are motivated by a desire to rule out
problems of misidentification. Indeed, diachronic phenomenal continua cases can occur during very short
periods of time, and synchronic cases are entirely unaffected. Thanks to Russell Wahl for drawing my
attention to Russell’s remarks about fleeting sense-data.
The George IV case in “On Denoting” concerns the following prima facie puzzle about the apparent failure of truth-preservation under the substitution of co-designating expressions in psychological attitude ascriptions:

If \( a \) is identical with \( b \), whatever is true of the one is true of the other, and either may be substituted for the other in any proposition without altering the truth or falsehood of that proposition. Now George IV wished to know whether Scott was the author of *Waverley*; and in fact Scott was the author of *Waverley*. Hence we may substitute *Scott* for *the author of “Waverley”*, and thereby prove that George IV wished to know whether Scott was Scott. Yet an interest in the law of identity can hardly be attributed to the first gentleman of Europe.\(^92\)

The puzzle seems to be this, in Russell’s estimation: If ‘Scott’ and ‘the author of *Waverley*’ both simply contribute Scott to the satisfaction conditions of thoughts or utterances in which they occur, then given the nature of numerical identity, we should be able to substitute one for the other in any thought or utterance without altering its truth-value. But substituting ‘Scott’ for ‘*the author of Waverley*’ in ‘George IV wished to know whether Scott was the author of *Waverley*’ does evidently change its truth-value, even though Scott did solely author *Waverley*. So why does the substitution of ‘Scott’ for ‘*the author of Waverley*’ lead from truth to falsity?

On Russell’s view, the Theory of Descriptions provides just what we need to solve the puzzle of George IV. Indeed, Russell maintains, the answer to the puzzle is simply to deny that ‘*the author of Waverley*’ means Scott, or anybody else.\(^93\) In other words, the descriptive phrase ‘*the author of Waverley*’ doesn’t contribute *any* object or feature into the satisfaction conditions of every thought or utterance in which it occurs. But it doesn’t contribute a Fregean sense to them either. Instead, descriptive phrases such as ‘*the author of Waverley*’ contribute quantificational complexes to the satisfaction

---

\(^{92}\) Russell (1905/2000), p. 215. We will consider this passage in greater detail below.

\(^{93}\) Russell (1959), p. 85.
conditions of the thoughts or utterances in which they occur, but not necessarily the same ones on every occasion. One and the same descriptive phrase might, for instance, contribute a wide-scope quantificational complex to the satisfaction conditions of one kind of thought or utterance, but a narrow-scope quantificational complex to those of another. But while descriptive phrases do not uniformly contribute the same constituents to the satisfaction conditions of every thought or utterance in which they occur, it does not follow that they lack meaning altogether. Indeed, so long as we can at least give them a “contextual definition” (that it, say what they contribute to the satisfaction conditions of each particular thought or utterance in which they occur), we can account for their meaningfulness.

The upshot of the Theory of Descriptions for the George IV case is straightforward. Since ‘Scott’ and ‘the author of Waverley’ contribute different elements to the satisfaction conditions of ‘George IV wished to know whether Scott was the author of Waverley’, it is no surprise that substituting one for the other does not preserve truth-value in all cases. What’s more, the Theory of Descriptions doesn’t require the Fregean distinction between the sense and reference of an expression (or its analogue in thought), as did Russell’s earlier attempts to account for the semantic features of denoting phrases. So we can avoid the problems of assigning the wrong satisfaction conditions to our thought and talk (particularly when it involves descriptive phrases) which are engendered by Fregean semantic theories. Thus, Russell’s treatment of the George IV case provides an attractive and powerful tool for explaining why failures of truth-

95 Another important feature of the Theory of Descriptions is that in giving a “contextual definition” for a denoting phrase, it is (in Russell’s words) “broken up” in the sense that “in any proposition in which it occurs the proposition, fully expressed, does not contain the phrase”. Russell (1905/2000), p. 217.
96 Russell (1903a, 1903b, 1903c, 1903d, and 1903e).
preservation (at least sometimes) occur when co-designating genuinely singular referential devices and descriptive phrases are substituted for one another in our thought and talk.

The prevailing reading of “On Denoting” takes Russell to intend for his treatment of the George IV case (together with his Descriptive Theory of Ordinary Names) to be a general solution to Frege’s Puzzle (and related substitution problems). On this reading, whenever we have an apparent instance of Frege’s Puzzle, Russell holds that at least one of the expressions is a denoting phrase (often a “truncated” one) which contributes a quantification complex to the satisfaction conditions of the thought or utterance in which it occurs rather than an object or feature. This is what fully explains potential substitution failures involving co-designating expressions and any corresponding differences in cognitive significance, or so the story goes.

There is some textual evidence in “On Denoting” that supports this reading. For one thing, Russell first presents the George IV case immediately after his criticisms of Frege’s sense-reference distinction, which makes it very natural to see his solution to the former as a complete alternative to the latter. For another, Russell explicitly states that “the usefulness of identity is explained by [the Theory of Descriptions]. No one outside a logic-book ever wishes to say ‘x is x’, and yet assertions of identity are often made in such forms as ‘Scott was the author of Waverley’”. If we take Russell’s remarks here at face value, then he certainly seems to be claiming that the one and only non-trivial use of identity consists in its application to cases like George IV, where at least one of the expressions flanking the identity sign is a descriptive phrase (“truncated” or otherwise).

97 I have in mind, for instance, Proops’ (2011) discussions of the “Saving (truth-)Value Principle”.
Hence, we have good reason (it seems) to take Russell’s treatment of George IV to be his general solution to instances of Frege’s Puzzle.

On the other hand, we noted at the beginning of this paper that the Theory of Descriptions can’t explain all instances of Frege’s Puzzle. Indeed, it is manifestly unable to explain differences in the cognitive significance between thoughts or utterances with the genuine logical form of ‘a is a’ and those with the genuine logical form of ‘a is b’ when ‘a’ and ‘b’ co-refer. Nor can the Theory of Descriptions explain differences in the cognitive significance between thoughts or utterances of the genuine logical form of ‘there is a unique F which is the unique F’ and those with the genuine logical form of ‘there is a unique F which is the unique G’ when ‘the unique F’ and ‘the unique G’ contribute one and the same quantificational complexes to the thoughts or utterances in which they occur. So it looks as if Russell intends for the Theory of Descriptions to account for all instances of Frege’s Puzzle even though it clearly can’t do so. But then what do we make of this? Does Russell simply make a colossal mistake?

This is where the received view of acquaintance begins to gain some real plausibility. For if Russell were to accept it, then true thoughts and utterances of the genuine logical form ‘a is b’ or ‘there is a unique F which is the unique G’ (where ‘the unique F’ and ‘the unique G’ contribute the same quantificational complexes) would seemingly turn out to be obviously and transparently true. And likewise for false ones. As a result, the only instances of Frege’s Puzzle that would remain would be ones to which we could usefully apply the Theory of Descriptions. So charity suggests that we should attribute to Russell an acceptance of the received view of acquaintance in order to
vindicate his claim that the Theory of Descriptions explains the usefulness of identity.

While this line of reasoning looks initially plausible, it is far more tenuous than first appearances suggest. First of all, even if Russell were to think that the Theory of Descriptions explains all instances in which “identity is worth affirming” there is no evidence that Russell took acquaintance to be part of the explanation for why this is the case.99 He never says that acquaintance shields a subject from the possibility of misidentifying the objects of acquaintance, either in “On Denoting” or elsewhere. He never even says that acquaintance is supposed to play a significant role in solving Frege’s Puzzle (at least apart from its role in all thought and talk). In fact, he never so much as mentions acquaintance in the passages in “On Denoting” in which he discusses the George IV case; nor does he mention the George IV case (or misidentification or substitution problems) in the passages in which he explicitly discusses acquaintance.100

What’s more, in “On Denoting” Russell seemingly does not accept important aspects of his later epistemological views which evidently are required for making the received view of acquaintance even remotely plausible. In particular, he does not accept his later views that the objects of our acquaintance are special “Russellian Objects” such as sense-data. Indeed, while Russell contends that we are not acquainted with such things as “the center of mass of the solar system”, “other people’s minds”, or “matter (in the sense in which matter occurs in physics)”, he doesn’t make the stronger claim that we are not (perceptually) acquainted with ordinary, medium-sized material objects, or with the

bodies of other people.\footnote{Russell (1905/2000), pp. 212 and 219. One might speculate that Russell is here only speaking loosely about “seeing” Scott in the distance in order to make a quite different point. However, it is well-known from Russell’s (1959) own remarks that he was a naïve realist about perception early on, and to my knowledge, there is no textual evidence that Russell gave up this view until roughly 1910.} On the contrary, Russell at one point suggests that we can literally see other people (though not their minds) when he remarks that George IV might have “seen Scott at a distance”\footnote{Russell (1905/2000), p. 218.} But ordinary, medium-sized material objects (such as bodies) and their perceptible features are precisely the kinds of things we are liable to misidentify, a point Russell surely recognizes.\footnote{And if Russell didn’t recognize this point, he would have no reason to accept the received view of acquaintance.}

In fact, Russell’s discussion about George IV seeing Scott at a distance seemingly is in direct conflict with the received view of acquaintance. Indeed, in the relevant passage, he says:

When we say, “George IV wished to know whether Scott was the author of *Waverley*”, we normally mean “George IV wished to know whether one and only one man wrote *Waverley* and Scott was that man”; but we may also mean “One and only one man wrote *Waverley*, and George IV wished to know whether Scott was that man.” In the latter, “the author of *Waverley*” has primary occurrence; in the former, a secondary. The latter might be expressed by “George IV wished to know, concerning the man who in fact wrote *Waverley*, whether he was Scott.” *This would be true, for example, if George IV had seen Scott at a distance, and had asked “Is that Scott?”*\footnote{Russell (1905/2000), pp. 217-8. Emphasis added.}

What is striking about this passage is this: When George IV wonders, ‘Is that Scott?’, he seems precisely to be wondering whether an (informative) identity of the genuine logical form of ‘*a is b*’ is true or false. Indeed, Russell consistently treats ‘Scott’ as if it were a genuine name throughout “On Denoting” (including in the first sentence of the cited paragraph), and he always holds that the demonstrative ‘that’ is a genuinely singular
referential device throughout his work. Moreover, there seems to be no question of whether George IV is acquainted with Scott given the points made in the previous paragraph. But George IV is manifestly unable to determine on the basis of experience alone whether or not the object of his occurrent perceptual acquaintance is Scott. So if it turns out that George IV is in fact seeing Scott in the distance (which is a live possibility), we not only have a clear instance of Frege’s Puzzle, we also have a decisive counterexample to the received view of acquaintance.

Given the points made in the previous three paragraphs, the received view of acquaintance seems to be in serious trouble. For rather than vindicating the received view, careful and charitable readings of “On Denoting” seem to tell against it. As a matter of fact, if we are to extend charity to Russell on any point, we seemingly should do so concerning his overly-strong remarks that the Theory of Descriptions explains the usefulness of identity. Indeed, Russell’s own case of George IV wondering whether ‘that is Scott’ seems to provide a clear case in which identity is “worth affirming” even though it makes no use of the Theory of Descriptions (or the Descriptive Theory of Ordinary Names). In all likelihood, Russell is simply overselling the virtues of his Theory of Descriptions to those antecedently inclined to think that there are no identities worth affirming (outside of logic-books). At any rate, what should be clear is that “On Denoting” provides scant evidence that Russell ever accepts the received view of acquaintance, or intends for his Theory of Descriptions to solve all instances of Frege’s Puzzle.

105 Even if we press the point that George IV’s thought or utterance is actually of the form ‘Is that the unique object named “Scott”’, there is no reason why he couldn’t wonder whether ‘this is that’ when deploying two co-referring demonstratives, as in Perry’s case of the dog Stretch. See Perry (2001), Chapter 4 Section 5.
One interesting (but to my knowledge unexplored) interpretive question is how much interest Russell has in providing a general solution to Frege’s Puzzle. Though Russell clearly has interest in failures of truth-preservation when co-designating genuine names and denoting phrases are substituted for one another in attitude ascriptions, his principal interests in discussing Frege’s sense-reference distinction in his works leading up to “On Denoting” concern the issues of reference-failure and getting the satisfaction conditions of our thought and talk right, with an eye to solving philosophical problems arising in logic and mathematics.\footnote{The relevant unpublished papers include Russell (1903b/1994), (1903c/1994), (1903d/1994), (1903e/1994), and (1905/1994). A relevant published work is Russell (1903a/1929).} He doesn’t appear to have independent interest in Frege’s Puzzle, or in issues of informative identities involving genuine names. In fact, in *The Principles of Mathematics* (1903a/1929), Russell shows remarkable ambivalence towards instances of Frege’s Puzzle involving genuine names when he flatly says:

> [Frege’s] theory of indication is more sweeping and general than mine, as appears from the fact that *every* proper name is supposed to have two sides [that is, sense and reference]. It seems to me that only such proper names as are derived from concepts by means of *the* can be said to have meaning, and that such words as *John* merely indicate without meaning.\footnote{Russell (1903a/1929), p. 502.}

But ordinary names for persons such as ‘John’ are just the sorts of expressions for which Frege’s Puzzle naturally arises, a point of which Russell is surely aware given that he is explicitly discussing Frege’s “Über Sinn and Bedeutung” (1892/2000). So it is very interesting that he makes no attempt to offer (or even gesture at the need for) a solution to the problem of informative identities either in the cited passage, or in his preceding and following remarks.

What’s more, it is quite striking that Russell never (at least to my knowledge) explicitly touts his Theory of Descriptions as even a partial solution to Frege’s Puzzle in
any of his subsequent discussions of the George IV case (or of Frege’s sense-reference distinction). A satisfactory investigation of the true extent of Russell’s interest in Frege’s Puzzle would require far more space than I have here. But I suspect that the answer to this question would be quite illuminating for our understanding of “On Denoting”, and for our understanding of Russell’s views about meaning and knowledge in general.

For a few examples of Russell’s subsequent discussions of George IV, see Russell (1910/1929), pp. 224-9, Russell (1918, 1919), Part VI, Russell (1919/1963), pp. 173-80, and Russell (1959), pp. 84-5. One might object that while Russell never explicitly touts his Theory of Descriptions as even a partial solution to Frege’s Puzzle, he nevertheless plausibly has Frege’s Puzzle squarely in mind as a target in his (1910/1929) “Knowledge by Acquaintance and Knowledge by Description”. After all, immediately after he concludes that the descriptive phrase “the man whose name was Julius Caesar” in the judgment expressed by “the man whose name was Julius Caesar was assassinated” must be broken up “if we are to be acquainted with all the constituents of the judgment”, he says that “this conclusion, which we have reached from considerations concerned with the theory of knowledge, is also forced upon us by logical considerations”, at which point he proceeds to discuss the alleged distinction between “meaning and denotation”. But then, one might reason, isn’t Russell at least suggesting that his views about restricting the constituents of thought to the objects of acquaintance are motivated at least partly by Frege’s Puzzle? And doesn’t Russell at least then restrict acquaintance to sense-data, abstracta, and (perhaps) the self? So why not simply conclude that Russell holds the received view of acquaintance in his 1910 paper?

While this line of reasoning might seem initially plausible, it does not hold up to scrutiny. For one thing, the conclusion that Russell in claiming to have reached from considerations of the theory of knowledge is that the phrase “the man whose name is Julius Caesar” must be broken up if we are to be acquainted with all of the constituents of judgments expressed by utterances containing it. And the basis for this conclusion is that we are not presented in experience with Julius Caesar, and so we can’t be acquainted with the person denoted by “the man whose name is Julius Caesar”. But nothing here conflicts with my interpretation of Russell above.

Moreover, when Russell turns his attention to the logical considerations for breaking up the descriptive phrase into quantificational complexes, what he wants to show is that identity statements involving descriptive phrases are assigned the wrong satisfaction conditions whether we take them to contribute their meanings or their denotations. Indeed, he argues, the meanings (or senses) of “the author of Waverley” and “the author of Marmion” are clearly not identical, so they can’t be what is contributed to the proposition expressed by “the author of Waverley is the author of Marmion”. But, he contends, the denotations of these two descriptive phrases can’t be what is contributed either. For one thing, he says, we can understand “the author of Waverley is the author of Marmion” without be acquainted with the denotation of either one, i.e. Scott. For another, it raises the very problems we considered on pages 26-7 above concerning descriptive phrases that lack denotation: that it either makes propositions that intuitively have truth-values truth-valueless, or it ushers in Meinogian objects (and all accompanying contradictions).

Finally, Russell argues, if descriptive phrases simply contributed their denotations, then “Scott is the author of Waverley” would be roughly analogous to “Scott is Sir Walter”, which Russell takes to state that “these are two names for one man”. But, he objects, “a man’s name is what he is called, but however much Scott has been called the author of Waverley, that would not made him be the author; it was necessary for him actually to write Waverley, which was a fact having nothing to do with names”. So the problem Russell sees is not merely that “Scott” and “the author of Waverley” differ in cognitive significance; they contribute different kinds of elements to propositions (as we saw on pages 31-2 above). But nowhere in Russell’s 1910 paper does he say anything which suggests that he intends to solve Frege’s Puzzle in full generality, or anything which supports the received view of acquaintance.
§6 CONCLUSION

In this chapter, I have shown that important theorists, such as Gareth Evans, Donald Davidson, John Campbell, and Ian Proops, have attributed to Russell what I call ‘the received view of acquaintance’: the view that acquaintance safeguards us from misidentifying the objects of our acquaintance. I have also shown that Russell’s discussions of phenomenal continua cases show that he does not accept the received view, or “Russell’s Principle” strongly interpreted. Indeed, I have argued, the possibility of misidentifying the objects of acquaintance follows from underappreciated aspects of Russell’s overall theory of knowledge and acquaintance. Finally, I have discussed the radical impact that Russell’s actual views on acquaintance have for our understanding of his well-known George IV case in “On Denoting”. In particular, I have argued that the George IV case is seemingly not intended to be a one-size-fits-all solution to Frege’s Puzzle (and similar substitution problems), and hence does not provide any compelling support for the received view of acquaintance.

The importance of this finding is not merely a historical one, however. Indeed, much recent work in the philosophy of mind, the philosophy of language, and epistemology takes Russell’s notion of acquaintance as a key starting point. But if I am correct in interpreting Russell’s actual views on acquaintance, then much of this contemporary work rests on a mistake. For one thing that is clear is that if we are to follow Russell’s lead we must reject “Russell’s Principle”, as strongly interpreted, wholesale. Is that plausible? Just how far can we depart from the strong readings of “Russell’s Principle”? We will see in the following chapters.
CHAPTER 2:
PERCEPTUAL ACQUAINTANCE AND THE EXPLANATORY ROLE OF EXPERIENCE

“On a Relational View, the qualitative character of the experience is constituted by the qualitative character of the scene perceived. I will argue that only this view, on which experience of an object is a simple relation holding between perceiver and object, can characterize the kind of acquaintance with objects that provides knowledge of reference.” – John Campbell, *Reference and Consciousness*¹⁰⁹

“Phenomenal character is manifest to us in our being aware of…external qualities. We cannot focus on it in any way that separates it from our focus on external things and qualities… On this view, the phenomenal character of the experience of red in a case of veridical perception is a feature of the surface the perceiver sees. The surface has the phenomenal character.” – Michael Tye, *Consciousness Revisited: Materialism without Phenomenal Concepts*¹¹⁰

§1 INTRODUCTION

Consider the case in which you are looking at a ripe red apple in standard viewing conditions. What is it that you are aware of? Intuitively, one is aware of the apple and its quality of redness. One is also in some sense aware of the quality of one’s experience of the ripe red apple; one is aware of “what it’s like” to perceive it visually in standard viewing conditions. What is the relation between these two things one is aware of, that is, between the qualitative character of one’s experience of seeing the apple and the redness of the apple itself? On the naïve realist view of perception, they are the same: the qualitative character of the experience of seeing red just is the redness that inheres in the surface of the apple. In the case of veridical perception, one is aware of the redness in the

¹¹⁰ Tye (2009), p. 120.
surface of the apple, and nothing else. On the indirect realist view of perception, on the other hand, what one is aware of directly is the qualitative character of one’s own mind-dependent sensation as of red, and one explicitly or implicitly infers the redness of the apple itself.

Many currently working on a Russellian notion of perceptual acquaintance and its role in perceptual experience tend to treat naïve realism and indirect realism as an exhaustive disjunction of possible views. In this chapter, I will articulate a third, intermediate view. This view of perceptual acquaintance is a form of direct realism, according to which one directly and literally sees the apple and its redness without seeing something mind-dependent and without making any intermediate inference. Nevertheless, it also maintains that the qualitative character of perceptual experience is a mind-dependent feature of our internal states of sentient awareness, and so is to be distinguished from the redness of the apple. Indeed, I believe that only this combination of direct realism and qualia internalism can provide an adequate characterization of our perceptual experience of things.

To foreshadow things a bit, on my picture of perceptual experience we are aware of the qualitative character of our sensations simply in virtue of having them, and we are non-inferentially aware of the apple and its redness by being attuned to what the occurrence of our sensations tells us about the rest of the world. Consequently, we are presented with, and thus acquainted with, both the apple (and its redness) and the qualitative character of our sensation of the apple, albeit in very different ways. Indeed,

---

111 I have in mind here Campbell (2002a, 2002b, and 2009) and Tye (2009). I will say more about how exactly we should understand the technical notion of ‘perceptual acquaintance’ in §2 of this paper.
112 Qualia internalism is the view that any duplicate of a subject will have the same qualitative experience regardless of the environment that he or she is in. Qualia externalism, on the other hand, is the view that the qualitative characters of experiences are constituted by features of the external environment.
perceptual experience has what I call a “two-faced presentational character”: we are presented all at once with external objects (and their perceptible features) and with the qualitative character of our perceptual experience itself.

My view here particularly contrasts with recent work by John Campbell and Michael Tye on Russell’s notion of acquaintance and its role in perceptual experience, which has tended to emphasize the naïve realism of early Russell (and G. E. Moore).¹¹³ For example, Campbell has recently claimed:

On a Relational View, the qualitative character of the experience is constituted by the qualitative character of the scene perceived…only this view, on which experience of an object is a simple relation holding between perceiver and object, can characterize the kind of acquaintance with objects that provides knowledge of reference.¹¹⁴

Similarly, Tye has argued that:

Phenomenal character is manifest to us in our being aware of…external qualities. We cannot focus on it in any way that separates it from our focus on external things and qualities… On this view, the phenomenal character of the experience of red in a case of veridical perception is a feature of the surface the perceiver sees. The surface has the phenomenal character.¹¹⁵

On such views, when we have perceptual acquaintance with things, what we directly and literally perceive are external objects and their perceptible features, where the qualitative characters of our perceptual experiences inhere in, or are constituted by, those external objects and features.

There are, however, a number of well-known objections to naïve realism. First of all, our current best scientific theories of perception subscribe to the view that the right sort of neurobiological states of the brain are (at least) causally sufficient for the

¹¹³ This recent work includes Campbell (2002a, 2002b, and 2009) and Tye (2009).
¹¹⁵ Tye (2009), p. 120.
occurrence of a perceptual experience as of external objects and features, even when there are no external objects present.\textsuperscript{116} And secondly, in cases of illusion or hallucination, subjects have the falsidical experience as of an external object having such-and-such features when either the object does not have those features, or there is no object whatsoever. Together, these objections suggest that the qualitative phenomenal features of which we are aware in perceptual experience are mind-dependent features of the experiential states themselves rather than mind-independent features of external objects.

Campbell and Tye are unconvinced by these arguments, however. They urge that we should embrace some variety of disjunctivism, the view that there is no “common factor” of phenomenal character or of content between veridical perceptions and the cases of illusion or hallucination.\textsuperscript{117} On such a disjunctivist view, in cases of veridical perception the perceived object or features are constituents of the perceptual experience itself and the qualitative characters of which we are aware inhere in them rather than in our internal states of sentient awareness. However, in cases of non-veridical perception or hallucination, the subject has an entirely distinct \textit{kind} of perception-like experience, one with a different phenomenal character and different content, yet one that the subject cannot introspectively discern to be different from the veridical case, despite the fact that it is.

As we will see, the move to disjunctivism is unpromising. In particular, it is difficult to see how disjunctivism can provide an epistemologically and metaphysically

\textsuperscript{116} I say “(at least) causally sufficient” because the relation between the relevant neurobiological states of the brain and the conscious perceptual experience will be even more intimate on any physicalist ontology.

\textsuperscript{117} In point of fact, Tye thinks that there \textit{is} a common factor of phenomenal character and of content schema, but no common factor of singular content. However, as we will see, it is hard to square this view with his professed qualia externalism.
plausible account of non-veridical perceptual experiences. I think we need to take a very different route in developing an account of perceptual acquaintance. I here argue that once we recognize the “two-faced presentational character” of perceptual experience, we can safely avoid the prima facie dilemma between naïve realism and indirect realism. Indeed, once we fully recognize the informational character of perceptual experience, we can see how our perceptual acquaintance puts us in direct experiential contact with external objects, features, and states-of-affairs even though the qualitative characters of our experiences are mind-dependent features of our own states of sentient awareness. Thus, we can have the best of both worlds; we can be direct realists about the content of perceptual experience and internalists about its qualitative character.

In this chapter, I proceed as follows: In §2, I briefly review the main lessons we learned in the previous chapter about how Russell conceived of acquaintance and its role in cognition, reference, and perception. In §3, I explain how Campbell and Tye work out their respective disjunctivist accounts of perceptual acquaintance. In §4, I criticize disjunctivism for failing to account adequately for phenomenal indiscriminability and for failing to deliver on Campbell’s explanatory demand on experience. In §5, I present John Perry’s theory of information and informational content, which I think is a useful framework for thinking about naturalized content in general. In §6, I explain how this theory of information and informational content helps us see how we can be direct realists about the content of perceptual acquaintance. Finally, in §7, I draw on recent work on perceptual intentionality by Searle to develop my own account of the “two-faced

---

118 It is worth noting that Russell’s own theorizing about the nature of perceptual acquaintance was plagued by this false dilemma. One goal of this paper is to show that we needn’t follow the early Russell in embracing naïve realism and its problematic qualia externalism, nor the later Russell in combining qualia internalism with an undesirable indirect theory of perceptual experience.
presentational character” of perceptual experience, which helps explain how we can be internalists about its phenomenal character while being externalists about its content.

§2 KNOWLEDGE OF THINGS AND KNOWLEDGE OF TRUTHS II

Recent interest in a Russellian notion of perceptual acquaintance has been fueled by the increased appreciation of Russell’s more general distinction between two kinds of knowledge: knowledge of things and knowledge of truths. Russell’s distinction here roughly corresponds with an ordinary distinction drawn in many natural languages between two uses of the term ‘know’. First of all, we ordinarily talk about knowing that such-and-such is the case. This sort of knowledge paradigmatically involves conceptually articulated beliefs and is propositional in character; it involves cognitive attitudes that can be evaluated in terms of their truth or falsity. However, there is also another way in which we use the term ‘know’ in ordinary language: we talk of knowing individuals in the sense that we have directly encountered them and therefore have some familiarity with them. This is roughly what Russell has in mind with his technical notion of knowledge of things, though it is not entirely equivalent to the ordinary one.

According to Russell, our knowledge of things is a distinctive epistemic relation that plays a fundamental role in reference and cognition. Indeed, Russell introduces the notion of knowledge of things, or acquaintance, in order to place a substantive cognitive constraint on what it takes for an individual to have genuinely singular thoughts about, or to make genuinely singular reference to, individuals. In other words, our knowledge of things explains what it takes to have thoughts or utterances whose contents are constituted by the very individuals and features, if any, that the thoughts or utterances are
about rather than by some way of getting at or identifying them. In fact, he argues that all cognition and linguistic designation ultimately rests on our fundamental epistemic capacity to be acquainted with, or consciously aware of, individuals and their features.

Russell’s basic picture of cognition and reference is this: There is some special class of individuals (concrete and abstract alike) and their features with which we have experiential contact. On the basis of this experiential contact, we are in a position to direct conscious attention to those individuals and features. Our conscious attention to these experienced individuals and features in turn grounds our ability to make demonstrative, singular reference to them (to designate them with a “logically proper name”) and puts us in a position to acquire knowledge of such demonstrative, singular reference. Furthermore, our conscious attention to experienced individuals and features also puts us in a position to pick up information about, and thereby form some conception of, these individuals and features. Finally, on the basis of our conceptions of these individuals and features and our more general conception of how the world works, we can designate objects and features with which we lack experiential contact. We do so by employing representations that encode identifying conditions that an object or feature must uniquely satisfy in order to be the designated individual or feature of the representation, or by employing representations that are purely quantificational in character.

But what exactly is acquaintance? Basically, Russell conceives of acquaintance as a fundamental experiential relation between a conscious subject and individuals or features in the world, where the ‘knowledge of things’ that it constitutes is logically independent of, but serves as the ultimate enabling condition for, our knowledge of truths.

---

119 I am being shamelessly terminologically anachronistic here, but I think doing so is helpful.
about, and conceptions of, the things or features with which we are acquainted. As Russell puts it:

Knowledge of things, when it is of the kind we call knowledge by acquaintance, is essentially simpler than any knowledge of truths, and logically independent of knowledge of truths, though it would be rash to assume that human beings ever, in fact, have acquaintance with things without at the same time knowing some truth about them.\textsuperscript{120}

For Russell, acquaintance is relational in the straightforward sense that it is a relation that obtains between a conscious subject and the individuals or features with which the subject is acquainted; both the subject and the individuals or features are constituents of the acquaintance relation.\textsuperscript{121} He holds that acquaintance is fundamental in that:

[Acquaintance is] the simplest and most pervading aspect of experience... All cognitive relations—attention, sensation, memory, imagination, believing, disbelieving, etc.—presuppose acquaintance.\textsuperscript{122}

Indeed, he holds that our acquaintance with things constitutes our most basic form of intentionality, or object-directedness, and grounds our ability to make genuine singular reference to the individuals and features with which we are acquainted. He contends that:

The faculty of being acquainted with things other than itself is the main characteristic of a mind. Acquaintance with objects essentially consists in a relation between the mind and something other than the mind; it is this that constitutes the mind’s power of knowing things.\textsuperscript{123}

Russell is quite clear in holding that acquaintance is essentially an experiential relation in which the conscious subject is presented with the objects of his or her awareness. In his 1913 \textit{Theory of Knowledge}, he remarks:

Now, since we have decided that experience is constituted by a relation, ...we shall employ synonymously the two words ‘acquaintance’ and ‘awareness’,
generally the former. Thus when $A$ experiences an object $O$, we shall say that $A$ is acquainted with $O$.  

Elsewhere, Russell asserts that:

The distinction between acquaintance and knowledge about is the distinction between the things we have presentations of, and the things we only reach by denoting phrases.

For Russell, an individual or feature is presented to a subject just in case it makes the right sort of difference to that subject’s conscious life such that the subject is aware of it.

Finally, Russell holds that the knowledge of things of which acquaintance consists is unreflective, non-conceptual, and non-propositional. Knowledge of things is “non-conceptual” in the sense that it does not require a subject’s having antecedent possession of any concept or conception of the thing or feature with which he or she is acquainted. Indeed, acquaintance is precisely the sort of unreflective and ‘objectual’ conscious awareness of which many non-linguistic sentient animals and small children are capable.

Knowledge of things is “non-propositional” in the straightforward sense that it lacks genuine veridicality conditions. Indeed, on Russell’s view, either a subject experiences such-and-such individuals or features, or the subject does not. It is in this sense (alone) that we know a thing “perfectly and completely” when we are acquainted with it. Of course, it goes without saying that the epistemic mechanisms and faculties through which we become acquainted with things are themselves fallible in that they can misrepresent the individuals and features we are experiencing. But acquaintance itself does not admit of error; it is all or nothing.

---

127 Note that Russell’s view here should not be understood as an endorsement of the “thesis of revelation”. 

§3 PERCEPTUAL ACQUAINTANCE AND DISJUNCTIVISM

Recently, Campbell and Tye have independently offered accounts of perceptual acquaintance that share many of the features noted in the previous section. However, they differ importantly from my account of perceptual acquaintance in that they both accept some version of disjunctivism. In this section, I present their respective accounts and say more about why they are led to embrace disjunctivist views about perceptual experience.

To anticipate things slightly, Campbell and Tye embrace disjunctivism in order to account for illusion and hallucination while respecting the prima facie intuitiveness of the “transparency of experience”.

To begin with, Campbell and Tye, like many recent naïve realists, are greatly moved by G. E. Moore’s remarks about the nature of sensation in his 1903 “The Refutation of Idealism”. In his classic article, Moore remarks,

[There are] in every sensation two distinct terms, (1) ‘consciousness,’ in respect of which all sensations are alike; and (2) something else in respect of which one sensation differs from another. It will be convenient if I may be allowed to call this second term the ‘object’ of a sensation: this also without yet attempting to say what I mean by the word.

We have then in every sensation two distinct elements, one of which I call consciousness, and another which I call the object of consciousness.¹²⁸

Many have found this passage interesting for at least two reasons. First of all, Moore draws in it the well-known distinction between the act of experience and the object, or objects, of the experience, where the act of experience is best thought of as a conscious relation between the subject of the experience and the object experienced (i.e. an acquaintance relation). And secondly, Moore contends that we can explain the differences between the characters of distinct conscious experiences by appealing wholly

¹²⁸ See Moore (1903), p. 444.
to differences in what we are conscious of, rather than to intrinsic differences in the sensations themselves. Indeed, he continues,

[When] we refer to introspection and try to discover what the sensation of blue is, it is very easy to suppose that we have before us only a single term. The term "blue" is easy enough to distinguish, but the other element which I have called "consciousness" -- that which sensation of blue has in common with sensation of green -- is extremely difficult to fix. That many people fail to distinguish it at all is sufficiently shown by the fact that there are materialists. And, in general, that which makes the sensation of blue a mental fact seems to escape us: it seems, if I may use a metaphor, to be transparent -- we look through it and see nothing but the blue … \(^\text{129}\)

On the standard reading of Moore, what he claims here is that our experience is “diaphanous”; in experience we are presented with aspects of the world rather than with intrinsic aspects of our experiences, if they have any. \(^\text{130}\) In other words, we see straight through our experiences to perceptible mind-independent objects and feature-instances in the external scene. Thus, if two experiences differ with respect to their characters, this difference is to be explained entirely by appeal to differences in the objects in the external scene that they are experiences of.

Moore’s remarks above are motivated in part by a crucial bit of purported phenomenological data: it seems that, in many cases, when we try to attend to the qualitative character of perceptual experience, we do so precisely by attending to the scene that our experience, when veridical, is an experience of. Moreover, when we engage in such phenomenological introspection, the argument runs, it is not as if we are aware of two qualities, one inhering in the attended-to object, and one intrinsic to the experience itself. Rather, the qualitative character is exhausted by the character of what

\(^{129}\) See Moore (1903), p. 446.

\(^{130}\) Detractors of the standard reading point out that Moore says only that it is “extremely difficult” to specify what all conscious sensations have in common and not that it is impossible. Ned Block has made this point in recent talks.
the experience is as of. So, Moore reasons, given that in veridical perception we are
directly aware of the perceptible qualities of the external scene, it follows that these are
the *only* qualities of which we are directly aware in perceptual experience.\(^\text{131}\) And so any
differences in the qualitative characters of different perceptual experiences are to be explained by differences in the qualities of the perceived external scene.

As I’ve noted, Campbell and Tye find Moore’s “radical transparency” very compelling, and hence they wed their accounts of perceptual acquaintance with the naïve realist theory of perception that Moore’s comments imply. However, they are also very aware of the problems that illusions and hallucinations pose for Moore’s transparency thesis. For in cases of illusion, we experience an object as having features that it lacks. And in cases of hallucination, we have the experience as of an object and its perceptible features, despite the fact that there is no external object to figure into the singular content of the experience. But then how can the qualitative characters of such non-veridical experiences inhere in the external perceptual scene? And if they do not inhere in the external scene in the non-veridical case, then why suppose that they do in the veridical case? For this reason, illusions and hallucinations are a prima facie threat to the naïve realism to which both Campbell and Tye are greatly sympathetic.

Campbell and Tye do not think that illusions and hallucinations ultimately undermine naïve realism, however. They argue that the reasoning above depends on the unwarranted assumption that veridical and non-veridical perceptual experiences

---

\(^{131}\) As we’ll see, Moore’s reasoning here is a little too quick; the same purported phenomenological data can be used to support the view that the qualitative characters of our experiences are exhausted by intrinsic features of our own experiential states. But this, I will argue, is not incompatible with the view that the objects of our experiences are external ones. It is also worth noting that Moore himself later gave up his naïve realism for a sense-data theory of perception due to cases such as blurry vision and double-vision, where changes in the qualitative character of a subject’s experience are not generally taken to be changes in the external scene perceived.
constitute a single, natural explanatory kind. However, they argue, the class of such experiences is in fact a disjunctive class, “more like the class of tigers or tables”, as Tye puts it.\(^{132}\) In the case of veridical perception, the object perceived enters into the singular content of the perceptual experience and its qualitative characters are constituted by the objects and features in the external scene. But in the case of non-veridical perceptual experiences, we have a fundamentally different kind of experience, one that has no “common factor” of phenomenal character or of content with the veridical case.\(^{133}\) Consequently, we cannot make legitimate inferences from the nature of non-veridical perceptual cases to the nature of veridical perceptual cases. This is, of course, a version of disjunctivism about perceptual experience.

Campbell and Tye admit, however, that veridical and non-veridical perceptual experiences can in principle seem the same to a subject upon introspection, despite the fact that they are different. As Tye (following Austin) illustratively puts it,

> The veridical experience and the hallucinatory experience are no more closely related than a lemon and a bar of soap that looks just like a lemon…The one experience is indistinguishable from the other via introspection, just as the lemon and the soap are indistinguishable from one another perceptually. Even so, the two are very different kinds of thing.\(^{134}\)

However, Campbell and Tye have slightly different explanations for why we cannot tell apart the veridical case from the non-veridical cases of perceptual experience in introspection. According to Campbell, veridical and non-veridical cases of perceptual experience are sufficiently similar in qualitative character such that they are subjectively indiscriminable. Nevertheless, Campbell maintains that the qualitative characters of the

---

\(^{132}\) Tye (2009), p. 78.

\(^{133}\) Strictly speaking, Tye thinks that there is a common factor of phenomenal character and of content schema, but no common factor of singular content. As we will see, it is hard to square this view with his professed qualia externalism, however.

\(^{134}\) Tye (2009), p. 178.
two cases are objectively different. According to Tye, in contrast, the reason why we cannot tell apart veridical and non-veridical cases of perceptual experience is straightforwardly because they have the same qualitative characters (and what he calls ‘content schemas’).\textsuperscript{135} What distinguishes the two cases, Tye contends, is not their qualitative characters, but rather whether or not the experiences have singular contents. In the rest of this section, I say more about Campbell’s and Tye’s respective views about perceptual acquaintance and how they aim to accommodate cases of illusion and hallucination.

To begin with, Campbell’s so-called ‘Relational View’ of perceptual acquaintance is driven by two related questions. First, what roles do experience and conscious attention play in our knowledge of things and in our knowledge of reference? And second, what is the nature of experience such that it can play the role that it does? His answer to the first question is that our experience of (and attention to) objects provides us with our knowledge of reference and explains how it is that we come by our conceptions of objects as categorical, mind-independent things. I will say much more about Campbell’s claims here in the next section of this paper. But for now, I will focus on his answer to the question about what the nature of perceptual acquaintance is as well as what he says about cases of illusions and hallucinations.

On Campbell’s view, perceptual acquaintance is a primitive three-place conscious relation between a subject, a standpoint, and an external scene. It is a primitive relation in the sense that it constitutes our knowledge of reference and is an enabling condition for our knowledge of truths, and it is relational in the sense that the subject, the standpoint,

\textsuperscript{135} Indeed, on page 113 he remarks, “The question now becomes ‘Why are [subjects] not conscious of any such difference?’ Surely the most straightforward and natural answer is that in some cases there is no difference; the phenomenal character is the same.”
and the external scene are all constituents of the relation (or relational fact). Moreover, he
thinks that our acquaintance with things is also primitive in the sense that it is irreducible
to causal relations between external objects and neuronal firing or information-processing
in the brain, though he takes it that acquaintance is (in part) causally sustained by
physiological processes in the brain.

Campbell does not fully spell-out how we are understand the notion of a
standpoint, but he thinks that we have a working sense of the sensory modality-specific
parameters (such as egocentric location, relative orientation, constancies, obstructions,
temporal dynamics, etc.) that constitute one. Indeed, he says,

Suppose you want to characterize the view that someone currently has of the Taj Mahal. We say which thing it is and which person is in question. Then to
describe the standpoint explicitly we have to say which sensory modality is
involved; and that will determine what further factors we have to fill in. For
example, suppose the modality is vision. Then we need, further, position, but
also the relative orientations of the viewer and object, how close the viewer is to
the object, whether there is anything obstructing the light between them, and so
on. In the case of hearing, a rather different set of factors would be relevant: not
just which object was in question, but what sounds it was making, and the
obstruction of light would not be to the point, though the obstruction of sound
would be. We do not usually spell out all these conditions, though we are
perfectly capable of articulating them when they are important in particular
cases.\footnote{Campbell (2009), pp. 20-1.}

In any case, what is crucial to the notion of a standpoint is that we always experience an
object from some standpoint or other, and differences in the standpoint from which we
experience an object can explain differences in the cognitive significance and
phenomenal character of our experience of it.\footnote{For more on standpoints, see Campbell (2009), pp. 18-22.}
Though the character of our experience of objects is always relative to some standpoint or other, Campbell nevertheless maintains that their phenomenal characters inhere in the external scene itself. As he puts it,

On a Relational View, the phenomenal character of your experience, as you look around the room, is constituted by the actual layout of the room itself: which particular objects are there, their intrinsic properties, such as colour and shape, and how they are arranged in relation to one another and to you.\textsuperscript{138}

What’s more, he continues:

On the Relational View, two ordinary observers standing in roughly the same place, looking at the same scene, are bound to have experiences with the same phenomenal character. For the phenomenal character of the experience is constituted by the layout and characteristics of the very same external objects.\textsuperscript{139}

What is clear from these passages is that Campbell sees perceptual acquaintance as a transparent relation that puts us in a position to experience, from a standpoint, the qualitative features of the external scene. He remarks that:

On a Relational View of perception, we have to think of cognitive processing as ‘revealing’ the world to the subject; that is, as making it possible for the subject to experience particular external objects…we have to think of experience of objects as depending jointly on the cognitive processing and the environment. Experience, on this view, cannot be understood simply as a matter of cognitive contents becoming subjectively available.\textsuperscript{140}

But if the qualitative characters of veridical perceptual experiences are features of the external scene, then how exactly does Campbell account for perceptual illusions and hallucinations? As I’ve already noted, Campbell simply denies that veridical and non-veridical perceptual experience constitute some unified, natural kind. He argues that:

On the Relational View…there is nothing intrinsic in common between the case in which there is a dagger to which you are consciously attending, and the case in which you are just having a hallucination. In the case in which there is a

\textsuperscript{139} Campbell (2002a), p. 116.
\textsuperscript{140} Campbell (2002a), p. 118.
dagger, the object itself is a constituent of your experience. The experience is quite different in the case of the hallucination, since there is no object to be a constituent of your experience.\footnote{\textit{Campbell} (2002a), p. 117.}

Unfortunately, Campbell does not make clear exactly how we are to treat the non-veridical cases except in saying that they are a fundamentally different sort of experience. What he does say, however, is that though the non-veridical cases seem to be the same with respect to their phenomenal character as the veridical case, they differ in some subjectively indistinguishable respect. There is, on his view, no “common factor” between the veridical and non-veridical perceptual experiences.

Like Campbell, Tye also thinks that veridical and hallucinatory experiences constitute a disjunctive category of perceptual experiences.\footnote{Tye has recently changed many of his views about what I discuss in what follows.} His reason for claiming this is that while veridical perceptual experiences have singular contents into which the objects of perception enter, hallucinatory experiences do not have singular contents precisely because there are no objects corresponding to them. In contrasting the two cases, Tye says:

Consider the china frog that is sitting by one of my house plants. As I view it, I think to myself that \textit{that} is a china frog. A little later, I look up from the book I am reading and again think to myself that \textit{that} is a china frog. Unknown to me, a mischievous demon has made the frog disappear in the intervening time period while making it appear to me that the frog is still present. In the second case, then, there is no frog for my thought to be about…My first thought, as I looked at the china frog, had a singular content, into which the china frog entered. The second thought did not have a singular content.\footnote{Tye (2009), p. 81.}

But though Tye agrees with Campbell that veridical and hallucinatory experiences are sufficiently different so as to constitute a disjunctive kind, he thinks nevertheless that there is \textit{some} common factor between the veridical and non-veridical cases of perceptual
experience. In particular, he thinks there is some element of content and phenomenal character in common between the two cases that explains why a subject can mistake one case for the other. Of course, given what he says about hallucinatory cases lacking singular content, the element in common between the two cases cannot be singular content but rather some other variety of content.

One natural proposal at this point is that what the two cases have in common is existential content to the effect that there is some object that is thus-and-so. However, Tye quickly rejects this proposal for failing to do justice to the apparent singularity of many hallucinatory experiences. For as he notes, hallucinatory experiences are often as of particular objects and features, despite the fact that those objects are not, on that occasion, actually perceived. This point, though not decisive against the existential thesis about perceptual content, suggests that if we are to locate an element of content in common between the veridical and hallucinatory cases of perceptual experience, we must look to something other than existential content.

In any case, Tye’s own proposal is that veridical and non-veridical perceptual experiences share what he calls a ‘singular (when filled) content’ or ‘SWF content’. So what, according to Tye, is SWF content? Returning to the case of the china frog, he says:

The natural proposal, I suggest, is that the second content [i.e. of the hallucination] is just like the first except where the first has a concrete object in it, the second has a gap. The two contents, thus, have a common structure. This structure may be conceived of as having a slot in it for an object. In the case of the first content, the slot is filled by the china frog. In the case of the second content, the slot is empty. I shall call such structures content schemas. On Tye’s view, A SWF singular content is a content schema that yields a singular content when its ‘slots’ are filled. According to Tye, in the case of a veridical perceptual

---

144 Tye (2009), p. 81.
experience, the objects that fill the slots have the perceptible features that the experience represents them as having. In the case of an illusory experience, the objects that fill the slots lack one or more of the perceptible features that the experience represents them as having. And in the case of a hallucinatory experience, one or more the slots of the content schema are simply empty.

In Tye’s estimation, SWF singular contents can usefully be conceived of “on the model of Russellian singular propositions having slots in place of objects”. When an object is experienced, the SWF content schema yields a Russellian singular proposition. When no object is experienced, the SWF singular content yields a ‘gappy proposition.’ When combined with the view that the experience is veridical if and only if the objects experienced have the properties the experience represents them as having, he argues, we get the correct result that hallucinatory experiences are non-veridical, precisely because there are no experienced objects.

Unfortunately, Tye does not provide a metaphysical characterization of SWF content schemas. Nevertheless, he does make a few remarks that suggest constraints on what they might be. First of all, Tye contends that SWF content is a species of non-conceptual content; it is content whose conditions of satisfaction needn’t be reflectively grasped by the subject of such experiences, nor need the subject antecedently possess the concepts of the objects and features represented by it.\(^{145}\) Secondly, SWF content must be such that the phenomenal character of experience supervenes on it. This fact explains why it is that a subject cannot tell apart veridical and non-veridical cases of perceptual experience. Thirdly, SWF content must be such that when a perceptual experience is

\(^{145}\) Tye (2009), p. 103.
veridical, it reveals the qualitative characters of the external scene to the subject. For, Tye contends,

Phenomenal character is manifest to us in our being aware of…external qualities. We cannot focus on it in any way that separates it from our focus on external things and qualities… On this view, the phenomenal character of the experience of red in a case of veridical perception is a feature of the surface the perceiver sees. The surface has the phenomenal character.\(^\text{146}\)

And lastly, in the case of the non-veridical perceptual experiences, however, the subject is presented with un-instantiated qualitative characters. While many might find this contention simply incredible, Tye finds it quite natural. He remarks that:

Perhaps it will be objected that this view is committed to the implausible claim that there are un-instantiated properties (for example, redness) in cases of hallucination. To this I reply that the claim is not in the least implausible. On the contrary, it is part of naïve commonsense. Suppose that you had never seen any red things and then, one day, you hallucinated a red car. Did you not then encounter redness in your experience?\(^\text{147}\)

In the next section of this chapter, I will evaluate Tye’s and Campbell’s views on perception and argue that they are deeply unsatisfactory.

§4 DISJUNCTIVISM AND THE REVENGE OF BERKELEY’S PUZZLE

Recently, John Campbell (2002a, 2002b, and 2009) has argued for disjunctivism on grounds other than just the need to respect the “transparency thesis”. For instance, in his “Berkeley’s Puzzle”, Campbell contends that the only way that we can respect what he calls “the explanatory role of experience” while at the same time avoiding Berkeley’s Puzzle is to adopt a “relational” version of disjunctivism.\(^\text{148}\) The classic statement of

---

\(^{146}\) Tye (2009), p. 120.
\(^{147}\) Tye (2009), p. 83.
\(^{148}\) Campbell (2002b), p. 128.
Berkeley’s Puzzle is the following passages from section 23 of Berkeley’s *The Principles of Human Knowledge* (1710):

But say you, surely there is nothing easier than to imagine trees, for instance, in a park, or books existing in a closet, and nobody by to perceive them. I answer, you may so, there is no difficulty in it: but what is all this, I beseech you, more than framing in your mind certain ideas which you call books and trees, and at the same time omitting to frame the idea of anyone that may perceive them? But do you not yourself perceive or think of them all the while? This therefore is nothing to the purpose: it only shows you have the power of imagining or forming ideas in your mind; but it doth not shew that you can conceive it possible, the objects of your thought may exist without the mind: to make out this, it is necessary that you conceive them existing unconceived or unthought of, which is a manifest repugnancy.\textsuperscript{149}

As Campbell sees it, Berkeley’s basic challenge here is this: If empiricism is true, then our conception of external reality must be acquired on the basis of our experience of the world. But, it is difficult to see how perceptual experience can provide us with a conception of the world as mind-independent if we accept the traditional empiricist view that what we are aware of in sense-perception are the qualitative characters of our own sensory experiences. Indeed, Campbell says:

The puzzle that Berkeley is addressing is that it is hard to see how our concepts of mind-independent objects could have been made available by experience of them. The resolution he finds is to acknowledge that we do not have concepts of mind-independent objects.\textsuperscript{150}

Campbell himself (rightly) thinks that it is unacceptable for us to conclude that we lack knowledge of the mind-independent world. But if we are to avoid Berkeley’s radical conclusion, he argues, we must answer his challenge head-on. Hence, we must adopt a theory of perception which respects what he calls “the explanatory role of experience”.

According to Campbell, the explanatory role of experience is best thought of as:

\textsuperscript{149} Campbell (2002b), p. 128  
\textsuperscript{150} Campbell (2002b), p. 128.
The principle…that concepts of individual physical objects, and concepts of the observable characteristics of such objects, are made available by our experience of the world. It is experience of the world that explains our grasp of these concepts.\footnote{Campbell (2002b), p. 128.}

Campbell’s idea here, echoing Berkeley, is that experience, and perceptual experience in particular, must explain how it is that we can form conceptions of individuals and their features as mind-independent. Furthermore, he maintains, perceptual experience must explain how we come by conceptions of individuals or features which characterize them categorically rather than merely in terms of their spatio-temporal and causal-dispositional profiles. Put slightly differently, perceptual experience must explain how it is that the external world is revealed to us, and how we can thereby have a conception of its intrinsic, categorical nature.\footnote{Campbell (2002a), p. 116 and pp. 137–42 and Campbell (2002b), pp. 141-3.}

Unsurprisingly, Campbell argues that only his disjunctivist “Relational View” of perceptual experience can adequately respect the explanatory role of experience. As we’ve seen, Campbell’s own view basically follows Russell’s in holding that perceptual experience must be more primitive than any knowledge of truths about, or conceptions of, the individuals or features of which we are aware and explain how we acquire them. In particular, Campbell sees perceptual acquaintance as playing an essential role in grounding the reference of our basic demonstrative concepts and providing us with knowledge of their reference. As Campbell puts it:

On the relational view, experience of objects is a more primitive state than thought about objects, which none the less reaches all the way to the objects themselves. In particular, experience of an object is what explains your ability to grasp a demonstrative term referring to that object.\footnote{Campbell (2002a), pp. 122-3.}
In claiming here that acquaintance “reaches all the way to the objects themselves”, he is not only siding with Russell in holding that perceptual acquaintance must be conceived as an experiential relation between a conscious subject and the perceived objects or feature-instances where both the subject and the objects are constituents of the relation (or, better, the relational fact). He also means to indicate his agreement with Moore (and the early Russell) that perceptual acquaintance itself is a “transparent” epistemic relation such that the qualitative characters of which we are aware inhere entirely in the objects and feature-instances in the external perceptual scene.

For Campbell, it is crucial that perceptual acquaintance is both non-representational (and hence non-conceptual) and transparent, since only such a combination provides us with a picture of experience on which the intrinsic, categorical nature of the external world is directly and straightforwardly revealed to us, rather than being inferred from the manifestations of its causal-dispositional powers. For it is precisely because we have such non-representational and transparent epistemic contact with the intrinsic, categorical nature of the external world that we are in a position to form conceptions of objects and feature-instances as categorical and as mind-independent.

In contrast, Campbell argues, it is entirely unclear how we would be in a positions to form such conceptions of external reality as categorical and mind-independent on any “common factor” view of experience on which the immediate objects of our acquaintance are the qualitative characters of our own experiential states and our knowledge of external reality is entirely inferential in character. Indeed, he contends:

Experience of objects has to explain how it is that we can have the conception of objects as mind-independent. The objection to the common factor view is that,
on it, experience of objects could not be what explained our having the conception of objects as mind-independent.\textsuperscript{154}

The reason why no such common factor view can provide us with conceptions of objects and feature-instances as categorical and mind-independent is, he argues, simply that:

On the common factor view, all that experience of the object provides you with is a conscious image of the object. The existence of that conscious image is in principle independent of the existence of the external object. The existence of the image, though, is dependent on the existence of the subject who has the conscious image. So if your conception of the object was provided by your experience of the object, you would presumably end by concluding that the object would not have existed had you not existed, and that the object exists only when you are experiencing it. We cannot extract the conception of a mind-independent world from a mind-dependent image; this is the traditional problem with Locke’s doctrine of abstraction.\textsuperscript{155}

But on the other hand, he contends:

It seems as though it ought to be possible, though, to extract the conception of a mind-independent world from an experience which has a mind-independent object as a constituent, which is what the disjunctive view ascribes to us.\textsuperscript{156}

For Campbell, this provides us with extremely compelling grounds for preferring his disjunctivist Relational View of experience to any on which there is a qualitative common factor between the veridical and non-veridical cases of perceptual experience.

Of course, he grants, these considerations don’t completely rule-out a common factor view of perceptual experience. After all, it could be that our mind-dependent perceptual experiences are constituted such that they \textit{present} themselves to us \textit{as if} they were revealing features of the world. If so, then we might suppose that we have the tools for a common factor view to answer Berkeley’s puzzle. In particular, we might be able to explain how we at least come by a conception of reality as mind-independent and

\textsuperscript{154} Campbell (2002b), p. 134.
\textsuperscript{156} Campbell (2002b), p. 135
categorical, even if we would lack a revelatory conception of it and hence could never be absolutely sure about its intrinsic, categorical nature or existence.

However, Campbell does not believe any such strategy will be satisfactory. The central problem with such a strategy, he argues, is that it takes for granted the intentionality, or object-directedness, of experience, an assumption which he thinks is illegitimate. He remarks that:

You might have thought that the immediate response of a common factor theorist to this argument is that the image provides the conception of an objective world simply by displaying the world as objective. Even if I am hallucinating, the objects I seem to see seem to be mind-independent objects. So a common factor image can present objects as mind-independent; and surely that is all that is needed. The problem with this reply is that it takes for granted the intentionality of experience.\textsuperscript{157}

His claim here will strike many as somewhat surprising, since perception is generally seen as a paradigmatic instance of an intentional phenomenon. But it is clear that he has something of an idiosyncratic view of intentionality. Basically, he imagines that the common factor view sees the intentionality of perceptual experience as a matter of grasping a proposition in thought. In particular, see thinks that the common factor view is committed to seeing perceptual experience as a matter of grasping a “demonstrative proposition” in thought. As he puts it:

It takes it for granted that experience of the world is a way of grasping thoughts about the world. To see an object is, on this conception, to grasp a demonstrative proposition. There are many ways in which you can grasp a proposition: you can grasp it as the content of speech or as the meaning of a wink or a sigh. One way in which you can grasp a proposition is as the content of vision. The common factor theorist says that ordinary vision involves grasping demonstrative

\textsuperscript{157} Campbell (2002b), p. 135. I’ll ignore Campbell’s somewhat strange view that propositions can be demonstrative.
propositions as the contents of experiences. And you could grasp such propositions whether or not the external objects exist.\textsuperscript{158}

The problem with this approach, he argues, is that it makes perception out to be just one among many ways of grasping a proposition and hence violates the explanatory role of experience. That is, it does not explain why experience plays a fundamental role in our acquisition of a conception of reality as categorical and mind-independent. And, he argues:

It is when we press the explanatory role of experience like this that we can see the force of the disjunctivist’s argument. We are not to take the intentional character of experience as a given; rather, experience of objects has to be what explains our ability to think about objects. This means that we cannot view experience of objects as a way of grasping thoughts about objects. Experience of objects has to be something more primitive than the ability to think about objects, in terms of which the ability to think about objects can be explained.\textsuperscript{159}

Thus, Campbell once again concludes that his disjunctivist Relational View (and presumably its close cousins, such as Tye’s view) is the only plausible theory of experience which can answer Berkeley’s Puzzle.

I believe that Campbell is mistaken on this point. While he (seemingly) is correct that a proper account of perceptual experience must be relational in character, he is wrong to think that we must be disjunctivists in order to be relationalists. Like many currently working on a Russell-inspired notion of perceptual acquaintance, he apparently sees naïve realism (and disjunctivism) and indirect realism as an exhaustive disjunction of possible views about the nature and character of perceptual experience. I believe that a proper account of perceptual experience is one in which its content is specified along the

\textsuperscript{158} Campbell (2002b), p. 135.
\textsuperscript{159} Campbell (2002b), p. 136.
lines of direct realism (and relationalism), but its *qualitative character* is a mind-dependent feature of our own experiential states.

Furthermore, I think that Campbell mischaracterizes the intentional character of such a “common factor” view in holding it to be a matter of grasping a demonstrative proposition in thought. Later in this chapter, I will advance what I take to be a far more plausible account of the intentional (and ultimately *informational*) character of perceptual experience. In doing so, I will provide valuable resources for a common factor view on which a subject might naively suppose that perceptual experience transparently reveals the intrinsic, categorical nature of the world, even though it does not. But at the same time, it will explain how we are directly presented with (or “intentionally” aware of) the intrinsic, categorical nature of the world (in a non-revelatory way), such that we can still form conceptions of it *as mind-independent* and *as having some categorical nature.* Hence, it provides a viable alternative to Campbell’s own disjunctivist Relational View.

In fact, I believe that my account ultimately fares *better* than Campbell’s (or Tye’s) in respecting the explanatory role of experience and avoiding Berkeley’s Puzzle. In my view, an adequate account of perceptual experience must not only provide an epistemologically and metaphysically acceptable story of how, when all goes well, it is able to provide us with a conception *of* the intrinsic, categorical nature of the external world *as* mind-independent and *as* having some categorical nature, it must also provide an acceptable story about the role of experience in providing us with mistaken conceptions of the world when all does not go well. However, neither Campbell nor Tye provide an epistemically and metaphysically acceptable account of the nature of illusory and hallucinatory perceptual experience. More specifically, their views are unable to
overcome a serious problem for all versions of disjunctivism which I call “the problem of phenomenal indiscriminability”. And because they seemingly cannot plausibly answer the problem of phenomenal indiscriminability, they are unable to provide a satisfactory response to Berkeley’s Puzzle that respects the full explanatory role of experience.

To see what the problem of phenomenal indiscriminability is, it is helpful to remind ourselves of why Campbell and Tye turn to disjunctivism in the first place. First of all, both Campbell and Tye aim to respect Moore’s thesis of the “transparency of experience” while at the same time respecting Russell’s notion of experience as a kind of epistemic contact with things that is independent of, and cognitively more fundamental than, any propositional knowledge of truths about things. For this reason, both Campbell and Tye embrace some version of naïve realism regarding cases of veridical perceptual experience. (Campbell is also led to this view out of considerations about the explanatory role of experience and his concern to answer Berkeley’s Puzzle). However, they are both equally aware that naïve realist accounts of perceptual experience have significant problems in explaining cases of illusory and hallucinatory experience, and are seemingly in tension with the widespread view of current scientific theories of perception that the right sort of neurobiological states of the brain are (at least) causally sufficient of perceptual experiences as of external objects and feature-instances.

But rather than following many philosophers and scientists of perception in concluding that the qualitative phenomenal features of which we are aware in perceptual experience are mind-dependent features of the experiential states themselves rather than mind-independent features of external objects, they simply reject the view that we have the same kind of experience in the problematic cases as in the veridical case. They rightly
point out that we cannot conclude from the fact that veridical and non-veridical perception-like experiences are indiscernible from the subject’s point of view that they are in fact instances of one and the same kind of experience. At best, they argue, what we have is prima facie evidence of identity, evidence that can be overturned by overriding theoretical considerations of the sort that they take themselves to have. So, they contend, rather than seeing veridical perceptual experience and non-veridical perception-like experience as forming a unified kind, we should see them as jointly constituting a disjunctive kind of the sort we’d get if we (mistakenly) lumped zebras and cleverly-painted mules together into a single kind.\footnote{Dretske (1970) uses this example for very different purposes.}

Of course, Campbell and Tye disagree about the particular details of what is going on in the non-veridical cases. Campbell argues that in illusory cases, we can account for why the perceptual scene is subjectively indiscernible from an entirely different veridically-perceived perceptual scene by appealing to differences in the points-of-view from which the subject experiences them. But he says very little about hallucinatory experiences, or ones abnormally caused by the direct manipulation of a subject’s neurobiological states of the brain, except of course that they are a different kind of experience. On the other hand, Tye accounts for the subject’s inability to tell the veridical case apart from the problematic cases by claiming that both kinds of cases share the same “singular-when-filled” content schemas, and that the phenomenal character of experience supervenes on such content schemas (though it is hard to see how this squares with his qualia externalism). But on either view, the key move in evading an internalist, “common factor” view about the qualitative character of experience is to deny that the \textit{subjective}
indiscriminability between the veridical cases and the problematic cases entails their sameness in kind.

But while Campbell and Tye are clearly right on this point, disjunctivism is not yet out of the woods. The problem they now face is in providing an epistemically and metaphysically satisfactory account of why subjects cannot tell the problematic cases apart from the veridical cases in a way that respects the full explanatory role of experience. This is the true heart of the problem of phenomenal indiscriminability, a problem for which neither Campbell nor Tye has a satisfactory answer.

We might initially pose the problem in the form of a question: What makes it the case that a subject cannot tell apart the veridical case from the problematic cases if they are not instances of one and the same kind of experience? Clearly, the most natural initial answer for why this might be so is that the veridical and non-veridical perception-like experiences have extremely similar qualitative characters. After all, we are quite familiar with the fact that we often make mistakes in concluding sameness of character when considering instances of qualities of very similar kinds.

But there is a problem for both Campbell and Tye in explaining how phenomenal indiscriminability could arise from a close similarity between veridical and non-veridical cases of perceptual experience, particular when hallucination (or abnormal direct neurobiological manipulation) is involved. This is because similarity is a supervenient relation, one that plausibly must be grounded by both the veridical and problematic cases having some qualitative character or other. But the problem now is: Where are we to locate the qualitative character of the hallucinatory experience which is needed to ground its extreme similarity to that of the veridical case? Obviously, it cannot be in the external
perceptual scene for, by hypothesis, in the problematic cases at issue there are no such objects or feature-instances in the external scene (or if there are, they are not playing their normal role in bringing about the so-called “veridical hallucination”). But then where in the world are they?

Campbell’s complete silence on the matter suggests his inability to come up with a plausible answer to this challenge. But Tye’s response fares little better. His strategy is to argue that phenomenal character supervenes on singular-when-filled content schemas rather than straightforwardly on the arrangement of objects and feature-instances in the external perceptual scene. This helps him to account for why a veridical perception of a floating dagger is subjectively indiscernible in character from a hallucination as of a floating dagger. In both cases, we have the same singular-when-filled content schemas; the difference simply is that in the veridical case the slots for the objects and feature-instances in the schema are filled by objects and feature-instances in the external perceptual scene, but in the hallucinatory case they are unfilled. But what his theory cannot account for are phenomenal differences between two very different kinds of hallucinatory experiences with the same number of slots for objects and feature-instances. Nor can he plausibly account for phenomenal differences between, say, a veridical perceptual experience of a zebra and an illusory perceptual experience in which a zebra appears to be an object other than it is, one with somewhat different features (fill in whatever details one likes).

Some disjunctivists, following Johnston (2004), argue that in the problematic cases, what a subject is aware of in experience is a un-instantiated “sensible profile”. I must confess that I cannot make sense of the idea that the qualitative character of a

---

hallucinatory perceptual experience is constituted by, and thus explained by, a structured cluster of feature-instances (and relations) that are not instantiated by any object and, in fact, are not instantiated at all. How could such un-instantiated structured clusters of features (and relations) possibly make a difference to the character of one’s experience? And how could their non-occurrence have any impact whatsoever on the behavior and cognitive life of the subject? Out of all of the un-instantiated features (and relations) there “are”, why do just these make such a difference to one’s experience? Frankly, I can make no sense of Johnston’s “obscure” idea, and I think it is best to be avoided.

At any rate, while I certainly haven’t explored every possible way in which a disjunctivist might try to answer the problem of phenomenal indiscriminability, I hope I’ve at least made it clear how difficult it is for them to answer in a satisfactory manner. The disjunctivist cannot claim that the qualitative characters of non-veridical perception-like experiences are mind-dependent on pain of implausibly multiplying their ontology or, even worse for them, giving up naïve realism in the veridical case as well. But the disjunctivist cannot simply put it in the external scene either, unless they are willing to deny that there is any such thing as hallucinatory experience. And lastly, the disjunctivist cannot simply maintain that the qualitative character is instantiated neither in the mind of the subject nor in the external scene on pain of obscurity and (in my view) incoherence. Clearly, none of the disjunctivist’s options are particularly attractive.

What’s more, even if we set aside such worries about what is going on in the non-veridical cases, there still remains a serious problem for the disjunctivist like Campbell who wants to answer Berkeley’s Puzzle. Simply put, the problem is that the relational character of veridical experience is what is supposed to explain, all by itself, how we are
able to form conceptions of the external world as mind-independent and as categorical. But we cannot tell from our own point of view whether our experience is of the veridical kind, or of a sufficiently qualitatively-similar hallucinatory kind in which we are not experientially related to anything. So if we cannot tell which circumstances we are in, how could one provide us with our conception of the external world as mind-independent and as categorical when the other does not? It is not clear that the disjunctivist can supply an adequate answer.

All things considered, I think it is very much worthwhile to look elsewhere for an adequate account of perceptual experience. And so in the rest of this chapter, I will advance my own account of perceptual experience, which I believe incorporates the respective strengths of naïve realism and common factor views while at the same time avoiding each of their respective shortcomings.

§5 INFORMATION AND INTENTIONALITY

In this section, I want to provide a framework that I believe is crucial for thinking about the nature of perceptual acquaintance. And while I ultimately aim to defend an internalist account of the qualitative characters of our perceptual experiences, I will here provide an account of the nature of perception that is, in its broadest features, compatible with both naïve realism and its competitors. Indeed, I think that any adequate account of perception must begin with a more primitive theory about information, informational content, and intentionality. In developing my theory of perception, I will rely on the theory of information and informational content advanced by John Perry in his “Information, Action, and Persons.”

---

162 Perry (2002). Perry’s earlier work with David Israel and Syun Tutiya is also excellent for thinking about the nature of information, but I will focus only on his later work on the topic.
To begin with, the basic unit of information on Perry’s account is a signal. A signal is an object having a property, where both the object and the property may be quite complex. Perry calls this object the carrier of information and its relevant property the indicating property.\textsuperscript{163} Given this technical notion of a signal, Perry characterizes information as:

What one part or aspect of the universe (the signal) shows about some other part or aspect [of the universe] (the subject matter).\textsuperscript{164}

However, as Perry rightly notes, the mere occurrence of some state or event in the world says nothing about how things are with the rest of reality unless its occurrence is against the background of the way the world works, or at least some conception thereof. Indeed, it is a common fault of informational-semantic theories that they fail to emphasize that signals only carry information about other parts or features of the universe given a background of constraints, where constraints are understood as states-of-affairs, some of which are true. Perry nicely puts this point as follows:

[That a signal carries information about some other part or feature of the universe] is possible only because events are constrained by laws of nature, or as I prefer, because of its more liberal, common-sense, loose, and non-reductive connotations, by the way that things happen. The information carried by a signal is what else things have to be like, for the signal to have occurred, given the way things happen.\textsuperscript{165}

Perry’s idea is that a signal $S$ carries the information that $P$ if there are principles of how things actually happen such that given those principles, the signal would not have occurred unless $P$ were the case.

One crucial thing to notice about his notion of information here is that it is factive. In other words, given the way the actual world works, it is not possible for a signal to carry the information that $P$ and for it not to be the case that $P$. This is one reason that

\textsuperscript{165} Perry (2002), p. 175.
information is not a viable candidate to which we can reduce intentionality, for
intentional states of systems can be directed at, can have as conditions of satisfaction,
states-of-affairs that are not the case. Another reason that we cannot reduce intentionality
to Perry’s notion of information is that any occurrence in the world can constitute a signal
that carries information about any other state of the world relative to the right constraints
about how the world actually works. But one of the essential features of genuine original
intentionality is that it is directed at, or places conditions on, particular objects or states-
of-affairs that, again, might or might not in fact obtain. Moreover, it is generally directed
in some particular way at them. Thus, it is not enough for the signal simply to causally-
covary with some particular indicated object or state-of-affairs.

Perry himself, of course, is quite aware of the fact that we cannot plausibly
identify intentionality with, or reduce it to, mere information. Instead, he maintains that
naturalized intentionality is to be identified with the right sort of informational content.
To see how informational content differs from mere information, it is helpful to have the
following rough general schema:

A signal $S$ has the informational content that $P$ if and only if, given constraints $C$,
$S$ occurs only if, and because, $P$.\(^{166}\)

In the case of mere information, the constraints against which a signal acquires its
informational content are true constraints governing the way the actual world works.
However, what is crucial for Perry’s notion of informational content is that the
constraints we rely on in determining the content of a particular signal need not be true, a

\(^{166}\) The “because” here should be interpreted quite liberally to capture logical, metaphysical, nomological,
conventional, etc., connections between the signal and what it indicates.
fact which allows us to consider counterfactual circumstances and which, as we’ll see, explains many cases of informational error.

Furthermore, Perry maintains that for any signal there are numerous layers of informational content they have depending on which constraints and circumstances are assumed in the background. Some of this information content is what Perry calls reflexive information, or information about the signal itself.167 Other levels of the signal’s informational content are about states or features of the rest of the world; Perry calls information of this sort incremental information.168 Perry thinks we get from the reflexive information about the signal itself to the subject-matter, incremental information it has by adding to the background constraints particular details about the circumstances of the signal’s occurrence.169 In other words, a signal will give us information about what some other part or aspect of the world must be like given the signal’s occurrence, the circumstances in which it occurred, and the way the world works. Borrowing one example from Perry, an x-ray has the informational content that so-and-so has a cavity in such-and-such tooth given the way that x-rays and tooth decay work and given the circumstances that the x-ray was exposed to so-and-so’s tooth.170

Perry thinks that many systems and devices depend on information and informational content having these features. Indeed, he thinks that many systems and devices are constructed in such a way that they harness the informational content of a signal in order to satisfy some goal. They do this by having some sort of architectural design such that the occurrence of a particular signal is also the cause an action that will

be successful in just those circumstances. To borrow one of Perry’s examples, the springing of a mousetrap carries the information that there is a mouse in the trap relative to the constraints that only mice will spring it, and it is designed to succeed in trapping mice in precisely those conditions in which a mouse is in it.\textsuperscript{171} If, however, the constraints are not satisfied, then the occurrence of the signal will fail in bringing about the designed effect.

Of course, in the case of the mousetrap, its informational content is ‘derived’ rather than ‘original’. That is, the occurrence of the trap’s signal has the informational content that there is a mouse in the trap given the relevant constraints and circumstances only because we have given it the job or goal of trapping mice. For this reason, Perry acknowledges that it is not enough for a signal simply to have informational content for it to have genuine intentionality. But what more is needed on Perry’s account to get to genuine original intentionality? His answer is that what we need is a system whose indicating states have natural jobs or goals to indicate some state-of-affairs in the world and to guide actions that make sense given their occurrence. However, unlike so-called ‘teleosemantic’ approaches to naturalized intentional content, Perry does not limit the sources of natural functions of states to those supplied by natural selection, learning, and social accretion.

In fact, one striking feature of Perry’s account is that the qualitative characters of many of our states of sentient awareness are a vital source of naturally supplied goals. In particular, he maintains that the pleasant or unpleasant dimensions of the qualitative characters of many of our experiential states provide us with a crucial source for deriving natural goals, a fact that natural selection and social accretion make good use of in

generating more sophisticated natural goals for us.\textsuperscript{172} Hence, Perry’s own view bears a strong resemblance to those (like Searle’s) according to which we cannot explain intentionality naturalistically, but only by appealing to consciousness itself. However, where he parts company with such opponents of naturalized intentionality is in holding that consciousness itself, and its qualitative characters, can in principle be naturalized. Of course, it is absolutely crucial to point out that Perry holds that the physical domain must be such that consciousness, in all of its subjective and qualitative richness, can arise within it. So he is not the kind of traditional naturalist who would seek to make consciousness out to be less than it in fact is simply with the goal of making it fit easier into an overly impoverished conception of physical reality.

\textbf{§6 INFORMATIONAL CONTENT AND PERCEPTUAL CONTENT}

In the previous section, I presented Perry’s account of informational content as what one aspect of the world tells us about another aspect of the world given relevant constraints and circumstances. But what exactly does this have to do with perception? I think that what is central to perception is not its causal character, as many believe, but rather its informational character. It is that the states of our perceptual systems have informational content about states of the world outside of our perceptual system in a way that we can harness this information to guide our thought and action. In fact, our perceptual systems have been designed through processes of natural selection, social accretion, and learning to provide us with information about our environment in order to help us be more successful in navigating it and thereby satisfying our various practical goals.

\textsuperscript{172} Of course, this can’t be the full story since many experiences, such as our experiences of color, do not have a recognizable dimension of pleasantness or unpleasantness.
Perception accomplishes this, of course, because we and our perceptual systems are, as Perry puts it, ‘attuned’ to more-or-less accurate regularities that hold between distal objects in our environment and the internal states of our perceptual systems. The crucial notion of attunement here is of an unreflective sensitivity to, or even differential responsiveness to, some more-or-less accurate constraints on the way that the world works. This sensitivity to constraints allows us to track and harness the information carried by states of our perceptual system about particular events in our environment. We might even construe attunement as a set of capacities, abilities, or know-how of the subject, or of his or her perceptual faculties.\(^{173}\) For example, many birds are attuned to the constraint that the path to any clearly visible object is unobstructed; indeed, their attunement to this constraint guides their (often highly complex) behavior even though they lack an explicit and reflective appreciation of the constraint and even though the constraint is false (given the advent of transparent windows).\(^{174}\)

One benefit of thinking of perception in terms of informational content is that it helps explain how direct realism might be true. For, as Searle remarks in a similar context, it is important to note that an account of perception does not get to declare itself a version of direct realism for free. Rather, direct realism should be a consequence of an independently motivated account of the intentionality of perception. And indeed, one crucial benefit of thinking of perception in terms of informational content is that we evidently can explain how direct realism about perception can be true. This is because the informational content of a signal is literally about the very external objects, feature-

---

\(^{173}\) I think that Perry’s notion of attunement to constraints bears important similarities to Searle’s notion of the “background” against which our intentional states have their particular conditions of satisfaction. See Searle (1983).

\(^{174}\) I borrow this example from Perry (2002), p. 184.
instances, or state-of-affairs, if any, for which, in the particular circumstances of its occurrence and relative to appropriate constraints, it has the job of indicating. And in the case of perceptual experience, the constraints governing our perceptual capacities are such that our states of perceptual awareness are directly about perceivable external objects, feature-instances, or state-of-affairs in our environment.

What’s more, this fact about the nature of informational content explains how it is that perception can have several other important intentional features noted by Searle and others.\(^\text{175}\) First, it explains why perception has an essentially indexical element; indeed, our perceptions are always about particular objects, feature-instances, or states-of-affairs that bear some indexical relation to us and our perceptual faculties. Secondly, it explains why perception is, as Searle puts it, ‘causally self-referential’. The basic idea here is that in cases of veridical perception, we are presented with the very objects, feature-instances, or states-of-affairs that cause our perceptual experience as of them. And indeed, it is very natural to see the causal self-referential character of all perception as a reflexive informational content that is a constitutive feature of perceptual signals. But we can also maintain the intuitive picture that the subject-matter or incremental informational content of perceptual experiences is exhausted by the objects, feature-instances, or states-of-affairs, if any, that constitute their conditions of satisfaction. In other words, we can straightforwardly maintain that a perceptual experience is veridical if and only if the particular things experienced are the way the perceptual experience represents them as being.

Most importantly, thinking of perception in terms of its informational content also helps explain how it is that we can have non-veridical perceptual experiences or

\(^{175}\) Searle (In Draft), pp. 19-25.
hallucinations, ones that have a qualitative ‘common factor’ with (or subjectively indistinguishable difference from) veridical cases. This is because a perceptual signal has its particular informational content only relative to constraints and circumstances to which we are attuned, which suggests a very natural explanation of why, in cases of non-veridical or perceptual experience, the states of our perceptual systems do not carry the information we take them to. Indeed, the basic idea here is that in such non-veridical or hallucinatory cases, we (or our perceptual systems) are attuned to false constraints on how the world works and we find ourselves in circumstances that bring out this fact. In cases of perceptual error and illusion, the conditions under which we perceive some object, feature-instance, or state-of-affairs are such that we have a perceptual experience whose qualitative character and informational content do not accord with the actual information that the occurrence of the perceptual state carries. Hence, we take the perceived object, feature-instance, or state-of-affairs to be other than it in fact is.

Moreover, in cases of hallucination, we find ourselves in such abnormal circumstances that we have the relevant perceptual experience without there being any sort of corresponding object, feature-instance, or state-of-affairs required for the satisfaction of its veridicality conditions.\(^{176}\)

In fact, once we recognize that the singular content of perceptual experience is incremental informational content, we have a natural explanation for why our perceptual experiences are error-prone in the first place. For like many information-harnessing devices, we needn’t be attuned to particularly accurate constraints and circumstances for

\(^{176}\) Obviously, this story is extremely over-simplified. For instance, we are also attuned to the fact that our perceptual faculties are fallible, which explains why we do not take visual blackouts, double-vision, blurry vision, etc., to be qualitative features of the world. We generally take such perceptual experiences to provide information about the improper functioning of our own perceptual faculties rather than information or misinformation about the world.
our perceptual faculties to perform well enough at helping us pick up information about the environment suitable for guiding successful actions. All that is required is that our perceptual faculties work well enough in the circumstances in which we most often find ourselves, and there is sufficient reason to suppose that they do.

In any event, my account of the informational content of perception provides an intuitive explanation about why there at least seems to be a common factor between the veridical and non-veridical cases of perceptual experience. The reason is that many of the possible errors can occur even if we have exactly the same signal. For, as I’ve just noted, the veridicality of a particular informational signal depends on factors external to it, such as background constraints and circumstances. Hence, if the occurrence of the signal itself suffices for the instantiation of the qualitative character of our perceptual experience, as I believe is the case, then we have a straightforward explanation for why there is a qualitative ‘common factor’. The common factor in such cases just is the occurrence of one and the same informational content-bearing signal.

§7 TWO KINDS OF AWARENESS

Now that we have at least a sketch of a framework of the informational character of perception, I want to consider how we might be internalists about the qualitative character of perceptual experience without being indirect realists about perception. My central claim in this section will be that given a proper informational framework for thinking about perception, we can accomplish this, provided that we are careful to distinguish between intentional and phenomenal awareness. I will also return to the question of how I see the resulting theory of perception as fitting into a broader Russell-
inspired account of acquaintance. What I hope to show, at least in outline, is that we can indeed have an adequate model of perceptual acquaintance without adopting naïve realism or disjunctivism.

In fact, I think we took the first steps towards developing an adequate model of perceptual acquaintance in the previous section when we made the distinction between signals and their informational contents. Indeed, what we saw was that we can usefully think of perceptual experiences as signals whose incremental informational content is directly about the external objects, if any, for which they have the job of standing, relative to appropriate constraints and circumstances. Thus, we can respect the naïve and direct realist insight that our perceptions are directly of, or about, external objects, feature-instances, and states-of-affairs. On the other hand, we can still suppose that the qualitative characters of our perceptual experiences are identical with, or supervene on, intrinsic features of the signal itself. In fact, I think that one of the crucial insights of Perry’s theory of information and informational content is that it explains why information is useful in the first place. For recall that, fundamentally, it is what one state or aspect of the world says about another state or aspect of the world, relative to constraints and circumstances, in virtue of the properties it (i.e. the signal) has. Indeed, the structure of information and informational content is important because it allows us to get at the properties of an accessible signal in order to acquire information about distal state-of-affairs to which we do not have as ready access.

Of course, we have to be extremely careful here. Unlike in many cases in which information is useful, we don’t literally look at, or perceive, the signals in the case of our own perceptual experiences. Indeed, supposing that we do so is exactly the mistake that
the indirect realist about perception makes. Instead, we simply have the perceptual experiences, and we are, in a relevant sense, aware of (or at least sensitive to) the qualitative features they possess just by having them.\textsuperscript{177} Moreover, it is in virtue of having, and thereby being aware of, such perceptual experiences that, in the right conditions, we directly and literally perceive external objects, feature-instances, or state-of-affairs, without any intermediate inference.

But how can we be aware of the qualitative features of our experiential states just by having them? Unfortunately, I cannot hope to answer such a deep question here. What I will say is that whatever metaphysical account of consciousness we offer, it must ultimately account for this somewhat remarkable fact. Indeed, I take it as a datum to be explained that we have experiential states where the having is, in the relevant non-reflective and non-propositional sense, the knowing.\textsuperscript{178}

This brings us to John Searle’s (In Draft) crucial distinction between the two senses of ‘of’ that he thinks are relevant to our perceptual awareness. According to Searle, there are two fundamentally different kinds of awareness relation, both of which are marked in English by the preposition ‘of’. The first is what he calls ‘the ‘of’ of constitution’, and the second is what he calls ‘the ‘of’ of intentionality’.\textsuperscript{179} To see what the distinction between these two relations are, Searle asks us to consider the following two cases:

(1) When I feel my pain, I am aware of, or conscious of, the pain.

\textsuperscript{177} I want to note that the locution “aware of” is somewhat problematic for characterizing the phenomenal awareness we enjoy simply in virtue of having a perceptual experience, since it suggests “intentional awareness” rather than “phenomenal awareness”. Unfortunately, I lack a better locution.

\textsuperscript{178} Galen Strawson makes roughly the same point in Strawson (2009), p. 286.

\textsuperscript{179} Searle (In Draft), p. 14. It is important to note that we do not have follow Searle’s somewhat dubious semantic thesis about the English preposition ‘of’ to agree with him that there is a distinction between the two sorts of awareness involved in perceptual experience.
When I see something red, I am aware of, or conscious of, the instance of red. What Searle hopes we’ll recognize is that the awareness relations involved in (1) and (2) are quite different. Intuitively, in the first case the awareness of the pain just is the pain, but in the second case, the awareness of the red is not itself the instance of red. Rather, in the case of (2), the instance of red is the intentional object of our (in this case) veridical awareness as of something red; it is what our perceptual experience as of red is directed at, or about. Searle quickly characterizes his distinction as follows:

In the case of pain, the pain is identical with the awareness or the consciousness. There aren’t two things, the pain and the awareness of the pain. That is what I call the “of” of constitution. In this case the awareness is constituted by the thing that I am aware of. But that “of” is different from the “of” of intentionality. When I see a red object, I am indeed aware of red, but the awareness in this case is of the red in the red object. The red is not identical with the awareness, the “of” is the “of” of intentionality.

Searle rightly thinks that it is absolutely crucial that we do not confuse the ‘of’ of constitution with the intentional relation we bear to objects, feature-instances, or states-of-affairs in perception. If one does so, he argues, then it is far too easy to form the mistaken belief behind indirect realism that:

Intentionality consists invariably of some sort of representation, and the subject who has the intentional state has some sort of [conscious] relation to these representations [rather than to their intentional objects].

What’s more, he continues:

---

180 I think that there are perhaps two different ways of interpreting Searle’s important notion of ‘the ‘of’ of constitution’. First, we can follow him in holding that the awareness of pain, for example, just is the pain. But second, we might instead hold that the awareness of pain is numerically identical with our having the pain. In my opinion, it is quite natural to say that when we have a perceptual experience, and are thereby aware of it, this just is the state of sentience, and so I prefer the second interpretation of Searle’s notion. However, I think either way of construing the ‘of’ of constitution is compatible with our English usage of ‘aware of’ and little turns on the issue.

181 I take it that Searle holds that in this second case, what we are aware of, in the sense of the ‘of’ of constitution, is the qualitative character of the experience as of red.


That is what forces the analogy between the intentional theory of perception and [the] idea that intentionality would be like reading a newspaper about the real world. I think, frankly, this is quite an absurd conception of intentionality of perception… [And] if you think that all intentionality is a matter of [conscious] relation to a representation, that the object of the intentionality is the representation or some element of it, and that on an intentionalistic account the awareness in the awareness of a hallucination must be the same kind of awareness as the awareness of an object in a veridical perception, then it will seem to you that an intentionalistic account of perception involves a denial of naïve realism.\(^\text{184}\)

In fact, Searle thinks that it is roughly this line of reasoning that forces one into the false dilemma of choosing naïve realism or adopting some version of indirect realism. However, Searle thinks that we shouldn’t be moved by this argument, and that we won’t be, provided we distinguish between the ‘of’ of constitution and the ‘of’ of intentionality.

One lesson we should take from Searle’s distinction is that there is a fundamental difference between what we might call ‘phenomenal awareness’ and ‘intentional awareness’. In the case of phenomenal awareness, we are non-reflectively and non-propositionally aware of the qualitative characters of our experiential states simply in virtue of having them. In the case of intentional awareness, on the other hand, we are aware of some state of the world in virtue of being in some representational state that is directed at it. Like our phenomenal awareness, such intentional awareness of things needn’t be reflective or propositional. Indeed, the subject need only be attuned to the informational content (and to relevant background constraints and circumstances) indicated by the representational state. However, unlike in the case of phenomenal awareness, intentional awareness can be reflective and propositional for sophisticated concept-wielding subjects. In fact, sophisticated enough concept-wielding subjects can

\(^{184}\) Searle (In Draft), p. 15. Note: Searle himself uses the term ‘naïve realism’ in the same fashion as I use ‘direct realism’.
even direct their reflective intentional awareness to think about their own states of
phenomenal awareness using higher-order so-called ‘phenomenal concepts’ of them. The
crucial point, however, is that this intentional awareness of our conscious experiences is a
further cognitive achievement from our more basic phenomenal awareness of them.

Returning to the central theme of this chapter, what I think these considerations
show is that there are various kinds of acquaintance involved in our conscious perception
of things. On one hand, in having a conscious perception, we are acquainted with the
perceptual experience itself. Our states of sentient awareness are in this sense ‘self-
presenting’; we are phenomenally aware of them when we have them in precisely the
same way that non-linguistic sentient animals are aware of their own conscious
experiences when they have them. And on the other, in having a conscious perception,
we are acquainted with the external intentional objects, if any, of the perceptual
experience. We are thereby presented with the objects, feature-instances, or states-of-
affairs, if any, of which it is a perception. This fact follows from the very informational
character of perception. Putting these two aspects together, we can say that conscious
perception has what we might call a “two-faced presentational character”; we are
presented in conscious perception both with the object of our perceptual experience, if
any, and with the subjective, qualitative character of the experience itself.

In saying this, we must be careful to note that for Russell, acquaintance does not
carry with it any commitment to dubious epistemic theses such as the so-called ‘thesis of
revelation’ or overly strong views about the ‘intimating’ character of the objects of our
acquaintance.¹⁸⁵ Though Russell holds that acquaintance presents us with its objects, he
does not claim that we thereby have an exhaustive presentation of the essential nature of

¹⁸⁵ Though as we saw last chapter, there are some unfortunate quotes that suggest otherwise.
that object, as the thesis of revelation would maintain. Nor do we automatically acquire any knowledge of truths about a thing simply by being presented with it in experience; I think we should take Russell’s contention that knowledge of things is logically independent of any knowledge of truths absolutely seriously.

This goes for the alleged ‘self-intimating’ character of our states of sentient awareness. We simply do not know that we are having such-and-such experiences simply in virtue of having them, at least if what we have in mind is reflective, semantically-articulated knowledge-that. Again, this is not to say that we sophisticated, concept-wielding adults don’t often think about them using phenomenal concepts, because we certainly do. The point is that this is a cognitive achievement above and beyond what acquaintance with our experiences provides all by itself.186

§ 8 CONCLUSION

In this chapter, I have argued that once we recognize what I’ve called the “two-faced presentational character” of perceptual experience, we can be direct realists about the content of perceptual experience and internalists about its qualitative character. We can be direct realists because the informational character of perceptual acquaintance puts us in direct experiential contact with external objects, features, and states-of-affairs. And we can be qualia internalists because of the special self-presenting character of the perceptual signals that have such informational content about the world. What I hope I have shown, at least in rough outline, is that we can have an adequate account of Russellian perceptual...

---

186 As we will see, one consequence of this view is that our phenomenal judgments are prone to error. Contrary to the opinions of many, I think this is a welcome consequence.
acquaintance without embracing problematic theories of perception such as naïve realism, disjunctivism, or indirect realism.
CHAPTER 3:  
PHYSICALISM AND THE PHENOMENAL CONCEPTS  
STRATEGY  

“Many philosophers have suggested in different ways that reflection on phenomenal concepts will play a major role in philosophy of mind, and in particular in the defense of physicalism, the doctrine that psychological truths supervene on physical truths. According to this strategy for the defense of physicalism—I will call it the phenomenal concept strategy—phenomenal concepts are importantly different from other concepts, and arguments against physicalism, in particular the conceivability argument and the knowledge argument, fatally neglect to take this difference into account.” – Daniel Stoljar, “Physicalism and Phenomenal Concepts” 187

§1 INTRODUCTION

Physicalists and dualists think that we have special knowledge of the experiential, that is, of experiences and the phenomenal properties that characterize what it is like to have them, by being acquainted with them. But exactly what kind of knowledge of the experiential does our acquaintance with it provide us with? Here there is little consensus. On one hand, dualists argue that acquaintance provides us with revelatory knowledge of the essence of the experiential such that we can a priori conclude that it is not physical. On the other hand, physicalists contend that our acquaintance with the experiential does not give us grounds to conclude that it is non-physical. Indeed, while they might grant that we have especially substantive knowledge of the nature of the experiential, knowledge of the qualitative character of experiences that is more intimate than other kinds of knowledge about the world, they deny that such knowledge reveals the

fundamental ontological nature of the experiential. The experiential is, according to them, wholly physical despite dualist intuitions to the contrary.

There are, however, a number of well-known dualist arguments that aim to show that the experiential cannot be physical given what we know about the experiential. First, there is David Chalmers’ (1996) zombie argument (or conceivability argument), which concludes that physicalism is false from the premises that zombies are conceivable and that if they are conceivable, they are metaphysically possible. Second, there is Frank Jackson’s (1982) knowledge argument, which aims to show that physicalism is false on the grounds that one can know all the physical facts without thereby knowing the facts about “what it’s like” to have particular conscious experiences. Third, there is Saul Kripke’s (1980) modal argument that physicalism is false because identity is a necessary relation, mind-body identity statements seem contingent, and we cannot explain away this appearance of contingency in the ways we can with other theoretical identities. Fourth, there is the explanatory gap argument, which concludes that physicalism is false because the physical is exhausted by the structural and causal-functional properties describable in an ideal physical theory and no such theory can render intelligible the relation between physical reality and our subjective, qualitative experiences.188 And fifth, there is Chalmers’ (1996, 2002, 2010) two-dimensional argument that the special semantic character of phenomenal and physical-theoretical concepts precludes the possibility that zombies are conceivable but not metaphysically possible. What all of these arguments have in common is the same basic structure; they begin by noting an epistemic or conceptual gap between phenomenal consciousness and physical reality, and

---

conclude that there is an underlying metaphysical gap responsible for it. In turn, this metaphysical gap shows that physicalism is false.

One of the most promising physicalist responses to these arguments is the phenomenal concepts strategy, according to which the epistemic gap between the experiential and the physical is a result of our special ways of thinking about the experiential. On this view, the prima facie force of the dualist arguments derives from the conceptual and cognitive gaps between our physical-theoretical concepts, on the one hand, and our phenomenal concepts, on the other. There is, the proponent of the phenomenal concepts strategy argues, a dualism of concepts rather than a dualism of ontology. Moreover, according to this phenomenal concepts strategy, the special nature of phenomenal concepts in part explains why phenomenal zombies are conceivable but not metaphysically possible, why there is the mere appearance of contingency in mind-body identity statements, why there is an explanatory gap between the world as described by our physical theories and our conception of conscious experience, and what it is that Jackson’s Mary learns after she leaves the black and white room and has her first color experience.

Of course, there is much dispute about what the special features of phenomenal concepts are that are responsible for lending (false) credence to the dualist arguments. The two leading accounts of phenomenal concepts are what I call “the demonstrative-recognitional model” and “the constitutional model”.

Demonstrative-recognitional accounts liken phenomenal concepts to familiar indexical and demonstrative expressions in natural language. First of all, they hold that phenomenal concepts can only be acquired when one stands in the right sort of contextual
relation to the experiential; one must have the experience before one can think about it using a phenomenal concept. Secondly, they maintain that phenomenal concepts refer to the experiential directly rather than through any descriptive ‘mode of presentation’. And third, they argue hold that the conceptual and epistemic gaps between thinking about the experiential and thinking about the physical are parallel to familiar gaps between “perspectival” thought and objective thought.189

Constitutional accounts hold that phenomenal concepts are at least in part (physically) constituted by instances of the phenomenal properties for which they stand. On such accounts, phenomenal concepts designate phenomenal properties directly by presenting those very phenomenal properties in our thoughts about them. Often this point is put by saying that phenomenal properties serve as their own ‘modes of presentation’ in thought. It is also sometimes claimed that thinking about the experiential is similar to linguistic quotation in that what is designated is exemplified in the thought about it. In either case, the idea is that phenomenal properties are (physically) present in our thoughts about them when we use phenomenal concepts, and this explains why we have substantive knowledge of the qualitative nature of the experiential.

I think that neither demonstrative-recognitional accounts nor constitutional accounts of phenomenal concepts are wholly satisfactory. The problem with demonstrative-recognitional accounts of phenomenal concepts is that the epistemic and conceptual gaps between perspectival thought and objective thought are too thin to explain fully the epistemic gaps between our thought about the experiential and our thought about the physical. This suggests that such accounts fail to identify the crucial feature of our epistemic situation with respect to consciousness (and to the physical) that

189 See Perry (1979).
explains why there is an epistemic gap but no corresponding ontological gap. On the other hand, the problem with constitutional accounts of phenomenal concepts is that they make the relation between the experiential and our thought about it too intimate. There are, as we’ll see, good reasons to suppose that the experiential and our thought about it are, as Hume puts it, ‘distinct existences’.\textsuperscript{190} In any case, it is unclear how the proposed physical intimacy between the experiential and our thought about it is supposed to constitute the epistemic intimacy for which proponents of the approach aim to account.\textsuperscript{191}

To foreshadow things a bit, I think that once we recognize the role that acquaintance plays in reference to, and cognition of, the experiential we can provide a better model of phenomenal concepts. On the account I will develop in the following chapter of this dissertation, phenomenal concepts are concepts of physical/phenomenal properties with a distinctive cognitive role; their reference is fixed by an inner demonstration to states to which we stand in special causal-informational relations (i.e. acquaintance relations), where these special relations are tied to special epistemic methods and capacities for picking up information about the states to which we are related and that provide us with distinctive conceptions of them. Indeed, what is fundamentally special about phenomenal concepts is that their reference is grounded on what I called in the previous chapter our ‘phenomenal awareness’ of the experiential.

Nevertheless, I contend that the special acquaintance relations that we bear to our own states of conscious awareness, those that ultimately ground our phenomenal concepts, do not have epistemic or semantic features that threaten physicalism. For as I argued in Chapter 1, when we are acquainted with a thing, we do not thereby have

\textsuperscript{190} Hume (1739).
\textsuperscript{191} Levine (2007), p. 163.
revelatory knowledge, complete semantic knowledge, or foundational knowledge about that thing. It is a mistaken presupposition underlying dualist arguments that given our acquaintance with the experiential and the physical, we can a priori determine that their natures are fundamentally different in kind. Our acquaintance with the experiential and the physical is simply not revelatory of their natures in the requisite sense. On the contrary, the radically different ways in which we are acquainted with the experiential and the physical—through our phenomenal awareness and intentional awareness respectively—explain why dualist intuitions seem powerful, even though they are mistaken.

But before I turn to my own account of phenomenal concepts in the next chapter, I proceed as follows. In §2, I say more about physicalism and present the main dualist arguments against it. In §3, I explain how the phenomenal concepts strategy aims to undermine these dualist arguments. In §4, I present the demonstrative-recognitional model of phenomenal concepts and show why it fails to explain fully the epistemic gap between the experiential and the physical. In §5, I present the constitutional model of phenomenal concepts and say why it too is unsatisfactory. In doing so, I will set the stage to discuss my own acquaintance-based model of phenomenal concepts in the next chapter of the dissertation. Ultimately, what I hope to show is that physicalists can explain why there is an epistemic gap between the experiential and the physical without concluding that there is an ontological gap or going beyond their explanatory resources.
§ 2 THE CASE AGAINST PHYSICALISM

The consensus among contemporary philosophers of mind is that the greatest challenge to physicalism is the problem of sentience, or conscious awareness. For the purpose of this chapter, I will follow Quine (1984) in describing physicalism as the view that “everything in the world comes down to elementary particles or microphysical events, whatever the point of view…Everything in the world. The qualification is significant, for it sets aside the question of abstract objects such as properties and numbers.” Of course, there are many ways we might understand the claim that “everything in the world comes down to elementary particles or microphysical events.”

Minimally, we can say that if physicalism is true, then the physical facts in the world fix all the facts in the world (aside from facts about the abstract). More strongly, we might say that all the facts in the world, however they might be described, just are physical facts. Even more strongly, we might say that physicalism is the view that there is a conceptual entailment between the world as described by a complete and ideal physical theory and all other true descriptions of the world. I think that this strongest claim is surely false, and in characterizing physicalism generally, I prefer the first, minimal way of understanding it. And in the specific case of the experiential, I think there is good reason to hold that facts about it are numerically identical to neurobiological facts described in physical-theoretical vocabulary.

One reason I think that the experiential is numerically identical with the physical is that I hold the following argument to be sound. First, the physical domain is a causally-closed system and there is no problematic causal over-determination. Second,

---

192 P. 308.
193 Version of this argument have been put forward by Lewis (1966), Kim (1989), Perry (2001), and Papineau (2002).
states of sentient awareness, and other instances of mental kinds, are real causes of some physical states and events, and real effects of others. Therefore, states of sentient awareness, and other instances of mental kinds, are numerically identical with some physical states or events. Of course, this argument leaves it entirely open whether, and in what sense, sentience is ‘reducible’ to some privileged level of physical reality. What it does not leave open, however, is the possibility that the physical domain lacks features that we know sentience to have, or vice versa. This fact follows from the very nature of numerical identity.

I want to note at the outset that this way of framing the problem of sentience assumes traditional ways of carving up reality that might strike some as fundamentally wrong-headed. John Searle (2004), for example, has repeatedly urged that we give up the often misleading traditional philosophical jargon of “the physical”, “the mental”, “physicalism”, “dualism”, etc. Instead, he argues, we should distinguish between those states or events that have a “first-person ontology” and those that have a “third-person ontology”, where this is to be glossed as the distinction between states or events whose existence depends on the existence of a conscious subject, and those whose existence does not. In a similar vein, Galen Strawson (2008) has recently argued that any sensible debate about the metaphysics of the mind must rather be couched in terms of “the mental” and “the non-mental”, or, better, “the experiential” and “the non-experiential”.194 While I agree with both Searle and Strawson that the traditional vocabulary in the philosophy of mind can be misleading, I find it to be sufficiently useful as long as we are careful, and so I will use it.

194 See pp. 8-9.
Setting such terminological issues aside, the problem of sentience can be broken into at least two components. First, there is the problem of explaining why, in a purely physical universe, our states of conscious awareness have qualitative, subjective features such that there is “something it is like” for us to be in them. Finding a place in physical reality for these so-called phenomenal properties, raw feels, or qualia has been dubbed “the hard problem of consciousness”. Second, there is the problem of giving a physically and epistemologically acceptable account of how sentience, by means of sense-perception, provides creatures like us with cognitive access to the world. In the previous chapter of this dissertation, I gave my answer to this second component of the problem of sentience. I argued that we are aware of the qualitative character of our sensations simply in virtue of having them, and we are non-inferentially aware of external objects and their perceptible features by being attuned to what the occurrence of our sensations tells us about the rest of the world. In this section, I turn to articulating the first component of the problem: the problem for physicalism for accounting for the qualitative, subjective features of conscious experience.

Here is the problem in a nutshell: If physicalism is true, then the experiential truths supervene on the physical truths with at least metaphysical necessity. That is, if \( P \) is a complete description of all of the physical truths in the world, and if \( E \) is a complete description of all of the experiential truths in the world, then necessarily if \( P \) is true, \( E \) is true as well. However, dualists argue that it is not the case that necessarily \( P \)

---

197 Stoljar (2005) calls the consequent of this sentence the ‘psychophysical conditional’. Following standard practice, I have framed the hard problem of consciousness in terms of supervenience rather than identity here. However, some dualist arguments, such as the modal argument, are aimed specifically at the identity theory. Moreover, since I advocate the identity theory with respect to the experiential, my arguments will defend a stronger form of physicalism than the supervenience version I articulate here.
metaphysically settles the truth of $E$. They claim that at most $P$ settles the truth of $E$ in this world because of contingent laws of nature that link the physical truths of this world with the experiential truths of this world. But in other worlds with very different psychophysical laws, the $P$ truths might be just as they are in this world, but the $E$ truths might be different. There are five arguments which aim to show that the experiential fails to supervene on the physical with metaphysical necessity: David Chalmers’ zombie argument, Frank Jackson’s knowledge argument, Saul Kripke’s modal argument, Joseph Levine’s explanatory gap argument, and Chalmers’ two-dimensional argument. Below, I will present each of these arguments in turn:

1. **The Zombie Argument**

   A philosophical zombie is a creature who is in every physical respect indiscernible from a normal conscious human being and yet who lacks conscious experience altogether. Whereas there is “something it is like” for a normal conscious human being to undergo various happenings, there is nothing it is like for a philosophical zombie to undergo them. Metaphorically, for such a creature, all is dark inside. Nevertheless, it behaves in every respect just as a normal conscious human would in the same circumstances. In fact, it even thinks and talks just as a normal conscious human would (assuming thought does not essentially involve consciousness); however, it simply has no experiential life.

   According to the zombie argument, it is at least conceivable that there are philosophical zombies (or at least that there are physically indiscernible creatures that differ with respect to some aspect of their experience from ours). What’s more, the argument goes, philosophical zombies are conceivable in the right sort of way so that
their conceivability tracks metaphysical possibility: an ideal reasoner can positively imagine a scenario in which there are zombies, so they are metaphysically possible. But if they are metaphysically possible, then it is not the case that necessarily the physical truths in this world settle the experiential truths. Hence, physicalism is false.

We can summarize the zombie argument as follows:

1) Zombies are conceivable (P & ~E is conceivable).

2) If zombies are conceivable, then they are metaphysically possible (if P & ~E is conceivable, then P&~E is metaphysically possible).

3) If zombies are metaphysically possible, then physicalism is false (if P&~E is metaphysically possible, then physicalism is false).

4) Therefore, physicalism is false.

The crucial premise here is premise two, which asserts that there is a genuine link between the right sort of conceivability and metaphysical possibility. However, it is important to note that it gets its strength from the intuition we have a sufficiently substantive grasp of the nature of the experiential and the physical to draw a priori conclusions about their metaphysical connection. For where we do not have sufficiently revealing conceptions, even a dualist will grant that conceivability is no sure guide to metaphysical possibility.

2. The Knowledge Argument

According to Frank Jackson’s famous knowledge argument, Mary is a brilliant scientist who is raised in a black and white room. In the black and white room, Mary

---

198 Some physicalists, such as Dennett (2006), deny that phenomenal zombies are even conceivable. This is an important response to the zombie argument, but I will not consider it here, partly because I am skeptical of its prospects. In any case, I think it is dialectically better for the physicalist to grant the dualist the premise that zombies and inverts are at least conceivable, but show why it doesn’t follow that they are metaphysically possible.
acquires complete knowledge of the workings of the physical domain and all of the physical information about her own perceptual faculties, the perceptual faculties of others, the perceptible properties of physical objects, and the physical processes involved in visual perception. However, she has never actually had a normal perceptual experience of chromatic color properties. One day, Mary is released from the black and white room and has a normal perceptual experience of red for the first time.

According to Jackson’s knowledge argument Mary clearly learns something new when she has her first perceptual experience of red; in particular, she learns what it is like for her to have normal perceptual experiences of red. What’s more, she learns what it is like for others, similarly endowed and situated, to have normal perceptual experience of red. These new items of knowledge that Mary learns are facts about conscious experiences that she didn’t already know in the black and white room. But by hypothesis, she already knew all of physical facts in the black and white room, so what she learns must be nonphysical facts. Therefore, physicalism is false.

We can summarize the knowledge argument as follows:

1) Before her release, Mary knows all of the physical facts.
2) After her release, Mary learns what it is like to have various experiences.
3) Thus, after her release, Mary learns new facts about conscious experience.
4) So before her release, Mary does not know all of the facts.
5) Therefore, the physical facts do not exhaust all of the facts, and physicalism is false.

The crucial premise of the knowledge argument is premise three, which asserts that the change in Mary’s epistemic life after her release consists in her acquisition of knowledge
of new facts that she didn’t already know in the black and white room. It is worth noticing, however, that the plausibility of this premise depends on our antecedent conception of what the physical and experiential facts are. For if we do not have a sufficiently substantive grasp of the nature of the experiential and the physical, we cannot conclude that Mary learns a new fact rather than an old fact in a new way.

3. The Modal Argument

Assuming that “the experiential” and “the physical” are rigidly-designating natural kind terms, then if the experiential is identical with the physical, it is necessarily so. This is an essential fact about the relation of identity, for one thing cannot be two things. It also follows from the nature of rigid-designation; a rigidly-designating term refers to the same thing in every possible world, so two rigid-designators that refer to the same thing in one world refer to the same thing in all worlds. On the other hand, if two rigid-designators refer to different things in any world, they refer to different things in all worlds. Thus, true identities are necessarily true, and false identities are necessarily false. There simply are no contingent identities.

Whether or not a particular identity is a priori knowable is a separate question. Many identities are not. For example, the identity between water and $H_2O$ is an empirical discovery, as is the identity of heat and mean kinetic energy. Still, these are necessary identities if they are identities at all. But then why does it at least seem contingent that water is $H_2O$ or heat is mean kinetic energy? Couldn’t it have turned out that water is not $H_2O$ or heat is not mean kinetic energy? If not, then what then are we imagining when we imagine it could have?

199 Some physicalists, such as Dennett (2006), will reject premise two and deny that Mary learns anything by having her first experience of chromatic color. Once again, I think this is an important response, but I find it a counterintuitive and dialectically unpromising position.
According to Kripke, when we imagine that water is not H₂O or that heat is not mean kinetic energy, what we are really imagining is that the sensations we use to pick out instances of water or heat could have been caused by something else. That is a coherent possibility, but not a possibility in which water is not H₂O or in which heat is not mean kinetic energy. For if the identities are true, there are no such possibilities.

In the case of the experiential and the physical, their putative identity also strikes us as contingent. We can imagine the experiential in a world devoid of the physical and the physical in a world devoid of the experiential. In fact, this is exactly the scenario we envisioned above when discussing the zombie argument. But is this scenario a coherent metaphysical possibility? It certainly seems to be, at least provided we are not making a similar mistake as when we imagine that water is not H₂O or heat is not mean kinetic energy. For, again, if we are making a similar mistake, then the possibility we think we are imagining, that the experiential is distinct from the physical, is not really a possibility.

According to Kripke, however, we are not making the same kind of mistake in the case of the experiential and the physical as we make when we take ourselves to be imagining that water is not H₂O or heat is not mean kinetic energy. Whereas in the case of water or heat we could exploit the fact that our method of picking out the relevant phenomenon (through our sensations) is only contingently related to the phenomenon itself, he argues that there is no contingency between the way we pick out the experiential and the experiential itself. Indeed, he contends, we pick out the experiential directly by grasping its essential nature—by grasping what it is like to have the relevant experience. For example, we pick out pain by what it is like to have pain, by what it feels like. But anything that feels like pain, is pain (and vice versa). So unless a physicalist can explain
what other sort of mistake we are making when we imagine the experiential and the physical to be distinct, we should simply conclude that there is no mere ‘illusion of contingency’. But there is no plausible way for the physicalist to explain away the prima facie illusion of contingency, so there is genuine contingency. Hence, physicalism is false.

We can summarize the modal argument as follows:

1) The relation between the experiential and the physical seems to be contingent.

2) If we cannot explain away why a relation seems to be contingent, then it is contingent.

3) We cannot explain away the apparent contingency between the experiential and the physical.

4) Therefore, the relation between the experiential and the physical is contingent, and physicalism is false.

The crucial premise of the modal argument is premise three, which contends that we cannot explain away the apparent contingency between the experiential and the physical. Its plausibility stems from two assumptions. The first assumption is that we grasp the essential nature of the experiential. The second assumption is that there is no other source of contingency in the offering that can explain away the illusion of contingency. Of course, if the physicalist can put pressure on either assumption (or both), premise three is false and Kripke’s conclusion does not follow.

4. The Explanatory Gap

Suppose that we have a complete theory that tells us that the experiential is identical with the physical. Suppose that the experiential is perfectly correlated with some
aspect of the physical domain, that they share a causal role, and that the simplest explanation of these facts is that the experiential and the physical are identical. Still, one might ask, why does it feel like this to be in such and such physical state rather than that? Why are particular physical states accompanied by particular experiential characters rather than others or none at all? Evidently, such questions are left open by a complete physical description of the world, which suggests that there is a deep explanatory gap between the experiential and the physical domain.

According to the explanatory gap argument, the deep explanatory gap here is one that is not found elsewhere in the physical domain. For example, once we have identified water with $\text{H}_2\text{O}$ or heat with mean kinetic energy, there is no further question about why we find water whenever we find $\text{H}_2\text{O}$ or find heat whenever we find mean kinetic energy. Yet patently we are faced with such a further question in the case of the experiential and the physical. For many, the novelty of this gap suggests that there is an underlying ontological gap responsible for it.

The problem, Joseph Levine argues, is that the nature of consciousness is such that we cannot rely on the standard explanatory schemes we use in other domains. Typically, he claims, what we do is give some structural-functional characterization of the phenomenon for which we wish to give a scientific explanation and then we identify the physical phenomenon with the right sort of features to fit our structural-functional characterization. In turn, we explain how the relevant physical phenomenon can possess the features that fit our structural-functional characterization by appealing to the more fundamental structural-functional properties of its constituent physical parts. However, Levine argues that this sort of explanation cannot work in the case of conscious
experience precisely because the subjective and qualitative character of conscious experience eludes structural-functional characterization. While he grants that physicalism might be true, he contends that we simply cannot see how given that we do not have any model in the physical sciences that can so much as make intelligible how the experiential might be physical.

Recently, Chalmers has transformed the brute intuition that the explanatory gap is grounded in an underlying ontological gap into an argument for dualism. Chalmers argues that the reason that there is an explanatory gap between the experiential and the physical is that the experiential is subjective and qualitative rather than structural or functional and physical accounts explain at most structure and function. That is, physical accounts only explain how the physical is distributed in space and time and how the physical behaves in various circumstances. However, neither explaining the distribution of the physical in space and time nor explaining how the physical behaves in various circumstances suffices to explain what it is like for a subject to have such and such experiences. Therefore, the experiential is distinct from the physical.

We can summarize the explanatory gap argument as follows:

1) Physical accounts explain at most structure and function.

2) Explaining structure and function does not suffice to explain consciousness.

3) If something cannot be explained by physical accounts, it is not physical.

4) Therefore, consciousness is not physical, and physicalism is false.

The crux of Chalmers’ argument is that we know enough about the physical to know that it can be explained entirely in terms of structure and function and enough about the

---

experiential to know that it cannot be explained in terms of structure and function. If either (or both) of these assumptions is false, then the explanatory gap argument fails.

5. The Two-Dimensional Argument

So far, we have seen four arguments that suggest that the physical facts about this world do not fix the experiential facts about this world with metaphysical necessity. In addition to these four standard arguments against physicalism, Chalmers has recently made a more general case against it on the basis of his sophisticated two-dimensional semantic theory. His argument aims to show both that the experiential is distinct from the physical and why we cannot explain away the apparent contingency between them by appealing to standard Kripkean considerations about a posteriori necessities. Thus, his argument supports Kripke’s contention that we cannot understand the putative identity of the experiential and the physical in the same way as other a posteriori identities. But it goes further in trying to show that the physicalist has no plausible way of resisting the dualist conclusions. In the next chapter of this dissertation, I will show that this two-dimensional argument is unsuccessful, and in doing so, I hope to bring out the core problem with all of the dualist arguments. In the rest of this section, I will simply present Chalmers’ argument.

On Chalmers’ view, linguistic expressions and concepts have two distinct patterns of evaluation. They have a primary intension which characterizes how their reference or extension depends on how the world turns out to be. They also have a secondary intension which characterizes their reference or extension given the way the actual world is.

__201__ Chalmers (2002).
The primary intension of a concept captures its ‘epistemic character’; it specifies what the concept picks out were various epistemically possible circumstances to obtain. When we evaluate the semantic value of a concept using its primary intension, we are evaluating what Chalmers calls ‘epistemic scenarios’ or ‘counteractual possibilities’. Counteractual possibilities are ways the world could be which the speaker cannot rule out a priori. Chalmers uses this to define epistemic or conceptual modality in the following way: A scenario $S$ is conceptually, or epistemically, possible if and only if it is true in some world considered as actual. A scenario $S$ is conceptually, or epistemically, necessary if and only if it is true in all worlds considered as actual. And a scenario $S$ is conceptually, or epistemically, impossible if and only if it is false in all worlds considered as actual.

The secondary intension of a concept captures how it behaves across the space of metaphysical possibilities given its reference or extension in the actual world. When we evaluate the semantic value of a concept using its secondary intension, we are evaluating its reference or extension across counterfactual possibilities. Counterfactual possibilities are, Chalmers argues, genuine metaphysical possibilities. This allows us to define metaphysical possibilities in the following standard way: A scenario $S$ is metaphysically possible if and only if it is true in some world considered as counterfactual. A scenario $S$ is metaphysically necessary if and only if it is true in every world considered as counterfactual. And a scenario $S$ is metaphysically impossible if and only if it is false in every world considered as counterfactual.

According to Chalmers, since the space of metaphysical possibilities turns on what a concept’s reference or extension is in the actual world, there is a crucial link
between its primary and secondary intensions. In particular, we can consider what a concept refers to counterfactually given how its reference varies across epistemically possible scenarios. For example, suppose that the concept “water” refers to the watery stuff around here in any epistemically possible scenario. In some worlds considered as counteractual, it will refer to H$_2$O and in others it will pick out XYZ. We can then determine that “water” refers to H$_2$O counterfactually given that the actual world is one in which the watery stuff around here is H$_2$O. But if the actual world turns out to be one in which the watery stuff around here is actually XYZ, then ‘water’ will refer to XYZ across all metaphysically possible worlds. What we end up with, Chalmers argues, is knowledge of special semantic conditionals: we know that if the world turns out to be one where the watery stuff around here is such and such stuff, then “water” refers to the such and such stuff across all possible worlds.$^{202}$

For Chalmers, most concepts are what we might call “semantically unstable concepts”; they are concepts whose semantic values vary across different worlds considered as counteractual and counterfactual. The concept “water” is a paradigm example of a semantically unstable concept. For as we’ve seen, if we suppose that the concept “water” picks out the watery stuff around here in any world considered as actual, it will pick out different watery substances in different worlds. But then given the link between the primary intension of a concept and its secondary intension, our judgments

---

$^{202}$ Chalmers (2002). Chalmers thinks that semantic competence with a term provides us with at least implicit knowledge of such special semantic conditionals. He maintains that we can a priori determine the primarily and secondary intensions of terms we are competent with across epistemic and metaphysical scenarios provided that we have the right sort of canonical description of the relevant scenarios. The right sort of world description for Chalmers is a complete qualitative characterization of the world, one that relies only on semantically stable expressions and from which all true statements about that world are entailed. While ordinary speakers do not have such complete qualitative descriptions available, he thinks that any counterfactual scenario that is not a priori contradictory will be part of, or entailed by, some such canonical world description. Hence, ordinary speakers can base their semantic judgments on them.
about the counterfactual extension of “water” will be sensitive to the shifting results of its counterfactual patterns of evaluation.

However, not all concepts are semantically unstable. Some concepts have the same reference or extension in all epistemic scenarios and so pick out the same things across all metaphysical possibilities. We might call such concepts “semantically stable concepts”. Chalmers himself describes them as “epistemically rigid” concepts since their “epistemic content picks out the same referent in every possible world (considered as actual)”. Epistemic rigidity is supposed to contrast with what he calls “subjunctive rigidity”, which captures the standard Kripkean insight that some concepts pick out the same referent in every possible world considered as counterfactual. However, Chalmers contends that (most) epistemically rigid concepts will also be subjunctively rigid. This should not be surprising given the alleged link between the primary and secondary intensions of a concept.

There are two sorts of semantically stable concepts that are crucial for Chalmers’ case against physicalism. First, he claims that physical-theoretical concepts are semantically stable. For example, he argues, the concept “hydrogen molecule” rigidly picks out whatever plays the causal role of hydrogen molecules in any world considered as actual. This is because all there is to being a hydrogen molecule is to play the relevant causal role. So given any appropriately described epistemic scenario, “hydrogen molecule” will pick out the same phenomenon. Consequently, the concept “hydrogen molecule” picks out the same phenomenon across all metaphysically possible worlds. For

---

the same reason, Chalmers thinks that our concepts for other fundamental physical phenomenon will share this semantic stability.

Second, Chalmers claims that our “direct phenomenal concepts” are semantically stable. Direct phenomenal concepts are concepts we use to think about conscious experiences when we have them and attend to them. According to Chalmers, direct phenomenal concepts are in part constituted by the very phenomenal properties for which they stand.\textsuperscript{205} Thus, there is an especially intimate manner in which direct phenomenal concepts pick out their referents; their referents are picked out in part by their being exemplified in the concept itself. Chalmers also thinks that there is an especially intimate manner in which we are aware of phenomenal properties when we think about them using direct phenomenal concepts. On his view, we are acquainted with them in a way that provides us with a substantive grasp of their essential natures. Together, the intimate constitution of direct phenomenal concepts and the intimate epistemic relation that makes our acquisition of them possible lends itself to a special sort of incorrigibility in our phenomenal beliefs about our occurrent conscious experiences. In particular, we cannot be mistaken in forming beliefs of the form “\textit{this} experience is such and such experience”, where the such and such experience is thought of under a direct phenomenal concept.\textsuperscript{206}

Chalmers thinks that these special features of direct phenomenal concepts are what make them semantically stable concepts. For, he argues, given these special features, direct phenomenal concepts allow us to pick out the same referents in any world considered as actual. Sometimes he puts the point by saying that phenomenal properties are part of the epistemic core that we take with us whenever we consider epistemic

\textsuperscript{205} We should notice right away that Chalmers’ account of direct phenomenal concepts is very close to physicalist constitutional accounts of phenomenal concepts.

\textsuperscript{206} Chalmers (2004), section 4.
scenarios. Indeed, he contends, they are not at a distance from our epistemic perspective on things, but an essential part of it. Consequently, they do not have shifting counterfactual patterns of evaluation and are both epistemically and subjunctively rigid.

Given that physical-theoretical concepts and direct phenomenal concepts are both semantically stable concepts, Chalmers argues, there is a deep problem for physicalism. The problem is that physicalists cannot explain why it is conceivable that there be scenarios in which the physical facts are the same but the experiential facts are different. According to Chalmers, in standard cases of true a posteriori identities, two concepts have different primary intensions but the same secondary intensions. That is, they pick out different referents in different worlds considered as actual but the same referents in all metaphysically possible worlds. This explains why it is conceivable that the relevant identities are false, but why they are necessarily true. However, in the case of physical-theoretical concepts and phenomenal concepts, we cannot appeal to their distinctness of primary intensions but sameness of secondary intensions, because their primary and secondary intensions are the same in all worlds. Thus, if their primary intensions are distinct, then so are their secondary intensions. It follows that if zombie worlds are epistemically possible, then they are metaphysically possible, and physicalism is false. And zombie worlds are at least epistemically possible. So, Chalmers concludes, physicalism is false.

§3 THE PHENOMENAL CONCEPTS STRATEGY

In the previous section, I presented the leading dualist arguments against physicalism: Chalmers’ zombie argument, Jackson’s knowledge argument, Kripke’s modal argument,
and Levine’s explanatory gap argument. I also presented Chalmers’ two-dimensional regimentation of the conceivable/modal argument against physicalism. What all of these arguments have in common is that they aim to show that the physical truths in the world do not settle the experiential truths in the world with metaphysical necessity. They do this by pointing to an epistemic or conceptual gap between the physical and the experiential and then inferring that there is an underlying ontological gap. In this section, I will present a common physicalist strategy, the phenomenal concepts strategy, for granting the epistemic or conceptual gap but denying that it has ontological implications. I will also show how this general strategy aims to undermine the standard dualist arguments.

As we saw in the previous section, physicalism is committed to the view that the physical truths in the world settle the experiential truths in the world with metaphysical necessity. We also saw that standard arguments for dualism try to put pressure on physicalism by noting that there is no a priori conceptual entailment between the world described by means of physical vocabulary and the world described by means of phenomenal vocabulary. From this dualists infer that the experiential is distinct from the physical. But this is too quick.

Many physicalists point out that physicalism is only committed to the view that the physical facts settle the experiential facts with metaphysical necessity, not logical necessity. And we know from standard Kripkean cases that there are metaphysical necessities that are not logical necessities. For example, we know that if water is H₂O, then it is necessarily H₂O despite the fact that there is no logical entailment between the concept “water” and the concept “H₂O”. That is why we could not discover that water is
H$_2$O simply by reflecting on the relevant concepts; we needed to do a posteriori empirical investigation to determine that fact. Likewise, we might need empirical evidence to conclude that Clark Kent is Superman, but Clark and Superman are one and the same person nonetheless (and necessarily so). So, many physicalists point out, it is not the case that metaphysical necessity implies logical necessity.

The proponent of the phenomenal concepts strategy argues that we are in roughly the same position with respect to the experiential and the physical as, we might imagine, people once were with respect to water and H$_2$O before it was discovered that water is H$_2$O. This is because we have two distinct kinds of concepts, physical-theoretical concepts and phenomenal concepts. There is no a priori logical entailment between them. But they rigidly pick out one and the same phenomenon.

But what are physical theoretical concepts and phenomenal concepts? And how do they differ from one another? To a first approximation, we can say that physical-theoretical concepts are the familiar concepts provided by the natural sciences. They are paradigmatically the sort of concepts one find in physics books, chemistry books, biology books, etc. While it is no trivial matter to specify necessary and sufficient conditions for a concept’s being a physical-theoretical concept, we can say roughly that the unifying feature of such concepts is that they have been introduced through empirical investigation and theory-building to stand for natural phenomena in the physical world. Their principal job, one might say, is to provide us with the explanatory resources to characterize how the world works.\textsuperscript{207}

\textsuperscript{207}I am here assuming a realist understanding of physical-theoretical concepts. Since I am only aiming to give an approximate characterization of physical-theoretical concepts, I am ignoring alternative anti-realist understandings of them.
Phenomenal concepts, on the other hand, are concepts we use to think about the qualitative character of conscious experiences. In their most fundamental uses, we use them to think about the character of our own conscious experiences in the familiar first-personal way. This allows us to form various kinds of useful beliefs about our own conscious experiences. For example, we can use phenomenal concepts to classify and compare our experiences in terms of what it is like to have them. We can also use them to remember what experiences we’ve had in various circumstances, and recognize those experiences when we have them again. We can use them to anticipate and imagine the experiences we might have in various circumstances so as to put ourselves in a position to avoid them or have them again. Finally, we can, at least in some circumstances, use phenomenal concepts to help make judgments about the epistemic credentials of our own experiential states.

In less fundamental uses, we use them to think about the conscious experiences of others. This allows us to understand and predict the behavior of other sentient beings or to make judgments about the epistemic credentials of their experiential states. Such third-personal uses of phenomenal concepts play a central role in the ‘folk psychology’ we use to understand and predict the behavior of others.

There are, however, important epistemic asymmetries between first-personal and third-personal uses of phenomenal concepts. When an individual uses a phenomenal concept to make a first-person avowal about the character of his or her current experience, we typically grant his or her avowal special authority, or even incorrigibility.\(^208\) On the other hand, when an individual uses a phenomenal concept to

\(^{208}\) As we will see, however, we needn’t (and shouldn’t) grant *infallibility* to first-personal uses of phenomenal concepts.
make a judgment about the character of another’s experience, we do not grant his or her judgment any special authority. One reason for this asymmetry is that in the first-personal case, an individual has special access to his or her experiential life on which to base the phenomenal judgment. In contrast, in third-personal uses of phenomenal concepts, individuals do not have special access to the experiential lives of others; indeed, they must rely entirely on third-personal methods (such as testimony and observation of behavior) for ascertaining what is going on in the consciousness of others.

Another special feature of phenomenal concepts is their special acquisition or possession conditions. Unlike many other kinds of concepts, an individual cannot acquire a genuine phenomenal concept through linguistic transmission alone; an individual cannot simply be taught a phenomenal concept. He or she must actually have the experience, and attend to it, before he or she can acquire a genuine phenomenal concept of it. Of course, that is not to say that an individual cannot refer to experiences using a public concept for them before he or she has the relevant experiences, for individuals can clearly accomplish that. But it is to say that such reference is not made by using a genuine phenomenal concept.\(^{209}\)

As for other special features of phenomenal concepts, there is little consensus. Some think that phenomenal concepts behave in crucial respects like demonstrative or recognitional concepts. Others maintain that a special feature of phenomenal concepts is that they are in part constituted by the very phenomenal properties for which they stand. We will discuss each of these proposals in greater detail in the next two sections of this

\(^{209}\) On my view, the reason that the public concept does not count as the same concept as the phenomenal concept is that the two concepts are grounded on very different ways of being acquainted with the relevant experiences: by hearing about them and by having them.
chapter. As for the rest of this section, I will briefly say a little about how the appeal to phenomenal concepts can help to undermine the standard dualist arguments.

In the case of the zombie argument, the deep conceptual gap between physical-theoretical concepts and phenomenal concepts can be used to block the inference from the conceivability of zombies to their metaphysical possibility. This is because we have two radically different concepts that rigidly pick out the same phenomenon. The experiential-physical identity is necessary, but a posteriori. So, just as in the case of water and H₂O and Superman and Clark Kent, we can conceive of the distinctness of the experiential and the physical, but our conception of distinctness here is misleading. Thus, physicalism is not threatened by the mere conceivability of zombies.

In the case of the knowledge argument, we also have a ready explanation of the change in Mary’s epistemic situation. What happens when Mary is released from the black and white room and experiences red for the first time is that she acquires a new concept, one she simply could not possess prior to her release. This new phenomenal concept allows her to think about her experiences and the experiences of others in ways she couldn’t in the black and white room. But it doesn’t follow that what she is thinking of is a new non-physical fact, for she could simply be thinking of an old fact in a new way. For instance, one does not think of two distinct facts when one thinks ‘water is wet’ and ‘H₂O is wet’. One simply thinks the same fact in two different ways. If the phenomenal concepts strategy is correct, then the same lesson applies in the case of thinking about the phenomenal/physical using phenomenal concepts and physical-theoretical concepts. So, while the physicalist can grant that Mary learns something when
she leaves the black and white room, it doesn’t follow that what she learns is a new fact. Hence, her purported new knowledge is no threat to the truth of physicalism.

In the case of the modal argument, there is less agreement about how phenomenal concepts help physicalism. Some hold that the appearance of contingency in psychophysical identities is the result of the real contingency of physical-theoretical concepts and phenomenal concepts both standing for the same phenomenon.\textsuperscript{210} Others hold that the appearance of contingency is the result of the fact that phenomenal concepts refer to phenomenal/physical properties by exemplifying them and physical-theoretical concepts do not.\textsuperscript{211} As we’ll see, I think that the illusion of contingency is a result of our radically different conceptions of the world that are provided by physical-theoretical and phenomenal concepts. Whatever the case may be, the hope is that some special feature of phenomenal concepts will explain why it at least \textit{seems} that the experiential is distinct from the physical, despite the fact that it is not. If such a case can be made, then the modal argument too loses its force against physicalism.

As with the modal argument, there is little agreement about how phenomenal concepts are supposed to help with the explanatory gap argument. Nevertheless, many have the intuition that the special nature of phenomenal concepts lets us think about the physical in a way that we do not when we use physical-theoretical concepts. The uniqueness of this special way of thinking is such that we don’t know how to connect what we thereby know with our standard explanatory models (which only use physical-theoretical concepts). But, the proponent of the phenomenal concepts strategy argues, the

\textsuperscript{210} Perry (2001).
\textsuperscript{211} Papineau (2002), Block (2007), and Balog (forthcoming).
explanatory independence of phenomenal concepts does not have metaphysical import. Hence, we have no reason to abandon physicalism.

§4 THE DEMONSTRATIVE-RECOGNITIONAL MODEL

In the previous section, I made some general remarks about the nature of phenomenal concepts and how they might help the defense of physicalism in the face of the standard dualist arguments. In this section, I present the demonstrative-recognitional model of phenomenal concepts in greater detail. First, I note the features that proponents of the model take phenomenal concepts to have in common with more familiar demonstrative and recognitional concepts. And second, I argue that though the demonstrative-recognitional model correctly identifies many of the special features of phenomenal concepts, it cannot explain the substantive epistemic gap between our conception of the experiential and our conception of the physical. This suggests that there is a more fundamental source of the epistemic gap than accounted for by the demonstrative-recognitional model.

One of the best recent discussions of the demonstrative-recognitional model of phenomenal concepts is Janet Levin’s in “What is a Phenomenal Concept?” In her paper, Levin argues that:

There are two essential features of demonstratives that phenomenal concepts, on this view, are taken to share: first, they pick out their referents from a particular point of view—the perspective of the demonstrator—and thus are not equivalent to any nonperspectival (discursive, objective) concepts; second, they can pick out their referents “directly,” without need of identifying modes of presentation.212

---

In what follows, I will say more about how we should understand each of these claims. However, I would like to first point out another essential feature of phenomenal concepts on the demonstrative recognitional model: that our acquisition of them depends on our standing in the right sort of contextual relation to the experiential; we must have the relevant experience in order to acquire a genuine phenomenal concept of it. We simply cannot think of experiences using phenomenal concepts, concepts of what it is like to have particular experiences, without having them, or at least without having relevantly similar experiences.

I will begin by turning to the second of the two features that Levin notes are shared by phenomenal concepts and ordinary demonstrative recognitional concepts—that phenomenal concepts pick out their referents directly rather than through identifying modes of presentation. There are basically two conceptions of semantics which have guided our theorizing about the contribution of referring devices to the meaning of our thought and talk. On the standard descriptivist conception of semantics, most often associated with Frege, our concepts of things contribute to the contents of our thought and talk not the objects our concepts are about, but rather some conceptually grasped identifying conditions, or ways of getting at those things. Worldly objects, properties, and relations are the subject matter of our thought and talk, according to this theory, in virtue of those objects, properties, and relations uniquely satisfying the identifying conditions (or “modes of presentation”) associated with our concepts. Such a conception of the semantics of referring devices has aptly been called an “indirect” theory of reference since, on this view, all designation of worldly objects, properties, and relations is
mediated through representational fit.\textsuperscript{213} What we think about, if anything, is that which fits the contents of our thought or talk.

There is a competing conception of semantics, however, according to which referring devices designate the worldly objects, properties, and relations they are about directly.\textsuperscript{214} On this directly referential conception of semantics, our concepts of things contribute to the contents of our thought and talk the objects or properties themselves rather than ways of getting at them. Often, this conception of semantics is described as holding that referring devices are mere “tags” for objects and properties, tags which have been assigned to their referents by linguistic convention or cognitive design.\textsuperscript{215} But though most self-professed direct referentialists eschew the idea that designation is mediated through representational fit, or that indentifying conditions are part of the content of our thought and talk, direct referentialists do not deny that there are explanations of the intervening mechanisms by which our concepts come to stand for their designated objects, properties, or relations.

Proponents of theories of direct reference have often pointed to causal and historical processes as the underlying ground of the referential relations our concepts bear to worldly items and their features. In its crudest form, the directly referential conception of semantics holds that the meaning of a concept consists in the worldly objects, properties, or relations that in part constitute its causal origin. For this reason, the direct theory of reference is often simply called “the causal theory of reference.”

\textsuperscript{213} Kaplan (1977/1989). For simplicity, I am here ignoring the important distinction between singular reference and direct reference.

\textsuperscript{214} This conception of semantics finds early articulation in the writings of Locke and Mill, but its contemporary resurgence can be traced back most directly to Russell work on acquaintance and logically proper names. For an excellent discussion of these issues, see A.Capuano (In Draft).

\textsuperscript{215} Opinions differ about exactly which causal-historical, phylogenetic, and ontogenic factors are responsible for imbuing our cognitive lives with their content, and to what extents. I will remain largely neutral on these issues.
One reason for preferring indirect theories of reference is that they have intuitive explanations of well-known co-reference and no-reference problems. On the other hand, direct theories of reference fare better with respect to capturing the counterfactual truth conditions, conditions for same-saying or –thinking, and practical/cognitive role of our thought and talk about things. Proponents of the demonstrative-recognitional model of phenomenal concepts find these latter considerations more compelling than those that favor indirect theories of reference (as they should), and hence think that all singular referential concepts, including phenomenal concepts, are directly referring devices.

In the case of phenomenal concepts, proponents of the demonstrative-recognitional model maintain that they refer directly to phenomenal/neurobiological properties that play the right sort of role in the life of the conscious subject. Typically, a phenomenal concept will refer directly to whichever phenomenal/neurobiological property kind (or instance) that constitutes the causal origin of the formation of the concept.\textsuperscript{216} Levin argues that:

\begin{quote}
Most physicalists who take phenomenal concepts to function like demonstratives suggest that they denote whichever neural properties are causally responsible for our application of these concepts in various introspective tasks.\textsuperscript{217}
\end{quote}

However, Levin contends that the reference of a phenomenal concept is not determined \textit{entirely} by its causal origin. In addition, the reference of a phenomenal concept is also partly determined by the subject’s dispositions to reapply the concept in further cases.

\textsuperscript{216} In atypical circumstances, perhaps a phenomenal concept will refer directly to whichever phenomenal/neurobiological property is relevantly similar to the causal origin of the formation of the concept. I have in mind here Hume’s missing shade of blue which we can perhaps form a phenomenal concept of on the basis of having experiences of similar shades of blue. I leave the matter open.

There are two reasons that the subject’s dispositions to reapply the concept are relevant to determining the reference of a phenomenal concept, according to Levin. First, the subject’s dispositions partly determine whether the concept is a token phenomenal concept or a type phenomenal concept. Token phenomenal concepts are concepts used to think directly about particular dated, unrepeatable experiences or phenomenal property-instances. As Levin puts it:

Token phenomenal concepts are those that can be used to pick out an instance of an experience with some salient qualitative character (Tye 1995), and they are taken to denote in the manner of token-demonstrative concepts. That is, whichever (neural) particular causes me to make introspective note of some experience I’m now having counts as the denotation of the token-demonstrative “that (experience I’m having now).”

Roughly speaking, a phenomenal concept is a token phenomenal concept only if a subject is disposed to use it to think about a particular dated, unrepeatable experience or phenomenal property-instance and is disposed not to use it to think about others.

Type phenomenal concepts, on the other hand, are concepts used to think directly about repeatable kinds of experiences or phenomenal properties. A phenomenal concept is a type phenomenal concept only if a subject is disposed to use it to pick out instances of a kind of experience or phenomenal property, particularly in the subject’s anticipatory, classificatory, or recognitional thinking about the experiential. Many proponents of the demonstrative-recognitional model place particular emphasis on the role that recognition plays in determining the reference of type phenomenal concepts. For example, Levin contends that:

The denotation of a phenomenal type-demonstrative will be the property—presumably physical—that’s causally responsible for the application of that concept in the introspective recognition or reidentification of an experience as

---

“that (kind) again” or “another of those.” These concepts…are taken to refer “directly”; that is, to have no reference-fixing “modes of presentation” or Kaplanian “characters” that can change reference from world to world…Rather, their references are determined solely by the causal and dispositional relations an individual has to her internal states that are effected by an introspective “pointing in”; that is, by the fact that she’s in causal contact with a certain property and is disposed to reidentify it on subsequent occasions.²¹⁹

It is worth noting that for Levin a subject needn’t be successful in every case in recognizing further instances of an experiential kind; it is sufficient that a subject is simply disposed to do so. This is important since it is implausible to suppose that subjects are infallible in judging when they are having such-and-such kind of experience again (despite the fact that many dualists seem to be committed to such an assumption).

The second reason that a subject’s dispositions to reapply the concept are relevant to determining the reference of a phenomenal concept, Levin argues, is that without such dispositions we couldn’t achieve determinate reference to experiential kinds. Levin’s worry here is basically a version of the well-known “qua problem” for causal theories of reference. The idea behind the qua problem is this: We often refer to things of various different sorts, but causation is too rough-grained to single out a thing of one ontological kind rather than another where the relevant sorts can be co-located or co-instantiated, as is often (always?) the case. For example, in the case of a phenomenal experience as of seeing a scarlet circle in standard viewing conditions, the subject stands in a causal relation to scarlet, to red, to a primary color, to a color, to a circle, to a shape, to a

²¹⁹ Levin (2007), p. 89. Some proponents of the demonstrative-recognitational model go on to explain a subject’s dispositions to reidentify instances of the same kind of experience or phenomenal property in terms of the subject having an internal ‘template’ for the relevant experiential kind. One might think of such templates as what John Perry calls a “Humean core”—a faint impression of the experiential kind which a subject can rely on in discriminatory and recognitional tasks. Whatever the case may be, most proponents of the model agree with Levin that recognitional capacities, however grounded, are essential for determining the reference of type phenomenal concepts. This is compatible with our getting it wrong on particular occasions, however.
concrete object, etc. But the causal relation itself arguably does not discriminate between these different sorts. So, which sort of thing does a concept grounded on the causal relation refer to? The worry is that either causation leaves it indeterminate whether we refer to one rather than another of these properties, or we must rely on something other than pure causal relations to single out a thing or property of one sort rather than another.

Levin’s solution to the *qua* problem is to supplement the pure causal relations between a concept and its referent with the subject’s dispositions to use (or not use) the concept in further cases. She maintains that:

The best way—perhaps the only physicalistically acceptable way—to determine whether someone’s current “pointing in” denotes what it’s like to see some particular shade of red, or a more course-grained phenomenal property…, or one of a number of phenomenal properties that are instantiated in an experience but impossible to attend to selectively at that time…, is to see what she is disposed to identify as *other* instances of that property.\(^{220}\)

While I disagree with Levin’s suggestion that a subject’s dispositions to use a concept in further cases is perhaps the only physicalistically acceptable way to solve the *qua* problem, I agree with her that a subject’s recognitional propensities can play an important role in grounding the determinacy of reference of many of our concepts, including of our phenomenal concepts.\(^{221}\)

Thus far, we have been considering Levin’s claim that phenomenal concepts are akin to familiar demonstrative-recognitional concepts in that their reference is direct rather than mediated through any descriptive mode of presentation. Now we will turn to her first claim that like familiar demonstrative-recognitional concepts, phenomenal concepts are essentially perspectival, and hence irreducible to nonperspectival concepts.

\(^{220}\) Levin (2007), p. 89.
\(^{221}\) In addition to a subject’s recognitional dispositions, I think that our attunement-guided selective attention puts us in a position to make determinate reference to things and properties.
In saying that phenomenal concepts are “essentially perspectival”, Levin seems to have two features in mind. First of all, the reference of a phenomenal concept is determined in a role-based manner in a demonstrative-permitting situation. That is, a phenomenal concept refers to whatever experiential kind (or instance) the subject has at the time of an inner, reference-fixing demonstration (typically on the basis of the subject’s selective attention). Secondly, the conceptual gap between phenomenal concepts and physical-theoretical concepts, she argues, mirrors the well-known gap between familiar demonstrative concepts and objective concepts of the same thing or property.

To illustrate this second point, let us consider Perry’s (fictional) example of his meeting Fred Dretske for the first time at a party. Perry asks us to suppose that prior to meeting Dretske, he admires Dretske’s work and knows it well. He knows, for example, that Dretske wrote *Knowledge and the Flow of Information*, *Naturalizing the Mind*, *Seeing and Knowing*, etc. However, he has never met Dretske and does not know what he looks like. Next, Perry asks us to imagine that he actually meets Dretske at a party but does not recognize that it is him. He begins to talk with Dretske about *Knowledge and the Flow of Information* and even suggests that Dretske read the book. When the man that Perry is talking to informs him that he wrote *Knowledge and the Flow of Information*, Perry’s doxastic life changes. In addition to becoming embarrassed about his mistake, Perry learns something new. But what exactly does he learn?

He doesn’t learn that Dretske wrote *Knowledge and the Flow of Information* because he already knew that prior to meeting Dretske at the party. What he learns is that this man (the one he is currently perceiving and talking to) wrote *Knowledge and the Flow of Information*. This second bit of knowledge is irreducible to the first bit of

---

222 Perry (2002), Chapter X.
knowledge. Indeed, we can say that there is an “epistemic gap” between the second piece of knowledge and the first, a gap which explains why Perry could know that Dretske wrote *Knowledge and the Flow of Information* and still make his social gaffe.

Of course, if we want to explain just how Perry’s doxastic life changes, we must be careful. We might be tempted to think that in coming to know the second bit of knowledge, Perry learns a new fact about the world. However, if we take “this man” to be a directly referential device (as we should), then Perry already knew the proposition it expresses about the world when he knew that Dretske wrote *Knowledge and the Flow of Information*. Indeed, both bits of knowledge place the same conditions on the world—that such and such an individual wrote *Knowledge and the Flow of Information*. But then how should we understand the change in Perry’s doxastic life?

One might say that what Perry learns is not a fact about the world itself, but rather a fact about the cognitive states he uses to think about the relevant state of the world. However, we must be careful here. Strictly speaking, his knowledge is about Dretske, and not about his own concepts. In learning that Dretske is this man, Perry is making a substantive empirical discovery about the world, and not merely a discovery about the incidental co-reference of two of his concepts. This issue, of course, is precisely what led Frege to give up his early treatment of identity statements in the *Begriffsschrift* in favor of his famous mature account involving the distinction between the ‘sense’ and ‘reference’ of a concept or linguistic expression.223

This is where Perry’s distinction between the “reflexive contents” and “incremental” or “subject matter contents” of a representation can help us to avoid Frege’s difficulties without taking on his problematic notion of ‘sense’. On Perry’s view,

223 Frege (1892/2000).
representations have many cognitively significant features, all of which can be classified in terms of the various conditions they place on the world. Some of these contents are *reflexive contents* in that the conditions they place on the world concern aspects of the representations themselves, given facts about the broader representational systems of which they are a part. Others are *incremental contents* in that they place conditions on the rest of the world. These incremental contents are, at least in the case of directly referential devices, the subject matter of our representations; they are what the representations are *about*.

Returning to the Perry-Dretske case, what his new knowledge is *about* is the same fact he knew before his discovery--the trivial, but necessary, fact he could have expressed all along with “Dretske is Dretske”. But the *reflexive* contents of his newly acquired knowledge are quite different from those of the knowledge he had of one and the same fact previously. His new knowledge involves very different representational vehicles, vehicles which play very different psychological roles and are related to Dretske in very different ways. And in learning that this man is Dretske, these cognitively distinct representations become linked by an internal identity sign where there was none before. Information flows from one to the other. The internal identity sign makes explicit that the representations are *of* the same thing and thus make the same contribution to the subject matter contents of thoughts or utterances in which they occur. This is true despite the fact that they make different contributions to the total cognitive significance of thoughts or utterances in which they occur.

What Perry hopes to show with this example is that there are cases involving familiar demonstratives in which there is an epistemic gap between two beliefs about the
world without a corresponding ontological gap. Indeed, prior to Dretske’s informing Perry that he wrote *Knowledge and the Flow of Information*, it was conceivable for Perry that the individual he thought of as “*this man*” was not the actual author of it. However, given that “*this man*” refers directly to Dretske, it is not metaphysically possible that he is not the actual author of *Knowledge and the Flow of Information*.

Both Perry and Levin agree that there is a similar epistemic gap, without a corresponding ontological gap, between the experiential/physical thought of under phenomenal concepts and the experiential/physical thought of under physical-theoretical concepts. What we have, they argue, are simply two radically different ways of thinking of one and the same (physical) facts that conceptually irreducible to each other. Indeed, to see the parallels between the epistemic gap between the physical and the experiential and the epistemic gap in the Dretske case, it is helpful to consider Perry’s treatment of the Knowledge argument.

In Perry’s discussion of the knowledge argument, Mary has a complete description of the world couched in physical theoretical vocabulary. However, from this, Mary cannot deduce any facts about the experiential couched in terms of phenomenal vocabulary, despite the fact that, by Perry’s lights, the experiential is physical and despite the fact that Perry’s Mary is a card-carrying physicalist. As a physicalist, she believes (rightly) that experiential kinds are identical with physical kinds, but prior to her release she has never actually had experiences of the chromatic color experiences before. After her release, she has chromatic color experiences for the first time and learns something new about herself and others. She learns what it is like for her, and others similarly
situated and endowed, to have the relevant kinds of experiences of chromatic colors. The
question is: Is what she learns a new, non-physical fact?

Perry thinks not. In characterizing the changes in Mary’s doxastic life, he asks us to begin by imagining the following:

Mary emerges from the black and white room, sees a fire hydrant, and has her first colour experience, call it $E$, of the type quale$_{\text{RED}}$. She thinks: ‘This$_i$ experience is the type I have when I see, in these conditions, the colour of that fire hydrant’. She has a certain relation to the experience: she has it. A less inquisitive person might have left it at that, but she also attends to the experience. $E$ is the referent of her thought ‘that$_i$ experience’ because of the relations it has to her: it is the one she is having and attending to.$^{224}$

On Perry’s account, what Mary has acquired is a token phenomenal concept of her particular experience of red. Like the perceptual demonstrative of Dretske, it refers directly to that which occupies a certain role relative to its user. That is, it refers directly to the dated, unrepeatable experience of red that Mary has and is attending to.

Next, Perry asks us to imagine that Mary forms a type phenomenal concept of the kind of experience, i.e. of red, of which Mary’s particular experience is an instance. He says:

Mary forms a concept of the type of colour experience $E$ exemplifies. She notes that $E$ is similar to the colour experience she has of the fire engine parked nearby, and not similar to the colour experience she is having of the grass next to the fire hydrant. She can introduce a term, ‘quale$_?\$’ and ask:

Is quale$_?\$ = quale$_{\text{RED}}$?\(^{225}\)

Mary’s new type phenomenal concept allows her to think various new kinds of thoughts that she couldn’t before her release. For example, she can anticipate circumstances under which she would have that kind of experience again. She can also wonder whether others have the same sort of experiences in similar circumstances. But most importantly for our

---


purposes, as Perry notes, she can wonder about the identity and ontological status of quale. As it turns out, Mary is a normal perceiver in standard viewing conditions and so quale is quale\textsubscript{RED}. But notice that Mary is not in a position to deduce this identity a priori. She might, after all, wonder whether the fire hydrant see is looking at is painted the standard color or not. If it is, then she can conclude that quale is quale\textsubscript{RED}. If not, then quale is not quale\textsubscript{RED} but some other kind of quale. What is crucial to note is that there is a gap here in Mary’s knowledge; even though she knows that “quale is quale?” and that “quale\textsubscript{RED} is quale\textsubscript{RED}” she does not know that “quale is quale\textsubscript{RED}” even though quale is quale\textsubscript{RED}. But if quale and quale\textsubscript{RED} are both directly referring devices, then all of these bits of knowledge have the same subject matter content; they place the same conditions of satisfaction on the world. What Perry hopes we’ll see is that we shouldn’t get at the gap in Mary’s knowledge here by focusing on subject matter content. What we need to recognize is that, just as in the Dretske case, the change in Mary’s doxastic life when she learns that quale is quale\textsubscript{RED} is at the level of reflexive content. The change in Mary is that she learns that the two concepts are of one and the same kind of experience.

Ultimately, Perry’s diagnosis of the epistemic gap in this case is that Mary is thinking of one and the same kind of experience in two very different ways. He contends:

Mary is thinking about the same type of experience in two different ways when she uses the two terms in thought or language. The referential relations are quite different; she is related to the quale in two quite different ways. On the one hand, it is the quale that two of her current experiences exemplify. On the other, it is the quale that her textbooks referred to, and identified as the type of

\footnote{In point of fact, Mary could evidently ask these sorts of questions with her token phenomenal concept as well.}
experience normal people have in favourable light when they see red objects. Her conceptions of the two are different.\textsuperscript{227}

Furthermore, Perry argues, the same lesson applies to the case of the identity of quale\textsuperscript{?} (and quale\textsuperscript{RED}) with a kind of neurobiological state of the brain thought of under a physical-theoretical concept, say of brain state “B\textsubscript{52}”. According to Perry, “quale\textsuperscript{?}” and “B\textsubscript{52}” are co-referring concepts. Nevertheless, there are epistemic gaps between them stemming from the radically different ways their reference is fixed to one and the same phenomenon. Indeed, he remarks:

> What magic would drive a mental identity sign between two such different concepts as these [i.e. the demonstrative-recognitional phenomenal concept and the physical-theoretical concept], even for someone as brilliant as we suppose Mary to be? Neither concept is defined in terms of the other; neither is introduced in terms of the other; neither makes reference to the other; they have no common parts.\textsuperscript{228}

What’s more, he continues:

> At this point the knowledge argument says: \textit{but then, what does \textit{Mary learn when she has the experience of red for the first time}?} But the answer is clear. It’s just the difference between having two concepts that are, in fact, of the same thing, and two concepts that are required by that internal identity sign to be of the same thing. She learns that the type of experience she is having, and so the type of experience her phenomenal concept is of, is the type of experience that her other, brain-science-based concept is of; namely, the [red] quale. The contents of her doxastic states change, in that the truth of her total doxastic state, abstracted from the referential relations of her concepts, requires that the two concepts are concepts of the same type of experience…That is how her beliefs change, and how her knowledge changes.\textsuperscript{229}

Perry’s argument here, I believe, is quite powerful. If phenomenal concepts are directly referential devices that allow one to think of physical facts in a new, irreducibly perspectival way, then the physicalist has a powerful tool for resisting at least one of the

\textsuperscript{227} Perry (2009), p. 235.
\textsuperscript{228} Perry (2009), p. 237.
\textsuperscript{229} Perry (2009), pp. 237-8.
dualist arguments. Therefore, we have good reason to accept much of what the
demonstrative-recognitional model says about the special features of phenomenal
concepts.\textsuperscript{230}

Unfortunately, I do not think that the demonstrative-recognitional model is wholly
satisfactory. On the demonstrative-recognitional model, phenomenal concepts are
irreducibly perspectival, directly referential concepts whose reference is fixed in a role-
based manner from a demonstrative-permitting situation. These features explain why
there is an epistemic or conceptual gap between phenomenal concepts and physical-
theoretical concepts despite the fact that they both co-refer to experiential/physical events
and properties. But the problem is that it does not have enough resources to capture what
is \textit{distinctive} about our epistemic situation with respect to consciousness: that our
phenomenal concepts provide us with conceptions of the subjective, qualitative
characters of our experiences.

We might usefully pose the problem as follows: What distinguishes phenomenal
concepts from more familiar kinds of demonstrative concepts, such as perceptual
demonstratives? One might initially be inclined to say that the distinctive feature of
phenomenal concepts is simple: they designate \textit{phenomenal} properties. However, this
cannot be the right response. For if the experiential is identical with some part of the
physical domain, as proponents of the strategy believe, then we can in principle perceive
that experiential/physical state or property and designate it with a familiar perceptual
demonstrative. We will not thereby, however, stand in the same epistemic situation with
the respect to the experiential as we do when we deploy phenomenal concepts to think

\textsuperscript{230} In fact, my own acquaintance-based account of phenomenal concepts accepts much of what the
demonstrative-recognitional account says about them.
about it. Nor will we be confronted with the kind of epistemic gap that we are looking to explain between the experiential and the physical. So, phenomenal concepts are not simply demonstrative-like devices that pick out phenomenal properties.

At this point, one might point to the fact that the reference of phenomenal concepts is fixed by introspective attention to physical/phenomenal properties rather than by perceptual attention to them. In fact, both Perry and Levin explicitly note quite often that the reference of phenomenal concepts is fixed by introspection or by an ‘inner demonstration’. But while they are surely correct in noting that the reference of phenomenal concepts is grounded in introspective awareness, it is not simply enough to point out a difference in the modality through which a subject is aware of the experience or phenomenal property. We lose grip on something important about the explanatory gap between the experiential and the physical if we treat introspective awareness as just one among many ways we might be aware of phenomenal properties.

Intuitively, the epistemic gap between the experiential and the physical is deeper than more familiar cross-modality gaps of the sort Molyneaux pointed out long ago to Locke.\textsuperscript{231} It is intuitively not as shallow as the gap between seeing a sphere and feeling one, for instance. Of course, one might happily “bite the bullet” here. But if one takes the epistemic gap between the experiential and the physical to be deeper than Molyneaux’s problem, our account of phenomenal concepts must say more about why introspective awareness yields an epistemic gap that is different in kind from those raised by cross-modal perceptual awareness of things.

\textsuperscript{231} Locke (1690), Book 2, Chapter 9.
§5 THE CONSTITUTIONAL MODEL

Recently, many proponents of the phenomenal concepts strategy have tried to account for the epistemic gap between the experiential and the physical by pointing to a different purported feature of phenomenal concepts: that they are partly constituted by instances of phenomenal properties. Indeed, in presenting the constitutional model, Katalin Balog contends that:

Phenomenal concepts are partly constituted by the phenomenal experiences they refer to. On this view, a current phenomenal experience is part of the token concept currently applied to it, and the experience – at least partly – determines that the concept refers to the experience it contains.\(^{232}\)

In a similar vein, David Papineau asserts that:

My phenomenal concepts involve a frame, which I have represented as ‘the experience: - - - ’; and, when this frame is filled by an experience, the whole then refers to that experience.\(^{233}\)

On this constitutional model, phenomenal concepts are analogous to devices for linguistic quotation in that what they refer to is exemplified in the representational vehicle itself. This, it is argued, explains why there evidently is such an intimate relation between our experiences and our thoughts about them. Moreover, our mistaken intuition of the distinctness of the experiential and the physical is explained by the fact that physical-theoretical concepts, on the other hand, do not similarly exemplify their referents. In this section, I examine Balog’s and Papineau’s recent versions of the constitutional model of phenomenal concepts. Ultimately, I argue that the constitutional model is unsatisfactory. It simply makes the relation between the experiential and our thought about it too intimate, and, moreover, it is unclear that it has the resources to explain this alleged epistemic intimacy.

---

\(^{232}\) Balog (Forthcoming), p. 18.

\(^{233}\) Papineau (2002), p. 117. Note that Papineau has since given up this view of phenomenal concepts.
To begin with, there are a number of respects on which the constitutional model agrees with the demonstrative-recognitional model about the nature of phenomenal concepts. First of all, like on the demonstrative-recognitional model, the constitutional model holds that phenomenal concepts are directly referential devices that pick out physical/phenomenal properties that bear the right sort of contextual relations to them. Secondly, the constitutional model maintains that one can only acquire or possess genuine phenomenal concepts if one has actually *had* the relevant kind of experience (or at least one suitably similar). However, thirdly, the constitutional model allows that we can still refer to experiences without having had them using a public concept for experiences, or what Balog calls an “indirect phenomenal concept”. And fourthly, the constitutional model holds that part of what fixes the determinacy of reference of our phenomenal concepts are our dispositions to apply them to further instances of experiential kinds. However, the constitutional model has very different explanations than the demonstrative-recognitional model for why phenomenal concepts have the first two of these four features.

One of the key claims of the constitutional model is that the direct reference of phenomenal concepts to phenomenal properties is achieved via mechanisms analogous to linguistic quotation. In other words, phenomenal concepts will pick out experiential kinds in virtue of exemplifying instances of the relevant experiential kinds in the very concepts used to think about them. We might also put this point by saying that phenomenal concepts pick out phenomenal properties that play the right sort of contextual role.

---

234 Balog (Forthcoming), p. 31.
235 Balog (Forthcoming), pp. 30-1.
relative to them—that of being the phenomenal properties that *constitute* the phenomenal concepts (or ones that appropriately resemble them).

In any case, the idea is that we can perform a cognitive operation whose conceptual role is similar to that played by linguistic quotation. Indeed, Balog argues:

The idea of an item partly constituting a representation that refers to that item is reminiscent of how linguistic quotation works. The referent of “__” is exemplified by whatever fills in the blank. In a quotation expression, a token of the referent is literally a constituent of the expression that refers to a type which it exemplifies and that expression has its reference (at least partly) in virtue of the properties of its constituent.\(^{236}\)

What’s more, she continues:

My proposal is that there is a concept forming mechanism that operates on an experience and turns it into a phenomenal concept that refers to a type of experience where the type is a qualitative property (a qualia) of the experience. Further, the operation, like linguistic quotation, can be explained in terms of its conceptual role.\(^{237}\)

Papineau similarly compares phenomenal concepts to quotational devices. In presenting his view, he remarks:

The model I wish to pursue, then, proposes that phenomenal concepts are compound term, formed by entering some state of perceptual classification or recreation into the frame provided by a general experience operator “the experience: - - -”…Such terms will have a certain sort of self-referential structure. Very roughly speaking, we refer to a certain experience by producing an example of it.\(^{238}\)

And he says:

It may be helpful to compare the model I am defending to the use of quotation marks. The referring term incorporates the things referred to, and thereby forms a compound which refers to that thing.\(^{239}\)

\(^{236}\) Balog (Forthcoming), p. 27.
\(^{237}\) Balog (Forthcoming), p. 27.
\(^{239}\) Papineau (2002), p. 117.
According to Balog, the conceptual role of a quotational expression consists in a disposition to accept instances of the following disquotational schema (where x stands for any word in the relevant language):

\[
\begin{align*}
L1 & \quad \text{“x” refers to } x \\
L2 & \quad \text{““x”” refers to “x”, etc.} \quad 240
\end{align*}
\]

Moreover, she argues, we can understand the conceptual role of the “mental quotation” involved in phenomenal concepts using a similar schema. That is, the conceptual role of a mental quotational device consists in a disposition to accept instances of the following schema (where “experience x” stands for an experience rather than a concept of an experience and “*” stands for the mental quotation operator):

\[
\begin{align*}
M1 & \quad \text{*experience x* refers to experience x} \\
M2 & \quad \text{**experience x** refers to *experience x*, etc.} \quad 241
\end{align*}
\]

Balog’s basic idea here is that the experience itself is operated on and thereby gets into our thoughts about it. Moreover, it is constitutive of being a phenomenal concept, on her view, that it has this structure since phenomenal concepts are individuated in part by having the conceptual role articulated by the mental quotational schema.

In addition, Balog and Papineau argue, the special structure of phenomenal concepts also explains why one needs to have particular experiences before one can acquire or possess phenomenal concepts of them. The reason for this is straightforward. Phenomenal concepts are by hypothesis partly constituted by phenomenal properties and thus are metaphysically dependent on instances of them. Put somewhat differently, one cannot operate on, and thereby mentally quote, an experience unless one has the

\[240\] Balog (Forthcoming), p. 27.
\[241\] Balog (Forthcoming), p. 28.
experience on which to operate. Therefore, no concept can satisfy the mental quotational schema unless it bears the proper relation to an instance of some experiential kind. This explains why one must have an experience in order to form a genuine phenomenal concept of it.

One motivation for adopting the constitutional model of phenomenal concepts is to make good on Loar’s (1990) well-known proposal that phenomenal properties serve as their ‘own modes of presentation’ without endorsing an indirect, Fregean view of reference. In endorsing what he sees as right about Loar’s proposal, Papineau remarks:

At one level, the idea that phenomenal properties can provide their ‘own modes of presentation’ may simply mean that they do not have to be picked out via some other contingently connected property they possess. There is only one property in play when a phenomenal concept refers to a phenomenal property: namely, the phenomenal property itself. No further property mediates between referring concept and referent.242

But, he continues:

It would seem badly to misrepresent this idea to say that phenomenal properties provide their own modes of presentation. This suggests a picture whereby the mind somehow already has the power to think about some phenomenal property, Ø, and then uses this ability to form a mode of presenting that property (‘the property which is property Ø’, perhaps). But this makes little sense. If we already have the ability to think about the phenomenal property Ø, we don’t need to construct some further mode of presentation to enable us to think about it.243

Here is where the notion of constitution is supposed to help. The basic idea here is that if phenomenal concepts refer to phenomenal properties exemplified in those very concepts, we have some idea of how it is that those phenomenal properties play a reference-determining role that is even more direct than other sorts of direct designation (and we needn’t do so in a way that ushers in Fregeanism). As Balog put it:

Direct phenomenal concepts pick out the referent in virtue of their being partly constituted by a token of their reference. In this they are unique among concepts. On this account, there is an intimate relation between a phenomenal concept and its referent; more intimate than any causal or tracking relation. It is also a way of cashing out the idea that experience serves as its own mode of presentation. The experience, so to speak, presents itself.\textsuperscript{244}

The reason that Balog and Papineau think that the reference of phenomenal concepts to phenomenal properties is more intimate than standard cases of direct reference is that they find the possibility of reference failure deeply problematic in the phenomenal case. In standard cases of direct reference, concepts and their referents are distinct existences related by some sort of causal, historical, informational, etc., relation. This evidently opens the possibility that one can token the concept even in the absence of a corresponding referent. But, Balog and Papineau contend, it is incoherent to suppose that one can deploy a genuine phenomenal concept in the absence of a corresponding referent. As Balog puts it:

\textit{On a “tracking” view, [a phenomenal concept] P, or rather, a concepts just like P could be tokened by someone in the complete absence of pain [for example]. A person like this would be a partial conceptual zombie; a \textit{conceptual duplicate} of a normal human, who, however, fails to have all the qualia the normal human has. But it seems to me that such a zombie is really impossible. Anybody who tokens a first personish phenomenal concept of pain purporting to refer to a current state is really in pain. The trouble with [the distinct existences account] is that it opens up the appearance/reality distinction for direct phenomenal concepts.}\textsuperscript{245}

Basically, Balog is endorsing here Kripke’s (1983) intuition that whatever strikes us as such and such experience \textit{is} such and such experience, and vice versa. In being phenomenally aware of experiences, the idea runs, we are aware of the qualitative

\textsuperscript{244} Balog (Forthcoming), p. 19.
\textsuperscript{245} Balog (Forthcoming), p. 18.
essence of experiences in a way that we cannot be mistaken about them. Indeed, Balog thinks that certain kinds of phenomenal judgments are infallible. She contends that:

A phenomenal concept may refer to a particular type of visual experience, say the experience typically caused by seeing red objects in ordinary light, etc. – call this type of experience reddish – by being constituted in part by a particular token of that type of experience. Then if I form the judgment ‘I HAVE R’ where R is a direct phenomenal concept of reddish, my judgment cannot fail to be true. Similarly, if E is a direct phenomenal concepts referring to a token reddish experience and R is also a direct phenomenal concept based on the same reddish experience E is based on, the judgment ‘E IS R’ cannot fail to be true.\(^{246}\)

Furthermore, Balog argues, in thinking about experiences using genuine phenomenal concepts, we can be assured that we are having the relevant experience. This explains both why we can be assured that we are not zombies and why the cognitive lives of zombies are so alien to us. For zombies, if possible, would presumably think they have experiences even though they lack experiential lives altogether. But this, Balog and many others have argued, just shows that there is something deeply problematic about the coherence of zombies.\(^ {247}\)

Another motivation for adopting the constitutional model of phenomenal concepts is that, unlike the demonstrative-recognitional model, it provides a possible explanation of the explanatory gap. On the constitutional model, there are two features of phenomenal concepts that are responsible for there being an explanatory gap between the experiential and the physical. First, phenomenal concepts designate phenomenal properties directly through mental quotation and, unlike physical-theoretical concepts, therefore do not in

\(^{246}\) Balog (Forthcoming), p. 21.

\(^{247}\) Chalmers thinks that the difference between us and zombies are that we have, and are acquainted with, evidence that the zombie lacks. Of course, the zombie will think that it has the very same evidence. Such puzzles about the cognitive lives of zombies have led several recent physicalists to contend that we are zombies. See, for example, Beisecker (2010).
any way describe the causal roles of their referents. Therefore, the argument goes, it is not possible to use our standard explanatory models of reducing natural phenomena by describing their causal roles and finding the lower level physical phenomena that realize those roles. Second, phenomenal concepts exemplify the physical/phenomenal properties they designate whereas physical-theoretical concepts do not. This radical difference in the way that we think about one and the same physical/phenomenal properties, Papineau argues, explains why we have such strong intuitions of distinctness between the physical and the experiential. It also explains why we feel so strongly that a complete physical description of the world, one using only physical-theoretical concepts, leaves something out. But if the constitutional model is correct, then our intuitions of distinctness are the result of the radically different ways we think about physical reality rather than any ontological gap in reality itself.

Unfortunately, while I agree with Balog and Papineau that there are many attractive features of the constitutional model, I do not think that it ultimately provides a plausible account of phenomenal concepts. In the rest of this section, I present a series of worries one might have about the constitutional model. While some of these worries are not decisive, I think that collectively they suggest that the constitutional model makes the relation between the experiential and our thought about it too intimate. Furthermore, there is a deep question about whether physicalist versions of the constitutional model have the explanatory resources to capture the robust epistemic and referential features that it

\[\text{Papineau (2002) puts the point as follows on page 143:} \]

The reason we cannot give any materialist ‘explanation’ of why the brain yields phenomenal properties is not that these properties are non-material, where those studied in other areas of science are material. Rather, it is that phenomenal concepts are not associated with descriptions of causal roles in the same way as pre-theoretical terms in other areas of science. This means that it is possible to understand identity claims in other areas of science as involving descriptions, and so open to explanation by materialist reductions in a way that is not open in the mind-brain case.
attributes to phenomenal concepts. Together, I think these worries give us good grounds to look for an alternative model of phenomenal concepts to guard physicalism against dualist arguments.

With that said, the most immediate and straightforward worry one might have about the constitutional model is that it runs counter to our pre-theoretic intuitions that experiences and our thoughts about them are distinct existences. Indeed, I take it that the commonsense picture is that pains, for example, are one thing, and thoughts about pains are another. What’s more, while we sometimes have pains when we think about them, we do not always have pain whenever we think about it. However, the constitutional model, in claiming that phenomenal properties partly constitute our phenomenal concepts, seems to predict that whenever we think about experiences using phenomenal concepts, we invariably have those very experiences. But this seems patently false. Intuitively, there are many imaginative, classificatory, and anticipatory uses of phenomenal concepts that are typically unaccompanied by the experiences they are about.

In response to this worry, proponents of the constitutional model draw a distinction between genuine phenomenal concepts and what Balog calls “indirect phenomenal concepts” or Papineau calls “phenomenally derived concepts”. These latter concepts are concepts of phenomenal properties which are derivative of genuine phenomenal concepts but are not themselves constituted by the phenomenal properties they designate. As such, they provide us with referential devices for thinking about experiences in the absence of them. Indeed, proponents of the constitutional model argue, indirect phenomenal concepts provide us with the cognitive resources needed for our
many imaginative, classificatory, and anticipatory thoughts about experiences that do not involve actually having them. As Balog puts it:

Indirect phenomenal concepts are applied to non-occurrent (e.g., past or future) experiences of oneself or to the experiences of other people. These concepts must be distinct from direct phenomenal concepts. Developing an account of these concepts is an absolutely essential task for understanding consciousness. My view is that these concepts are individuated in part by conceptual roles that link them to direct phenomenal concepts.249

However, though it is not a decisive objection, it is worth pointing out that this bifurcation of concepts into direct phenomenal concepts and indirect phenomenal concepts is at least somewhat counterintuitive. Typically, we take it that we deploy one and the same concepts when thinking about our own occurrent experiences, our non-occurrent experiences, and the experiences of others. In addition, if proponents of the constitutional model are correct, then the explanatory power of the appeal to genuine phenomenal concepts is somewhat diminished by the fact that most of our thoughts about experiences are arguably not about occurrent experiences. Again, this is not a decisive objection, but it is a substantial consideration against the constitutional model.

Another common worry about the constitutional model of phenomenal concepts is that there are some kinds of phenomenal judgments that become problematic on it. In particular, it becomes impossible for an individual to think truly “I am not now having such and such experience” where he or she uses a phenomenal concept to think about the relevant experience.250 This is because in thinking of an experience using a phenomenal concept, one tokens that very experience in one’s thoughts. Hence, in thinking that one is not occurrently having such and such experience, one ensures that one is. But it is an

249 Balog (Forthcoming), p. 31.
250 Papineau (2007).
extremely counterintuitive result that we cannot think that we are not occurrently having such and such experiences.

As in the case of the previous worry, proponents of the constitutional model appeal to indirect phenomenal concepts to circumvent the problem. They claim that when one thinks “I am not now having such and such experience”, one uses an indirect phenomenal concept, one that does not exemplify what it designates, rather than a direct phenomenal concept. But while this provides a way of avoiding the objection, it is an unduly restrictive and counterintuitive move. After all, our pre-theoretical intuitions are, once again, that we use one and the same concept whether we think that we are having such and such an experience or that we are not. Thus, we again have reason to suspect the constitutional model of phenomenal concepts.

So far, the worries about the constitutional model are the result of its conflict with certain pre-theoretical intuitions. However, pre-theoretical intuitions only hold so much weight. If all that were objectionable about the constitutional model were that it had a few counterintuitive results, its proponents would have good reason to feel that it is the pre-theoretical intuitions, rather than the theory, that must go. For, as we have seen, the constitutional model has several nice features for avoiding the dualist arguments. Nevertheless, there are deeper problems for the constitutional model which suggest that it cannot ultimately be correct.

The first worry is that it is hard to see how the notion of ‘constitution’ can do the heavy explanatory lifting for the proponent of the constitutional model in accounting for the epistemic and referential intimacy they see between phenomenal concepts and the phenomenal properties they designate. For if the constitutional model is to provide the
physicalist with a defense against the dualist arguments, then it must respect what Joseph Levine calls “the materialist constraint”. That is, it must only use explanatory resources that are available to a physicalist in characterizing the relevant notion of ‘constitution’.

The problem is that physical constitution is not an epistemic or semantic relation. This explains why Balog remarks that:

On [the constitutional] view, a current phenomenal experience is part of the token concept currently applied to it, and the experience – at least partly – determines that the concept refers to the experience it contains. Of course, by “part” I do not mean “spatial part” but rather part in the sense that it is metaphysically impossible to token the concept without tokening its referent.

The challenge for the proponent of the constitutional model is to characterize this necessary metaphysical dependence in a way that is acceptable by the lights of physicalism. Their strategy at this point, of course, is to appeal to an alleged analogy between direct phenomenal thought and linguistic quotation. Phenomenal properties, the argument runs, are part of the phenomenal thoughts about them in the same way that expressions are part of the linguistic quotations that designate them. This, of course, is not a mere matter of spatial constitution, a fact which is exploited by the proponent of the constitutional model for explaining mental quotation.

However, there is a deep shortcoming of the analogy between linguistic quotation and our thoughts about phenomenal experience: linguistic quotation is a semantic phenomenon, not an epistemic one. Hence, while the constitutional model can perhaps account for how phenomenal concepts directly designate the phenomenal properties they are about, it does nothing to dispel the mystery of how we are intimately aware of phenomenal properties when thinking about them using phenomenal concepts. Moreover,

---

251 Levine (2007).
252 Balog (Forthcoming), pp. 18-9.
it is not clear how the constitutional model can cash out the notion of “constitution” in a way that can ground an appropriate epistemic relation. There is a simple reason for this. Constitution just doesn’t seem to be an epistemic-enabling relation. Of course, it is always possible that some future proponent of the constitutional model will be able to make intelligible how constitution lends itself to epistemic intimacy. Still, I think that the constitutional model has the heavy burden of explaining how any kind of constitution can ground epistemic awareness, let alone intimacy—a burden that it has as of yet failed to take up. What is clear is that analogies to linguistic quotation are inadequate to the task.

The second, and perhaps even more decisive, worry about the constitutional model is that the phenomenal continua cases that I discussed in chapter one of this dissertation show that phenomenal concepts and the phenomenal properties they designate are indeed distinct existences. To recap the discussion in chapter one, phenomenal continua are cases in which one quality transitions into another quality such that no difference can be discriminated between two distinct points A and B in the transition, nor between two distinct points B and C, and yet some difference can be discriminated between points A and C. According to Russell, such phenomenal continua cases can be constructed for all cases in which qualities admit of gradual change. Phenomenal continua cases fall into two basic varieties. First, there are diachronic phenomenal continua cases in which a quality gradually transitions to another quality over some period of time such that we cannot distinguish A at t_1 from B at t_2 or B at t_2 from C at t_3 but we can distinguish A at t_1 from C at t_3. Recall that in illustrating such diachronic cases, Russell says:

Suppose, for example, a person with his eyes shut is holding a weight in his hand, and someone noiselessly adds a small extra weight. If the extra weight is
small enough, no difference will be perceived in the sensation. After a time, another small extra weight may be added, and still no change would be quite easily perceptible; but if both weights had been added at once, it may be that the change would be quite easily perceptible.\(^{253}\)

Second, there are *synchronous* phenomenal continua cases in which a quality gradually transitions into another such that we cannot distinguish A and B or B and C pair-wise but we can distinguish A and C pair-wise. Recall that in describing such synchronous cases, Russell remarks:

Or, again, take shades of colour. It would be easy to find three stuffs of such closely similar shades that no difference could be perceived between the first and second, nor yet between the second and third, while yet the first and third would be distinguishable. In such a case, the second shade cannot be the same as the first, or it would be distinguishable from the third; nor the same as the third, or it would be distinguishable from the first. It must, therefore, though indistinguishable from both, be really intermediate between them.\(^{254}\)

As we’ve seen, what Russell sees as important in each of these cases is that our capacities for introspective discrimination are outrun by changes in the items or qualities that we are experiencing such that we might erroneously conclude that there are no differences or boundaries between them. This in turn shows that *even in the case of our own “sense-data”* (i.e. the subjective, qualitative characters of our own conscious experiences), acquaintance does not secure us against the possibility of misidentifying the objects of our thought and talk. What’s more, they show that when we are acquainted with a thing we do not thereby have revelatory knowledge of its essential nature.

Recall that for Russell, knowledge of the numerical identity or diversity of A and B is *knowledge of truths* about A and B, which he holds to be logically distinct

from our more fundamental acquaintance with them. We cannot know, simply on the basis of being acquainted with A and B, whether A and B are identical or distinct—we cannot know whether we are acquainted with two distinct things or one thing twice over. Russell says:

To know that two shades of colour are different is knowledge about them; hence acquaintance with the two shades does not in any way necessitate the knowledge that they are different.255

Moreover, he contends, even if we designate A and B with genuinely singular, directly referential concepts and consider whether A is B, we might not know what the right answer is. This point is extremely important for Russell, and he leans heavily on it when arguing that we cannot, from our acquaintance with experience alone, determine whether or not it is constituted by discrete “sense data”.256

Setting aside Russell’s own philosophical concerns, phenomenal continua cases also highlight a deep problem with the constitutional model: that it rules out errors we have good reason to think that we make. As I just noted, if phenomenal continua cases are possible (and they manifestly are), then our judgments about our occurrent experiences admit of the possibility of errors of misidentification. But the alleged tight metaphysical relation (i.e. constitution) between phenomenal properties and our thoughts about them seems to rule out this kind of error, a feature which we’ve seen Balog cites as a virtue of the model. Indeed, errors of misidentification evidently require that our thoughts and their corresponding referents are “distinct existences”; presumably, we must be able to apply (or misapply) its constituent concepts to different things. But then it cannot be the case that the relevant concepts

---

256 See Chapter 1 of this dissertation.
are partly *constituted* by referents which are exemplified whenever they are tokened. So unless the constitutional model can find some other way of accounting for phenomenal continua cases (and other instances of introspective error), it looks to be doomed.

§6 CONCLUSION
In this chapter, I have tried to set the stage for a discussion of my own take on phenomenal concepts in the following chapter. In section §2, I presented a rough characterization of physicalism. I also explained how Chalmers’ zombie argument, Jackson’s knowledge argument, Kripke’s modal argument, Levine’s explanatory gap argument, and Chalmers’ two-dimensional argument aim to show that physicalism cannot be true because there is an epistemic gap between the physical and the experiential that is, they claim, indicative of an underlying ontological gap. In section §3, I presented the phenomenal concepts strategy and explained how it aims to block the dualist conclusion that an ontological gap is responsible for the epistemic gap between the experiential and the physical. In section §4, however, I argued that the demonstrative-recognitional model of phenomenal concepts does not suffice to explain completely our epistemic situation with respect to the experiential. And in section §5, I argued that the constitutional model of phenomenal concepts makes the experiential and our thought about it too intimate. In the next chapter, I will provide an account of phenomenal concepts that is grounded on the special way in which we are acquainted with our own experiences. What’s more, I will show how such an acquaintance-based model of phenomenal concepts, unlike its competitor models, can provide a satisfactory answer to the dualist arguments against physicalism.
“When it comes to the properties of our immediate experience, we stand in a kind of epistemic relation to them that is more intimate, more substantive, than the kind of relation that obtains between our minds and other items. The properties of experience are, to use my other phrase, cognitively present to us.” – Joseph Levine, “Phenomenal Concepts and the Materialist Constraint”\(^{257}\)

“Acquaintance is a relation that makes possible the formation of pure phenomenal concepts, and…pure phenomenal concepts embody a certain sort of lucid understanding of phenomenal properties. So acquaintance is a relation that makes this sort of lucid understanding possible.” – David Chalmers, “The Content and Epistemology of Phenomenal Belief”\(^{258}\)

§1 INTRODUCTION

The previous chapter began with the following guiding question: “Exactly what kind of knowledge of the experiential does our acquaintance with it provide us with?” Does it, we might continue, provide us with revelatory knowledge of the essence of the experiential such that we can a priori conclude that it is not physical? Or is the nature of our acquaintance with the experiential such that we cannot draw such a priori conclusions about its ontological status? In the previous chapter, I argued that the assumption that acquaintance provides us with revelatory knowledge of the essence of the experiential plays a vital role in the leading dualist arguments against physicalism. For, I suggested, it is this purported substantive knowledge of the nature of the experiential (and the

---


\(^{258}\) Chalmers (2003), p. 29.
physical) that pushes us to conclude from the apparent epistemic or conceptual gap between the physical and the experiential that they are ontologically distinct.

I also presented one of the most promising physicalist strategies—the so-called “phenomenal concepts strategy”—for accepting the existence of an epistemic or conceptual gap between the physical and the experiential while denying that it is symptomatic of an underlying ontological gap. According to the phenomenal concepts strategy, the epistemic gap is the result of our having two radically different, cognitively isolated kinds of concepts, phenomenal concepts and physical-theoretical concepts, for thinking about one and the same physical states of the world. Moreover, this strategy holds that the special nature of phenomenal concepts in particular explains why the epistemic gap between the experiential and the physical so strikingly suggests an underlying ontological gap.

Unfortunately, as I argued in the previous chapter, neither of the two leading accounts of phenomenal concepts is satisfactory. While the demonstrative recognitional model correctly characterizes phenomenal concepts as irreducibly perspectival, directly referential concepts whose reference is fixed in a role-based manner from a demonstrative-permitting situation, it does not have the resources to explain the distinctive character of the epistemic gap between the experiential and the physical. On the other hand, while the constitutional model allegedly has the resources to explain the distinctive character of the epistemic gap, it makes experiences and our thoughts about them too intimate. Indeed, in claiming that phenomenal concepts are partly constituted by the phenomenal/physical properties that they designate, the constitutional model predicts infallibility in our phenomenal judgments where phenomenal continua cases show there
to be fallibility. What’s more, it is doubtful that there is a physically acceptable notion of ‘constitution’ that can explain the epistemic intimacy between experiences and our thoughts about them. Constitution just doesn’t seem to be an epistemically-enabling relation.

In this chapter, I advance my own acquaintance-based account of phenomenal concepts. I argue that phenomenal concepts are concepts of physical/phenomenal properties with a distinctive cognitive role; their reference is fixed by an inner demonstration to states to which we stand in special causal-informational relations (i.e. acquaintance relations), where these special relations are tied to special epistemic methods and capacities for picking up information about the states to which we are related and that provide us with distinctive conceptions of them.

I also show the advantages that this acquaintance-based account of phenomenal concepts has over its competitor views. Unlike the demonstrative-recognitional model, my acquaintance-based model has the resources to say what is distinctive about the epistemic gap between the experiential and the physical. And unlike the constitutional model, my acquaintance-based model has the resources to explain the special epistemic relation between experiences and our thoughts about them without committing us to the infallibility of our phenomenal judgments. But most importantly, my acquaintance-based model undercuts the leading dualist arguments in a more direct way by challenging the crucial dualist assumption that acquaintance provides us with revelatory knowledge of the essence of the experiential.

I proceed as follows. In §2, I develop my acquaintance-based account of phenomenal concepts by drawing on my claim in chapter two that we should understand
Russell’s notion of acquaintance in terms of the right sort of “original” information relations between phenomenally conscious signals and states of the world that phenomenally conscious signals are self-presenting. Then, in §3-7, I show, in succession, how my acquaintance-based account of phenomenal concepts deals with Chalmers’ zombie argument, Jackson’s knowledge argument, Kripke’s modal argument, Levine’s explanatory gap argument, and Chalmers’ two-dimensional semantic argument against physicalism. By showing how an acquaintance-based account of phenomenal concepts can defuse the leading dualist arguments, I aim to show how an antecedent physicalist can grant that there is an epistemic gap between the physical and the experiential without being forced to conclude that there is a corresponding ontological gap.

§2 THE ACQUAINTANCE MODEL OF PHENOMENAL CONCEPTS

In recent years, an increasing number of physicalists have seen the notion of “knowledge by acquaintance” as playing an important role in the defense of physicalism, particularly against Jackson’s knowledge argument. For example, Churchland (1985, 1989), Conee (1994) and Tye (2009) have each proposed that the knowledge Mary intuitively acquires when she sees red for the first time consists entirely in her newfound acquaintance with the experience of red, which is to be distinguished from (and is perhaps more fundamental than) any kind of discursive, propositional “knowledge by description” of the experience of red, or from any kind of know-how. If this view were correct, then it would evidently explain how Mary could know all of the physical facts in the black and

---

259 For the distinction between “original” and “derived” informational content, see Searle (1992), Perry (2002), or Chapter 2 of this dissertation.
white room and yet learn something new upon experiencing red for the first time without thereby learning any non-physical facts. The reason for this is simple. Knowledge by acquaintance is by hypothesis entirely distinct from knowledge of any kind of fact. So, if all Mary acquires is knowledge by acquaintance, then she doesn’t learn any facts at all. Consequently, there is no threat of having to admit non-physical facts into our ontology.

One thing to notice about this use of the notion of ‘knowledge by acquaintance’ is that it makes no appeal to, and provides no account of, the special features of the phenomenal concepts we use to think about the experiential. In fact, Tye (2009) has recently touted the view as a distinct alternative to the phenomenal concepts strategy for defending physicalism. However, I think it is a mistake on the part of physicalist proponents of acquaintance to ignore the role that it can provide in giving an account of phenomenal concepts. One reason for this is that, intuitively, Mary’s epistemic life changes in more ways than just becoming acquainted with the experience of red. For as Bigelow and Pargetter (2006) point out, Mary also acquires, or at least is in a position to acquire, new ways of thinking and talking about the experience of red that were unavailable to her prior to her experience of it. For example, she can for the first time think or say things like “such and such is like this property of my experience”. Prior to her experience of red, she simply could not refer to it as “this property of my experience”. But intuitively, this change in her ability to think about, talk about, and refer to the experience of red suggests a genuine conceptual change in Mary.

Another reason for thinking that we should use the notion of “knowledge by acquaintance” to develop an account of phenomenal concepts is that it is unclear how the appeal to acquaintance all by itself can help the physicalist disarm other anti-physicalist

260 P. 362
arguments, such as the zombie argument, the modal argument, or the explanatory gap argument. On the other hand, one virtue of the phenomenal concepts strategy is that it promises a unified solution to all of the leading anti-physicalist arguments. And given that we have (defeasible) reason to prefer unified solutions to ones that are not unified, I take it that we therefore should prefer defenses of physicalism that put the notion of knowledge by acquaintance to use in characterizing our phenomenal concepts to ones that do not. Hence, in the rest of this section, I will use the notion of knowledge by acquaintance to develop my own account of phenomenal concepts. However, I will first give a brief summary of what I’ve said about “knowledge by acquaintance” in previous chapters.

Following Russell, I conceive of “knowledge by acquaintance” as a fundamental experiential relation between a conscious subject and individuals or features in the world, where the “knowledge of things” that it constitutes is logically independent of, but serves as the ultimate enabling condition for, our knowledge of truths about, and conceptions of, the things or features with which we are acquainted. I also follow Russell in holding that our acquaintance with things constitutes our most basic form of intentionality, or object-directedness, and that it is essentially an experiential relation in which a subject is presented with the objects of his or her awareness.

Furthermore, acquaintance is a distinctive epistemic relation that plays a fundamental role in reference and cognition. Indeed, Russell introduces the notion of acquaintance, or knowledge of things, in order to place a substantive cognitive constraint on what it takes for an individual to have genuinely singular thoughts about, or to make genuinely singular reference to, individuals. In other words, our knowledge of things
explains what it takes to have thoughts or utterances whose contents are constituted by
the very individuals and features, if any, that the thoughts or utterances are about rather
than by some way of getting at or identifying them. In fact, he argues that all cognition
and linguistic designation ultimately rests on our fundamental epistemic capacity to be
acquainted with, or consciously aware of, individuals and their features.

As I’ve noted several times already, Russell’s basic picture of cognition and
reference is this: There is some special class of individuals (concrete and abstract alike)
and their features with which we have experiential contact. On the basis of this
experiential contact, we are in a position to direct conscious attention to those individuals
and features. Our conscious attention to these experienced individuals and features in turn
grounds our ability to make demonstrative, singular reference to them (to designate them
with a “logically proper name”) and puts us in a position to acquire knowledge of such
demonstrative, singular reference. Furthermore, our conscious attention to experienced
individuals and features also puts us in a position to pick up information about, and
thereby form some conception of, these individuals and features. Finally, on the basis of
our conceptions of these individuals and features and our more general conception of
how the world works, we can designate objects and features with which we lack
experiential contact. We do so by employing representations that encode identifying
conditions that an object or feature must uniquely satisfy in order to be the designated
individual or feature of the representation, or by employing representations that are
purely quantificational in character.

As I’ve also noted previously, there are a number of strong epistemic theses that
are often (erroneously) associated with Russell’s notion of acquaintance. For example,
many have interpreted Russell as holding that when we are acquainted with a thing, we thereby attain infallible knowledge of at least some truths about that thing, and that these truths constitute the epistemic foundation of all of our knowledge of the world. Others have held that when we are acquainted with a thing, we are presented with the essential nature of the thing and thereby acquire, or are in a position to acquire, non-inferential revelatory knowledge of the essential nature of that thing. Most frequently, it is maintained that acquaintance constitutes an especially strong kind of knowledge—which, one that provides us with our knowledge of demonstrative, singular reference and that shields us from the possibility of reference-failure and misidentification.

However, as I argued in Chapter 1 of this dissertation, I do not think that we should build any of these epistemic theses into our notion of knowledge by acquaintance. For one thing, Russell is very explicit in asserting that our knowledge of things is logically independent of any knowledge of truths about those things, including presumably knowledge of semantic truths about them. For another, Russell clearly states that “all our knowledge of truths is infected with some degree of doubt”, which directly conflicts with the view that acquaintance provides us with infallible knowledge of at least some truths about its objects. Most importantly, however, Russell’s treatment of phenomenal continua cases demonstrates that acquaintance does not provide us with revelatory knowledge of, or infallible knowledge—which about, the objects of our acquaintance. We can, such phenomenal continua cases show us, stand in acquaintance relations multiple times over to one and the same thing and not know that we do.

In fact, the only sense in which we know a thing “perfectly and completely” when we are acquainted with it is this: if we stand in the acquaintance relation at all, we

---

must stand in it towards some existing object.\textsuperscript{262} There is no such thing as partial
acquaintance, and we cannot be acquainted with nothing at all. Thus, if we know \textit{that} we
are acquainted at all, then we know that there is an object of our acquaintance, and we
can be assured that any demonstrative, singular reference made on the basis of our
acquaintance will be immune to the possibility of reference-failure. However, contrary to
some of Russell’s own remarks, it is evident that being acquainted with a thing does not
automatically entail knowledge that one is acquainted with anything at all. Indeed, small
children and animals are acquainted with all kinds of things without even having the
capacity to know that they are.

Though Russell himself does not undertake to provide an ontological account (or
“analysis”) of the nature of acquaintance, it is evident that any physicalist who wishes to
use it in defense of physicalism must do so in a way that is consistent with his or her
physicalist commitments. In chapter two of this dissertation, I took up this challenge and
argued that given its ascribed epistemic features and role in reference and cognition, we
should identify acquaintance with the right sorts of original information relations between
a subject’s experiential states and corresponding states of the world. In other words, a
conscious subject is, on my view, acquainted with a thing if and only if he or she has a
phenomenally conscious state the occurrence of which puts him or her in ‘original’ direct
informational contact with that thing.

In chapter two, I went on to argue that we should understand what it is to be in
direct informational contact with a thing in terms of John Perry’s (2002) naturalistic
theory of information and informational content. According to Perry, \textit{information} is
“what one part or aspect of the universe (the signal) shows about some other part or

\textsuperscript{262} Russell (1912/1997), pp. 46-7.
aspect [of the universe] (the subject matter)” given the particular circumstances of the signal’s occurrence and given the way the world actually works. Infor-

mational content, in contrast, is what one part or aspect of the universe (the signal) shows about some part or aspect of the universe (the subject matter), if any, given the particular circumstances of the signal’s occurrence and assuming some (possibly false) constraints on the way the world works. Finally, original informational content is what one part or aspect of the universe (the signal) shows about some part of aspect of the universe (the subject matter, if any, given the particular circumstances of the signal’s occurrence, given some (possibly false) constraints on how the world works, and given the natural job of the signal (where the natural job of the signal is determined, together with its subjective, qualitative character, if any, by its selective, ontogenetic, and social history).

Given Perry’s framework, my proposal is that we are in the right sort of direct informational contact with a thing, and thereby acquainted with it, if and only if we have a phenomenally conscious signal bearing original informational content directly about that thing (i.e. not by means of the signal indicating it through its representational fit to identifying conditions encoded in the signal). Thus, my ultimate view is that our singular thought about, and reference to, things is enabled by our having experiences whose original informational content is directly about them.

I also contended that this identification of acquaintance with the right sort of information relations between experiences and corresponding states of the world has

---

264 However, my account of the conditions for singular thought and reference includes a social dimension that is absent from Russell’s own account. On my view, we have the capacity to have singular thoughts about, and make singular reference to, things only if we are personally acquainted with them or we use concepts acquired through communicative chains (and with the explicit or implicit intent to co-refer with those from whom we have acquired the concepts) where the concepts’ reference was initially grounded in some subject’s personal acquaintance with them.
important consequences for our understanding of the nature of perceptual acquaintance. In particular, I argued that in veridical perceptual experience, we are in original direct informational contact with, and thus acquainted with, both external objects (and their perceptible features) and with the qualitative characters of our experiential states themselves, albeit in radically different ways.

On one hand, we are acquainted with, in the sense of being “phenomenally” aware of, the qualitative characters of our conscious perceptual states themselves simply by having them. Indeed, our experiential states are, we might say, in the relevant non-reflective and non-propositional sense “self-presenting”. On the other hand, we are acquainted with, in the sense of being “intentionally” aware of, external objects, features, and states-of-affairs by being attuned to what the occurrence of our conscious perceptual states tell us about the rest of the world. Putting these two aspects together, we can say that conscious perception has what we might call a “two-faced presentational character”; we are all at once, and without intermediate inference, presented in conscious perception with both the object of our perceptual experience, if any, and the subjective, qualitative character of the experience itself.\(^\text{265}\)

With the above account of knowledge by acquaintance in hand, I am now in a position to present my acquaintance-based account of phenomenal concepts. On my view, there are five features which, when conjoined, distinguish genuine phenomenal concepts from our more familiar physical-theoretical concepts provided by the natural sciences (and folk versions thereof). First of all, genuine phenomenal concepts are genuinely singular, directly referential devices that pick out experiential states,

\(^\text{265}\) One important upshot of the “two-faced presentational character” of perception, I have argued, is that it shows how we can be direct realists about its content while being internalists about its qualitative character.
properties, or kinds. Secondly, their reference is fixed in a role-based manner by an inner demonstration on the basis of a subject’s selective introspective attention. Thirdly, they are for this reason “irreducibly perspectival” and hence not reducible to any discursive or descriptive concepts. Fourthly, their inner reference-fixing is ultimately made on the basis of our standing in special informational relations, or acquaintance relations, to our own experiential states. In particular, the reference of genuine phenomenal concepts is fixed by an inner demonstration (on the basis of selective introspective attention) to our self-presenting experiential states of which we are “phenomenally” aware simply by having them. And fifthly, because the reference of genuine phenomenal concepts is grounded in our standing in special informational relations to our own experiential states, relations which are tied to special epistemic methods and capacities for picking up information about those states, they provide us with distinctive conceptions of them, conceptions of their subjective, qualitative characters as subjective and qualitative. Nevertheless, these conceptions do not constitute revelatory knowledge of the essence of our experiential states, do not play a role in determining the reference of our phenomenal concepts, and do not figure into the subject matter content of thought and talk using such phenomenal concepts.

Let us consider each of these five features in turn. As I noted above, genuine phenomenal concepts are genuinely singular, directly referential devices that pick out experiential states, properties, or kinds. On my view, the distinctive feature of a genuinely singular concept is that its semantic contribution at the level of “what is thought” or “what is said” to the contents, or truth conditions, of our thought and talk is constituted by the very individuals or features, if any, that the device is about rather than by some
way of getting at or identifying them. Furthermore, I hold that genuinely singular concepts are “rigid designators”; that is, their semantic contributions to the subject matter content of our thought and talk are invariant across modal contexts.²⁶⁶

As for directly referential concepts, my view is that they are devices whose contributions to the subject matter contents of our thought and talk are determined by mechanisms other than the representational fit of that which is designated to some grasped conception of, or way of getting at, it.²⁶⁷ What’s more, the subject who deploys a directly referential concept in thought or talk needn’t have a reflective grasp of the (often external) mechanisms which determine its semantic contribution to the contents of our thought and talk. Rather, the subject need only be attuned to such reference-determining mechanisms.

As I argued in chapters one and two of this dissertation, both Russell and I hold that genuinely singular, directly referential concepts play a fundamental role in our reference to, and cognition of, worldly individuals and features, including experiential states and properties. For one thing, genuinely singular, directly referential concepts provide us with the capacity to think about, and refer to, things and properties in a way that is more primitive than, and does not presuppose, any knowledge of truths about, or conception of, them. Indeed, it is only after we acquire genuinely singular, directly referential concepts about things with which we are acquainted that we have the conceptual materials required for designating things descriptively or quantificationally.

For another thing, genuinely singular, directly referential concepts play an indispensable role in tying our thought and talk down to the individuals and features in

²⁶⁶ Kripke (1980).
²⁶⁷ Thus, as I noted in chapters two and four, my view about direct reference differs from Kaplan’s. On my view, what Kaplan calls “direct reference” is what I would call “genuinely singular designation”.

188
our environment. Because the designation of singular denoting devices in non-modal contexts is determined by representational fit, there is no assurance that our conceptions allow us to latch onto the objects we encounter in the world, for there is always the genuine possibility that our conceptions better fit remote objects, too many objects, or none at all. Hence, genuinely singular, directly referential devices are needed to anchor our thought and talk to the things and features in the environment with which we regularly interact.\(^{268}\)

Finally, genuinely singular, directly referential concepts provide creatures with finite cognitive capacities like us indispensible tools for navigating “Strawsonian Worlds” (like ours) of persisting objects and features in space-time.\(^{269}\) Indeed, it is absolutely crucial that our thought and communicative practices permit us to pick up information about the persisting individuals and features that we encounter (or might encounter) in experience in ways that allow us to interact successfully with them, to correct our false or incomplete conceptions of them, and to keep track of them through actual and possible changes in their features.

In the case of genuine phenomenal concepts, there is good reason for thinking that their reference is fixed in a role-based manner by an inner demonstration on the basis of a subject’s selective introspective attention. The key feature of concepts whose reference is fixed in a role-based manner is that they directly designate whatever things, properties, or kinds that bear the right sort of contextual role relative to the subject who possesses and

\(^{268}\) Some Fregeans have tried to solve this problem by claiming that there is always an indexical element in our conception-guided reference to things. However, there is no plausible descriptivist account of indexicality.

\(^{269}\) P. F. Strawson (1959), Perry (2002), and Føllesdal (1986, 2004).
deploys them. In the case of phenomenal concepts in particular, they directly designate experiential instances, properties, or kinds the conscious subject has and selectively attends to at the time of an inner, reference-fixing demonstration.

Of course, in saying this, we must be careful here not to equate concepts whose reference is fixed in a role-based manner with familiar role-based indexicals and demonstratives. Such concepts differ from familiar indexical and demonstrative expressions in natural language (and their analogues in thought) in that their reference doesn’t regularly shift in a role-based manner with their context of use; rather their reference becomes invariant across contexts of use after their initial contextually-based reference-fixing. For this reason, we should not think of genuine phenomenal concepts as indexical concepts, despite the fact that their reference is initially fixed in a role-based manner by an inner demonstration.

In any case, as noted in the previous chapter, I believe that we can appeal to at least two factors in explaining how such role-based phenomenal concepts achieve (more-or-less) determinate reference to experiential instances, properties, or kinds. First, we can appeal to the subject’s dispositions to reapply the concept in further cases at the time of initial reference-fixing. Indeed, many have effectively deployed this kind of strategy in the context of answering the qua problem for causal theories of reference, and I think it can be straightforwardly and fruitfully applied to cases of role-based reference-fixing as well.

---

270 I think that we can usefully think of roles as capturing patterns of contextual relations we bear to things where we abstract away from the particular things that stand in such relations to us.

271 The following account is meant to be compatible with the possibility that phenomenal concepts are vague.

272 This qualification is important because a subject’s dispositions to reapply a concept in further cases might change over time without thereby shifting its reference.
In addition, such dispositions to reapply a particular phenomenal concept in further cases partly determine whether it is a *token* phenomenal concept or a *type* phenomenal concept. Token phenomenal concepts are concepts used to think or talk directly about particular dated, unrepeateable experiences or phenomenal property-instances. Roughly speaking, a phenomenal concept is a token phenomenal concept only if a subject is disposed to use it to pick out a particular dated, unrepeateable experience or phenomenal property-instance and is disposed not to use it to pick out others. Type phenomenal concepts, on the other hand, are concepts used to think or talk directly about repeatable *kinds* of experiences or phenomenal properties. A phenomenal concept is a type phenomenal concept only if a subject is disposed to use it to pick out instances of a kind of experience or phenomenal property.

Second, we can explain how role-based phenomenal concepts achieve determinate reference by appealing to the nature of selective introspective attention itself. On my view, selective introspective attention works roughly as follows. First, we have experiences with some determinate character of which we are unreflectively sensitive to given the self-presenting character of phenomenal experience. On the basis of our unreflective sensitivity to our self-presenting experiential states, we become unreflectively sensitive to the determinate qualitative features such states have, and become attuned to the roles that such states and features play in our experiential lives. Next, we can rely on our unreflective sensitivity to our experiential states (and their qualitative features) and on our attunement to the roles that they play in our experiential lives to select out, and introspectively attend to, particular determinate experiential states or features with which we are presented. Finally, following what was said above, we can
appeal to our dispositions at the time of selective introspective attention to determine whether our selection of particular experiential states or features is an act of selecting out particular dated, unrepeatable experiences (or phenomenal property-instances) or an act of selecting out instances of repeatable experiential kinds.\textsuperscript{273}

One important consequence of the fact that the reference of genuine phenomenal concepts is fixed in a role-based manner is that they are therefore relatively uncontroversial instances of “irreducibly perspectival” concepts. The key feature of an “irreducibly perspectival” concept is that we cannot determine a priori what it picks out given a purely discursive, third-person description of the world; we can only determine its reference by determining which object, feature, or state-of-affairs occupies the right sort of contextual role relative to a relevant point-of-view (where most often the relevant point-of-view is that of the user of the concept). Simply put, to determine the reference of an “irreducibly perspectival” concept, we must look to see how the concept and its user are situated in the world.

Because genuine phenomenal concepts are ‘irreducibly perspectival’ in this way, they generate an epistemic or conceptual gap analogous to the familiar gap between indexicals or demonstratrices and co-referring nonperspectival concepts. For as Perry (1979) pointed out some time ago, we can know all the facts about the clueless guy that is leaving a trail of sugar on the supermarket floor as expressed with nonperspectival concepts without thereby having the self-locating belief we would express as “\textit{I am the guy making the mess}”. Similarly, as we saw in the previous chapter, we can know everything objective there is to know about Dretske without thereby knowing that “\textit{that}

\textsuperscript{273} This view of selective introspective attention is noncommittal about the epistemic and explanatory priority between phenomenal concepts and perceptually-based concepts.
man (the one I now see in front of me) is Dretske”. In likewise fashion, I believe, we can know all the physical facts as expressed entirely by objective physical-theoretical concepts without thereby knowing which part of the physical domain, if any, a genuine phenomenal concept picks out. Thus, the demonstrative-recognitional model is correct in asserting that there is an epistemic or conceptual gap between phenomenal concepts and physical-theoretical concepts that mirrors more familiar gaps between the perspectival and the nonperspectival.

Nevertheless, the epistemic or conceptual gap between the experiential and the physical isn’t wholly captured by the gap between the perspectival and the nonperspectival. Rather, the distinctive epistemic or conceptual gap between the experiential and the physical is the result of our phenomenal concepts, on one hand, and our physical-theoretical concepts, on the other, being grounded by two radically different kinds of original direct informational relations, or acquaintance relations, to the world, where these relations are tied to different epistemic methods and capacities for picking up information about it, and thereby provide us with radically different conceptions of it.

Indeed, as I noted above, I identify the notion of “knowledge by acquaintance” with the right sort of original direct informational relations between a subject’s phenomenally conscious signals and corresponding parts or aspects of the world. And typically, when are acquainted with something, we are acquainted with some part of the world distinct from our phenomenally conscious signal in virtue of being attuned to what the occurrence of that signal tells us about the relevant part of the world with which we are in direct informational contact (given the circumstances of the signal’s occurrence
and our sensitivity to how the world works). We therefore have what I’ve previously called “intentional” awareness of that part of the world.

However, in the case of our acquaintance with the subjective, qualitative characters of our own experiential states, things are quite different. Indeed, when we are acquainted with the subjective, qualitative characters of our own experiential states, we are in original direct informational contact with them in virtue of our being attuned to what the occurrence of our phenomenally conscious signals tell us about themselves (again, given the circumstances of their occurrence and our sensitivity to how the world works). We therefore have what I’ve previously called “phenomenal” awareness of the subjective, qualitative characters of such “self-presenting” experiential states.

Furthermore, my view is that we have such “phenomenal” awareness of the subjective, qualitative characters of our experiential states simply by having them. When it comes to the subjective, qualitative characters of our own experiential states, the having is, in the relevant non-reflective and non-propositional sense, the knowing.

Of course, there are many deep questions about the relation between the subjective, qualitative characters of experience and their self-presentational character. Are they merely two side of the same coin, as John Searle seems to think? Does the fact that our experiences put us in original direct informational contact with themselves explain their subjective and qualitative characters, which would vindicate certain aspects

---

274 This view is compatible with, but does not entail, the popular view that we have experiential states only when we have higher-order representations about our first-order representational states. Personally, I am skeptical of the viability of higher-order theories of consciousness. But since my goal here is not to solve the hard problem of consciousness but rather to show why there is an epistemic or conceptual gap between the experiential and the physical that is compatible with physicalism, I take no stand on how to explain consciousness itself. And, in my opinion, the hard problem of consciousness just is the problem of explaining how it is that physical states can be self-presenting such that we are “phenomenally” aware of their subjective, qualitative features simply by having them.

275 Strawson (2009) makes roughly the same point on page 286.
of higher-order theories of consciousness? Or conversely, is the special self-presenting character of our conscious experiences importantly explained by their subjective qualitative characters? I will not here take a stand on these issues. While the answer to them will shed important light on the hard problem of consciousness, my concern here is not to solve that problem. Rather, it is simply to explain why there is an epistemic gap between the experiential and the physical, and why the epistemic gap is not indicative of a corresponding ontological gap.

One advantage of identifying “knowledge by acquaintance” with the right sort of original direct informational relations between a subject’s phenomenally conscious signals and corresponding parts or aspects of the world is that it makes clear how acquaintance can be a species of “knowledge” that is distinct from, yet an enabling condition for, our knowledge of truths about, and conceptions of, the objects of our acquaintance. Indeed, following Recanati (2010), we might say that unlike many relations (including most kinds of contextual relations), direct informational relations are “epistemically rewarding” relations for systems that can exploit them. Basically, an epistemically rewarding relation is one that enables the right kind of system to pick up information about the objects standing at the other end of the relation. I follow Perry (2002) in calling systems capable of exploiting the direct informational relations they bear to things to pick up information about them “information harnessing systems”. I also follow Perry in holding that we are basically sophisticated, biological information harnessing systems that exploit our acquaintance with things to pick up information about them that, if accurate, will help us to interact successfully with them, or with other parts of the world.
Because they are typically designed only to perform relatively simple tasks, simple information harnessing systems are most often constructed in ways that allow them to exploit only a very limited number of direct information relations they bear to things, and only by means of relatively simple and goal-specific methods and capacities for picking up information about them. On the other hand, complex information harnessing systems often exploit a variety of different direct information relations they bear to things to pick up various sorts of information about them, often by deploying various kinds of information-gathering methods and capacities. And by exploiting a variety of different direct information relations to things to pick up various sorts of information about them, complex information harnessing systems are capable of performing far more complex and sophisticated tasks than simple information harnessing devices, and in a much wider range of circumstances.

As far as we know, we are without question the most complex and sophisticated information harnessing devices in the universe. Not only do we have a number of exteroceptive, interoceptive, and proprioceptive faculties for picking up various sorts of information about our environment, our internal organs, and our movement and bodily orientation, respectively, we also have special introspective methods for picking up information about our own mental states, including the subjective, qualitative characters of our own experiential states. These various epistemic faculties allow us to deploy a number of different kinds of information-gathering methods, in a range of different circumstances, to pick up different kinds of information about things by exploiting the different sorts of direct information relations we bear to them. Not only does this provide

---

276 Of course, we needn’t (and shouldn’t) conceive of “introspection” literally. That is, we needn’t think of introspection as inwardly “perceiving” or “looking at” our own mental states.
us with the means to acquire knowledge about numerous different kinds of things or features in the world, it also provides us with the means to pick up different kinds of information about one and the same thing or feature. We are thereby able to acquire different conceptions of, or knowledge about, one and the same thing or feature.

Nevertheless, we must be careful not to underemphasize the radical differences between our introspective information-gathering methods and those of our other information-gathering faculties. As I noted above, the direct information relations we bear to our own experiential states are importantly different in kind from those we bear to other parts or aspects of the world, including those parts of aspects of the world that our interoceptive and proprioceptive faculties tell us about. Indeed, whereas our acquaintance with other parts or aspects of the world consists in our having a phenomenal experience that has some distinct part or aspect of the world as its original direct informational content, our acquaintance with the subjective, qualitative characters of our own experiential states consists in our having a phenomenal experience whose original direct informational content is itself. Given this, we should not be surprised to find that our introspective information-gathering methods are radically different in kind from other information-gathering methods, or to find that they provide us with a radically different kind of information about, and thus conception of, our own experiential states.

And in fact, because our genuine phenomenal concepts are grounded in our special acquaintance with our own experiential states, they provide us with conceptions of their subjective, qualitative features as subjective, qualitative features. Of course, given the nature of acquaintance, we must be careful here. In claiming that genuine phenomenal concepts provide us with conceptions of the subjective, qualitative features
of our experiences *as* subjective, qualitative features, we are not thereby committed to the view that acquaintance *reveals* their essential natures. Indeed, there remains the genuine possibility that our special acquaintance-based conceptions of the subjective, qualitative features of experience are mistaken or incomplete, and in any case, they play no role in determining the reference of our phenomenal concepts and do not secure us against the possibility of misidentification.

With that said, we are now in a position to compare genuine phenomenal concepts with physical-theoretical concepts as well as with indexical concepts and perceptual demonstratives. *Phenomenal concepts*, we have seen, are irreducibly perspectival, genuinely singular, and directly referential concepts that pick out our own self-presenting experiential states, properties, or kinds *as* experiential states, properties, or kinds in a role-based manner on the basis of our introspective selective attention. They importantly differ from *indexical concepts*, which are role-based, irreducibly perspectival, genuinely singular, and directly referential concepts, in that they specifically designate self-presenting experiential kinds *as* experiential kinds, and in that their reference does not shift from context to context. Furthermore, they differ importantly from *perceptual demonstratives*, which are role-based, irreducibly perspectival, genuinely singular, and directly referential concepts, in that they do not have the job of picking out, on the basis of selective perceptual attention, perceptible objects, features, and states-of-affairs in the environment of which we have “intentional” perceptual awareness.

Finally, and most importantly, genuine phenomenal concepts differ radically from genuinely singular *physical-theoretical concepts* in two crucial ways. First, unlike genuine phenomenal concepts, physical-theoretical concepts have the job to pick out
physical entities, states, properties, or kinds either by descriptive fit, or by role-based, directly referential perceptual demonstration (on the basis of selective perceptual attention to things with which we have “intentional” perceptual acquaintance). And second, whereas the job of genuine phenomenal concepts is to provide us with conceptions of the subjective, qualitative features of experience as subjective, qualitative features, the primary job of physical-theoretical concepts is, roughly speaking, to provide us with conceptions of what things there are in the world and what role they play in its workings.  

Nevertheless, despite their radical differences, we must keep in mind an extremely important respect in which physical-theoretical concepts and phenomenal concepts are alike. Indeed, as in the case of genuine phenomenal concepts, the reference of physical-theoretical concepts is ultimately grounded in our fundamental acquaintance with things, though in this case it is our perceptual acquaintance with them that is involved. Consequently, we have no reason to suppose that physical-theoretical concepts provide us with revelatory knowledge of the essential nature of the physical, or that they secure us from the possibility of misidentification. But once this is granted, the dualist is in trouble. For as we’ll see in the following sections, once we recognize that neither physical-theoretical concepts nor phenomenal concepts provide us with revelatory knowledge about, or infallible discriminating conceptions of, the individuals and

---

277 In point of fact, there is a rather complicated story about just what sorts of conceptions of things are provided by our physical-theoretical concepts. However, telling the full story, or even a moderately adequate one, would be an entirely different (and quite frankly enormous) project.

278 Indeed, as Russell taught us, even when the reference of a physical-theoretical concept is fixed by description, the descriptive materials required for its initial reference-fixing are ultimately acquired on the basis of our (typically perceptual) acquaintance with things.
features they designate, we cannot infer an ontological gap between the physical and the experiential from the epistemic or explanatory gap between them.

§3 THE ZOMBIE ARGUMENT

As we’ve seen, (phenomenal) zombies are hypothetical creatures who are stipulated to be indiscernible from us in every physical respect but to lack experiential lives like ours altogether. We’ve also seen that their creator, David Chalmers, designed these gruesome creatures to serve a singular purpose: to show that physicalism is false by putting pressure on its claim that the physical facts of this world settle the experiential facts of this world with at least metaphysical necessity. Chalmers makes his case with the following “zombie argument” against physicalism:

1) Zombies are conceivable (in the right way).
2) If zombies are conceivable (in the right way), then they are metaphysically possible.
3) If zombies are metaphysically possible, then physicalism is false.
4) Therefore, physicalism is false.279

Chalmers argues that zombies are conceivable in the right way since an ideal reasoner could positively imagine a (epistemic) scenario in which there are creatures that are indiscernible from us in every physical respect but that lack experiential lives altogether. He also maintains that conceivable circumstances are metaphysically possible

---

279 As we saw in the last chapter, we can also formulate the argument as follows (where \( P \) stands for a complete description of all of the physical facts in this world and \( E \) is a complete description of all the experiential facts in this world): (1) \( P & \neg E \) is conceivable (in the right way), (2) if \( P & \neg E \) is conceivable (in the right way), then \( P & \neg E \) is metaphysically possible, and (3) if \( P & \neg E \) is metaphysically possible, then physicalism is false; therefore, (4) physicalism is false. The advantage of this formulation is that it works equally well for physically indiscernible phenomenal inverts, and other denizens in the anti-physicalist arsenal. See Chalmers (2002).
circumstances just in case an ideal reasoner could positively imagine a (epistemic) scenario in which those circumstances obtain (or are “verified” to use Chalmers’ preferred terminology). Thus, he concludes, zombies are metaphysically possible and therefore physicalism is false.

Physicalists have quite naturally objected to various parts of Chalmers’ zombie argument. Dennett (1998), for instance, has questioned both the conceivability of zombies and the coherence of the very notion of a zombie. Beisecker (2010) has questioned whether the metaphysical possibility of zombies would render physicalism false on the grounds that we might turn out to be zombies. But while I am somewhat sympathetic to both of these worries, I agree with the majority of physicalists in holding that the weakest part of the zombie argument is its claim that the mere conceivability of zombies (in the right way) warrants conclusions concerning their metaphysical possibility.

Pointing to standard Kripkean cases of a posteriori identities, these physicalists rightly note that there are conceivable circumstances that are nevertheless not metaphysical possible. For instance, it is (in some sense) conceivable that water is not H$_2$O, but it is not metaphysically possible that water is not H$_2$O given that “water” and “H$_2$O” rigidly designate one and the same stuff. However, most physicalists who question the link between the conceivability and metaphysical possibility of zombies grant dualists that the relevant psycho-physical identities are not entirely analogous to other instances of a posteriori identity. The difference is, they argue, that unlike most cases of a posteriori identity (which identify the referents of two distinct physical-theoretical concepts), psycho-physical identities identify the referent of a genuine phenomenal

---

280 However, it is not clear that Beisecker’s strategy can work given how I have stipulated zombies here.
concept with the referent of a physical-theoretical concept. But despite the special character of psycho-physical identities, they maintain, Kripke’s key insights about *a posteriori* identity apply equally to them.

Unsurprisingly, dualists have resisted the strategy of using special phenomenal concepts to undermine the zombie argument. One of their most common objections is that *a posteriori* identities do not sever the link between conceivability and metaphysical possibility in quite the way physicalists imagine. While dualists grant that *we* are not always in a position to use conceivability to track metaphysical possibility when an *a posteriori* identity is at stake, an ideal reasoner with sufficient knowledge of the world could *a priori* determine whether or not the identity is a true one. For instance, they argue, an ideal reasoner would be in a position to conclude *a priori* that water is H$_2$O provided he or she grasps the meanings of “water” and “H$_2$O” and has a sufficiently complete qualitative description of the world (supplemented with an “I am here” marker). And while we are not *ideal* reasoners with complete qualitative knowledge of the world, we can nevertheless trust our (positive) conceivability to track metaphysical possibility in cases where we have a sufficiently revealing grasp of the meanings of our concepts and have no reason to suspect that we are making any mere logical mistake.

Dualists argue that the same basic considerations apply to the case of psycho-physical identities. If the experiential is indeed identical with the physical, they maintain, then an ideal reasoner with a sufficiently revealing conception of the physical and the

---

281 Many of their objections involve aspects of the modal argument, the knowledge argument, the explanatory gap, or the two-dimensional argument against physicalism. Since we will consider these arguments in the following sections of this paper, I will here focus on issues that are mostly restricted to the zombie argument.

282 Presumably, the reason why *we* needed to do *a posteriori* investigation to discover that water is H$_2$O is that we needed to acquire sufficiently revealing conceptions of water and H$_2$O before we could draw the relevant *a priori* conclusion.
experiential should be able to make the identification of them \textit{a priori} (at least given enough qualitative information about the world). However, they argue, psycho-physical identities are intuitively not knowable \textit{a priori} even to an ideal reasoner with mastery of the relevant concepts and a sufficiently complete qualitative description of the world (supplemented with the “I am here” indexical knowledge). In fact, they maintain that \textit{we} have sufficiently revealing conceptions of experience and the physical (or at least of its general character) such that we are in a position to trust our (positive) conceivability concerning them to track genuine metaphysical possibility. So, they conclude, Kripkean insights concerning \textit{a posteriori} necessities cannot help defuse the zombie argument.

In my view, it is quite doubtful that even an ideal reasoner with complete qualitative knowledge about the world would be able to determine all true identities \textit{a priori} simply on the basis of his or her mastery of the relevant concepts. Hence, I am skeptical about any straightforward link between (the right kind of) conceivability and metaphysical possibility. But rather than entering into a speculative dispute about the intellectual powers some hypothetical ideal reasoner might possess, I will say why, given the nature of our phenomenal concepts, \textit{our} conceivability regarding consciousness is no sure guide to metaphysical possibility. At the same time, I will explain why there is such a strong \textit{intuition} of distinctness in the case of the experiential and the physical.

I want to begin by drawing attention to an aspect of conceivability that I believe has not been sufficiently appreciated by proponents of the zombie argument: On one hand, the conceivability of such-and-such circumstances is driven by the \textit{conceptions} that are tied to the concepts we use in so conceiving. But on the other hand, the metaphysical possibility of such-and-such circumstances depends on how the things and features, if
any, our concepts are *about* might otherwise be. Hence, even positive conceivable is a guide to metaphysical possibility (at best) only to the extent that it relies on conceptions that sufficiently reveal the nature of the things, if any, our concepts are about (and thereby how they might otherwise be). So, conceivable can only provide us with genuine insight into the ontological nature of the experiential and the physical given that we have conceptions of them that sufficiently reveal their natures. But do we have sufficiently revealing conceptions of the experiential and the physical?

There are compelling reasons for thinking that neither our conception of the experiential nor our conception of the physical is sufficiently revealing to warrant the conclusion that the experiential and the physical *are* distinct from the mere conceivable of their distinctness. On the physical side, Russell (1912/1997, 1927a/1954, and 1948) and others have made a strong case that while physical-theoretical concepts provide us with conceptions of what things there are in the physical domain, how they are distributed in space and time (or space-time), and how they are disposed to behave in various circumstances, they do not provide us with conceptions that characterize the intrinsic, categorical nature of the physical entities they are about. And on the experiential side, Russell’s discussion of phenomenal continua cases, which we considered in Chapter One of this dissertation, show that our acquaintance with the experiential does not provide us with a revelatory knowledge of its nature, though it does put us in a position to form conceptions of its subjective, qualitative character *as* subjective and qualitative. So in either case, there is no assurance that our conceptions

---

283 Of course, this won’t serve as a general characterization of metaphysical possibility since I appeal to the *possible* way things might be to explain the *possible* circumstances that might obtain. I remain neutral on this matter.
provide us with sufficiently accurate and complete knowledge of their referents such that conceivability can track their numerical identity or diversity.

In fact, the radical *differences* in the conceptions associated with our physical-theoretical concepts and our phenomenal concepts provide us with no traction for drawing conclusions about the numerical identity or diversity or their referents. In cases where the conceptions associated with distinct concepts are of the same general character but say conflicting things about the same kinds of features of their respective referents, then we are typically in a position to use conceivability to conclude either that they are concepts of two distinct things or that one or more of the conceptions is false. But when distinct concepts provide conceptions with very different characters and say very different sorts of things about their respective referents, then it could be that they are concepts of two different things, but it could be that they instead provide two different, (mostly non-overlapping) partial takes on one and the same thing (or feature). In this second kind of case, we cannot simply depend on conceivability alone to settle the matter; we must bring outside considerations to bear, or simply withhold judgment. And this is precisely the situation we are in with respect to the experiential and the physical.

But even setting aside these important general worries about our lacking sufficiently revealing conceptions of the physical and the experiential to draw ontological conclusions, there are more specific reasons to distrust our conceivability when applied to the case of psycho-physical identities. Indeed, if my account of phenomenal and physical-theoretical concepts is correct, they have a number of additional features which plausibly sever the link between (the right sort of) conceivability and metaphysical possibility. For

---

284 In my view, it is these radical differences in conceptions that explain what Papineau (2002) calls the “intuition of distinctness” that lingers even after we’ve become convinced of the truth of psycho-physical identities, an intuition we do not similarly retain in most cases of *a posteriori* identity.
one thing, they are grounded very different sorts of acquaintance relations, ones which put us in the position to pick up different kinds of information about the things that stand in very different roles relative to our cognitive (and practical) lives.\textsuperscript{285} For another, their respective references are (typically) fixed on the basis of two different sorts of attunement-guided attentive demonstration (i.e. introspective and perceptual demonstration) which generates well-known perspectival-objective gaps. And as proponents of the demonstrative-recognitional model rightly point out, these perspectival-objective gaps render circumstances conceivable that are not metaphysically possible (e.g. that \textit{this} man [the one I see in front of me now] is not Dretske when he is in fact Dretske).\textsuperscript{286} Finally, as we saw in the previous chapter, they play very different (and independent) roles in our cognitive lives. There is no conceptual or logical link between them. They are not defined in terms of one another. And they are likely implemented by very different mechanisms in our brains. Hence, there is the prima facie possibility that the conceivability of zombies is only possible due to contingent features (or limitations) of our own cognitive systems.\textsuperscript{287}

In short, even if zombies are conceivable (in the right way) for us, we have insufficient reason to take their conceivability as indicative of an ontological gap between the physical and the experiential. As many physicalists have pointed out, phenomenal concepts have a number of special features which generate various kinds of epistemic gaps which have no corresponding ontological gaps. But even more importantly, I have argued, we have good reason to hold that neither phenomenal concepts nor physical-theoretical concepts provide us with sufficiently revealing conceptions of the experiential

\textsuperscript{285} See Perry (2009) for similar remarks.
\textsuperscript{286} Perry (1979, 2001a, and 2009) and Levin (2007).
\textsuperscript{287} For similar remarks, see Perry (2009), Churchland (1996), McGinn (1989), and Stoljar (2006).
and the physical such that we can take the radical differences in our conceptions of each
to indicate that they are numerically distinct. To my knowledge, dualists have not been
pressed sufficiently on this point, but it cuts straight into the heart of the zombie
argument. For if I am correct, even a dualist should grant that the ideal positive
conceivability of zombies provides us with no reason to conclude that they are
metaphysically possible. Hence, the zombie argument fails to show that physicalism is
false.\textsuperscript{288}

\section{The Knowledge Argument}

In the previous chapter, I presented Frank Jackson’s (1982) famous knowledge argument.
The basic form of the argument is this: one can (in principle) have a complete physical
description of the world without having had such-and-such experience and yet learn
something substantive about oneself (and others) by actually having the relevant
experience, thereby (purportedly) showing that a complete physical description of the
world doesn’t capture all the substantial aspects of (concrete) reality. Hence, dualists
conclude, there is more to (concrete) reality than \textit{physical} reality.

More precisely, we can formulate Jackson’s knowledge argument as follows:

1. Prior to having some experience of type $E$, a subject $S$ can know all of the
   physical facts about (or relevant to) $E$.

2. After having an experience of type $E$, $S$ learns what it is like for $S$ (and others
   similarly situated and endowed) to have experiences of type $E$.

\textsuperscript{288} As noted above, dualists often supplement the zombie argument with various features from the other
leading anti-physicalist argument. Hence, I have not fully addresses every possible version of the zombie
argument. Even still, I will indirectly undermine these additional incarnations of the zombie argument by
attacking the remaining anti-physicalist arguments directly. For if they all fail individually, we have little
reason to think that they will succeed when conjoined in various ways.
(3) Thus, after having an experience of type $E$, $S$ learns new facts about experiences of type $E$.

(4) So before $S$ has an experience of type $E$, $S$ does not know all of the facts about (or relevant to) experiences of type $E$.

(5) Therefore, the physical facts about (or relevant to) experiences of type $E$ do not exhaust all of facts about (or relevant to) experiences of type $E$, and physicalism is therefore false.\(^{289}\)

Jackson illustrates this argument, of course, with the well-known case of Mary, who (somehow or other) knows all the physical facts relevant to chromatic color perception without having actually had the perceptual experience as of seeing red in standard viewing conditions, but who learns something new and substantial about “what it is like” to have perceptual experiences of chromatic color once she is no longer confined to her black-and-white laboratory. Jackson takes it that this new knowledge of Mary’s is knowledge of new non-physical facts, which is why she didn’t already know them on the basis of her pre-release exhaustive physical knowledge about human color perception. Hence, he concludes, physicalism is false.\(^{290}\)

\(^{289}\) In point of fact, Jackson’s (1982) initial presentation of the knowledge argument turns on whether possession of complete physical information about chromatic color perception exhausts all of the information there is to be had about chromatic color perception. However, Jackson never provides a substantive theory about the nature of “information”, nor does he explain how physical information is to be distinguished from other sorts of information. And since Jackson clarifies the knowledge argument in terms of physical facts in his (1986) response to Churchland, I have followed suit and formulated it in terms of knowledge of facts rather than information.

\(^{290}\) It is worth noting that Jackson no longer sees the knowledge argument as a sound one. On his current view (2004), the inference from (2) to (3) is fallacious since what $S$ learns in having an experience of type $E$ is what it is like to represent the world as being $E$. For instance, what it is like to experience red objects visually simple consists in visually representing the world as having red objects. But, Jackson argues, acquiring new ways of representing the world as being thus and so does not entail that what is represented is a new fact. Moreover, he contends, such representational features of our perceptual systems are a priori derivable from a complete physical-theoretical description of the world (at least for an ideal reasoned), and thus pose no threat to physicalism. I think that there are deep and underappreciated problems with such
There is no shortage of objections to the soundness of the knowledge argument, one of the most promising of which is the phenomenal concepts strategy. For in addition to having the experience of chromatic color experience (and thereby becoming acquainted with what it is like to have them) and acquiring new abilities to imagine, recognize, and classify such experiences in terms of their qualitative characters, Mary is in a position to acquire new concepts of chromatic color experiences—concepts that enable her to think and talk about what it is like for subjects (similarly situated and endowed) to have chromatic color experiences. Furthermore, her acquisition of new concepts of chromatic color experiences are plausibly accompanied with new capacities to entertain thoughts about conscious experience whose truth or falsity turn on what the facts are with respect to those conscious experiences. In this way, Mary’s first experience of chromatic color evidently puts her in a position to acquire new factual knowledge of what it is like for subjects (similarly situated and endowed) to have the relevant kinds of experiences.

The crucial question, however, is whether this new factual knowledge is knowledge of new non-physical facts. The phenomenal concepts strategy, we have seen, provides physicalists with a powerful tool for explaining how Mary acquires new factual knowledge without acquiring knowledge of new non-physical facts. According to proponents of the phenomenal concepts strategy, what Mary acquires is simply a new concept which (directly) picks out one and the same physical/phenomenal properties about which she already had physical-theoretical concepts in the black-and-white room.

---

representational accounts of the qualitative character of experience, but I must leave my objections for another occasion.

291 In fact, the new abilities Mary acquires are most naturally explained in terms of her acquisition of new concepts and conceptual abilities.
So while Mary is in a position to entertain (and express) new factual thoughts about conscious experiences after her first experience of chromatic color perception, such thoughts (and utterances) concern one and the same facts she could entertain (and express) all along by means of co-referring physical-theoretical concepts.

However, one lesson from the previous chapter is that a plausible physicalist account of phenomenal concepts must do more than explain how Mary might think (and express) new, conceptually irreducible thoughts about physical facts she already knew in the black-and-white room. It must also provide an adequate and plausible characterization of the various changes in Mary’s epistemic situation with respect her own conscious experiences (and those of others) given her new experience of chromatic color. But neither of the leading models of phenomenal concepts is successful in doing so. On one hand, the demonstrative-recognitional model does not provide enough materials to explain the distinctive epistemic changes Mary undergoes after having the experience of chromatic color. And on the other hand, the constitutional model suggests far too intimate of a relation between Mary’s new experiences and her newly acquired capacities to think about them using genuine phenomenal concepts.

Fortunately, I believe that my acquaintance-based model of phenomenal concepts fares far better than either theory with respect to capturing the substantive epistemic changes in Mary’s epistemic situation regarding conscious experience. To see how, let us consider Mary’s epistemic situation with respect to chromatic color experiences at various stages in Jackson’s thought experiment.

First, let us consider Mary’s epistemic situation with respect to chromatic color experiences while she is confined to the black-and-white laboratory, which following
Perry (2001a) I will call “the Jackson Room”. The first thing we should note is that even when Mary is confined to the Jackson Room, she is capable of possessing knowledge of several important facts about chromatic color experiences. For instance, it is widely granted that Mary can possess testimonial knowledge that individuals with normal vision will have such and such chromatic color experiences when viewing such and such physical objects and properties in such and such circumstances. What’s more, given her advanced physical-theoretical knowledge of physics, neurobiology, and perceptual psychology, she will know that such and such physical and physiological states or processes occur in individuals with normal vision when viewing such and such physical objects and properties in such and such circumstances.²⁹²

She will presumably also know what physical and physiological states or processes occur in individuals with abnormal vision in the very same circumstances, as well as which chromatic color experiences these physical and physiological states or processes correlate with, if any, in individuals with normal vision in the very same circumstances. Hence, assuming she has complete physical-theoretical knowledge about the normality (or abnormality) of her own vision, the physical objects and properties she will encounter on leaving the Jackson Room, and the circumstances under which she will view them, Mary will seemingly be in a position in the Jackson Room to predict which chromatic color experiences she will have upon her release. Mary can possess all of this

²⁹² If Mary is a physicalist, as on Perry’s (2001a) account, she will also believe that the qualitative characters of chromatic color experiences are identical with such and such physical phenomena involved in the perception of chromatic color. However, Mary’s antecedent ontological commitments are inessential to whether the knowledge argument warrants the conclusion that qualia are non-physical properties.
knowledge (given sufficient empirical evidence) simply by possessing the right physical-theoretical concepts and the right public folk-psychological concepts.\textsuperscript{293}

We can even imagine that Mary has visually encountered (in black-and-white and grayscale) the relevant physical and physiological states individuals with normal (and abnormal) vision undergo when perceiving such and such physical objects and properties in such and such circumstances (as well as all of the aforementioned objects, properties, and circumstances). In other words, we can imagine that Mary is \textit{acquainted} with all of the physical phenomena involved in chromatic color experience, at least in the sense of having had direct “intentional awareness” of them, by having black-and-white and grayscale perceptual experience of them.\textsuperscript{294} So she can even have demonstrative-recognitional knowledge of all of the relevant physical phenomena, though only the knowledge of this sort that is attainable on the basis of black-and-white- and grayscale-based perceptual demonstration.

Hence, while confined to the Jackson Room, Mary will have a quite rich conception of chromatic color experiences. Nevertheless, Mary is subject to a number of epistemic prohibitions regarding chromatic color experiences, prohibitions which leave her with an incomplete conception of them. The most important of these is the simple fact that he has never \textit{had} experiences as of chromatic color. As such, she has never been acquainted with chromatic color experiences in the special first-personal way; she has never had “phenomenal awareness” of their self-presenting phenomenal characters. She

\textsuperscript{293} If we put this point in terms of Chalmers’ (2004) classifications, Mary can in principle know all of the facts about chromatic color experiences expressible in terms of “community relational phenomenal concepts” and “individual relational phenomenal concepts” in addition to those expressible in terms of physical-theoretical concepts.

\textsuperscript{294} We should notice that standard acquaintance responses to the knowledge argument, such as Churchland (1985, 1989), Conee (1994) and Tye (2009), do not have enough resources to allow for this possibility and still explain how Mary’s epistemic situation with respect to chromatic color experiences changes upon her release.
has never been in a position to select and attend to chromatic color experiences
introspectively, or to deploy an introspective-based demonstrative to designate them.
Hence, she has never been in a position to deploy our special first-personal epistemic
methods and capacities for picking up information about the qualitative characters of our
own chromatic color experiences and thereby form a conception of such qualitative
characters as subjective and qualitative. Consequently, she is unable to use such
introspectively-based conceptions of the subjective, qualitative characters of chromatic
color experiences to reflect on “what it is like” for others to have similar experiences. In
sum, Mary will lack a great deal of substantive factual knowledge about chromatic color
experiences in the Jackson Room.

After being released from the Jackson Room, we can imagine that Mary enters
what Perry (2001a) calls the “Nida-Rümelin Room”. In the Nida-Rümelin Room, Mary
has her first experiences of chromatic color experiences in the form of randomly scattered
patches of chromatic color on the walls with nothing marking which color is which.
Moreover, we can suppose that for whatever reason, Mary does not yet exercise her vast
physical and physiological knowledge to determine which neurobiological states and
processes she undergoes, which color patches have which reflectance profiles, or what
the current viewing conditions are. Instead, Mary simply enjoys, even savors, her first
experiences of chromatic color.

Mary undergoes a number of significant epistemic changes after initially entering
the Nida-Rümelin Room. First of all, she will have chromatic color experiences for the
first time (and will undergo the corresponding neurobiological states and processes for
the first time). By having such self-presenting chromatic color experiences for the first
time, she will become acquainted with them (in a new way) in the sense of having newfound “phenomenal awareness” of them. In turn, she will be in a position for the first time to select and attend to them introspectively; she will therefore be in a position for the first time to deploy an introspectively-based demonstrative to designate their subjective, qualitative characters. Moreover, she will be in a position for the first time to deploy our special first-personal methods and capacities for picking up information about our own chromatic color experiences and thereby forming a conception of their subjective, qualitative characters as subjective and qualitative.

In my view, Mary will now possess genuine phenomenal concepts. These new concepts provide her with new conceptions of what it is like for her (and others) to have various chromatic color experiences. She will thereby acquire new conceptual capacities to identify, classify, and recognize chromatic color experiences in terms of what it is like to have them. She will be able to entertain many new thoughts about chromatic color experiences in terms of what it is like to have them. But what Mary is thinking about when deploying these genuine phenomenal concepts is not determined by representational fit. Rather, what Mary thinks about when deploying these genuine phenomenal concepts are the phenomenal properties of the experiences Mary actually has, attends to, and introspectively demonstrates at the time she forms the relevant

---

295 At this point, Mary forms what Chalmers (2004) calls an “E” concept (or an “indexical phenomenal concept”). Of course, I am greatly oversimplifying what it takes to move from acquaintance with something to forming a full-blown concept (demonstrative or otherwise) of that thing. One can tell any number of detailed stories here.

296 Chalmers (2004) would say that Mary now possesses a “pure phenomenal concept”. On his view, a pure phenomenal concept picks out a phenomenal property in terms of its “intrinsic phenomenal nature”. Consequently, he argues, pure phenomenal concepts have the same epistemic and subjunctive intensions in all possible worlds. As we will see in §7, our genuine phenomenal concepts do not have these strong epistemic features. So we should reply to Chalmers either by denying that we have any pure phenomenal concepts, or by saying that he has mischaracterized the epistemic features of our pure phenomenal concepts.
concepts. Indeed, Mary’s conceptions of designated phenomenal might be (and in my view will be) incomplete, inaccurate, and/or not uniquely distinguishing. Even still, her genuine phenomenal concepts stand directly for the relevant designated phenomenal properties rather than others in virtue of the fact that they are the ones with which she is phenomenally acquainted.

Now let us consider what happens when Mary either leaves the Nida-Rümelin Room, or uses her advanced physical-theoretical knowledge to determine which color patch is which color, what the viewing conditions are, and what physiological states and processes she is undergoing when viewing them. One thing she will learn is what it is like for her to have such and such experiences while she is viewing such and such color patches, with such and such reflectance profiles, in such and such conditions, while she undergoes such and such corresponding neurobiological states and processes. Using this piece of knowledge, she will be in a position to infer that similarly situated and endowed individuals will have the same type of experience while viewing color patches of the same reflectance profiles, at least so long as they undergo the same types of neurobiological states and processes. She can even begin to make hypotheses about what sorts of experiences individuals with abnormal vision might experience while viewing similar color patches in similar conditions, or what sorts of experiences individuals with normal vision might have in very different circumstances and/or when viewing different objects and properties.

However, such hypotheses will be quite tentative. Indeed, because there is an epistemic gap between her newly acquired genuine phenomenal concepts and her

---

297 Presumably, Mary could also use predictive knowledge she formed in the Jackson Room to the same effect.
antecedently-possessed physical-theoretical concepts, she will never be able to know with certainty that individuals invariantly have such and such experiences given such and such physical circumstances. In fact, Mary will be unable to rule out a priori the possibility that others have very different experiences from those she has in one and the same narrow and wide physical circumstances. Nor can she a priori rule out the possibility that they lack experiences altogether. For all that she can deduce a priori, she is the only conscious subject in a world populated by phenomenal zombies.

On the other hand, she has no reason to discount the possibility that her newly acquired phenomenal knowledge is new factual knowledge about physical facts she already knew in the Jackson Room by means of her physical-theoretical concepts. She cannot know a priori whether her first experiences of chromatic color acquainted her with new non-physical phenomena, or whether she instead became acquainted with “old” physical phenomena in a new way. Hence, she cannot know a priori whether her newly acquired conceptions of conscious experiences as subjective and qualitative are conceptions of newly experienced non-physical phenomena, or whether they are new conceptions of “old” physical phenomena that were not already provided by her physical-theoretical concepts.

If she simply reflects on how radically different the conceptions of reality provided by her genuine phenomenal concepts and her physical-theoretical concepts are, she will likely find the intuition that her two kinds of concepts refer to distinct phenomena quite powerful. But her intuitions of distinctness will be no sure guide to the metaphysical issue of the numerical identity or distinctness of conscious experience and some aspect of physical reality. This is because we can account for all of the changes in
Mary’s epistemic situation with respect to chromatic color experiences in terms of what Perry (2001a, 2001b) calls “reflexive content” without introducing any new “subject matter content”. That is, these changes can all be accounted for by explaining changes in Mary’s representational system (and the contextual relations or roles the newly designated phenomenal properties stand in relative to it) while abstracting away from what exactly her physical-theoretical concepts and genuine phenomenal concepts refer to.

Moreover, Mary might have powerful reasons to distrust her intuitions of the distinctness of the experiential from the physical. For one thing, she might simply find dualism antiquated and utterly mysterious. She might be attracted to physicalism due to its promise of explanatory and ontological simplicity and/or unity. She might be convinced (as I am) that some form of physicalism must be true as the result of causal considerations. Whatever the case might be, if Mary is an antecedent physicalist of some form or another, the fact that she acquires quite substantial factual knowledge about chromatic color experience upon leaving the Jackson Room gives her absolutely no reason to conclude that the experiential is distinct from the physical. On the contrary, she would have good reason to suppose that this new substantial factual knowledge is knowledge of physical facts she already knew in a different way, by means of very different concepts, and ultimately by being acquainted with reality in a very different way.

298 See §2 of Chapter 3 of this dissertation.
299 Some dualists will be unconvinced by this line of argument on the alleged grounds that a complete and ideal physical theory will provide an exhaustive and revelatory conception of physical reality such that a subject could a priori deduce any physical fact from it. In my opinion, reflection on the nature and aim of physical theorizing should cast significant doubt on the presumption that even a complete and ideal physical theory will provide an exhaustive and revelatory conception of physical reality. I am also skeptical of the two-dimensional semantic theory this dualist objection seemingly presupposes. But it is worth noting that a physicalist could accept both of the former assumptions and still deny the dualist conclusions by
§5  THE MODAL ARGUMENT

Kripke’s modal argument begins with his well-known insight that if an identity is expressed with rigid-designators, the identity is either necessarily true, or necessarily false. This is because, by definition, a rigid-designator picks out one and the same thing in every possible world (or at least in every possible world in which it exists); so if two rigid-designators pick out one and the same thing in any particular world, they will both pick it out in every possible world. And if they pick out different things in any particular world, they pick out different things in every possible world. Thus, there are no contingent identities (at least when rigid-designators are involved). But, Kripke quite plausibly continues, whether an identity is analytic or knowable a priori is a separate matter. Some identities, though necessary, must be discovered through empirical investigation. In such cases of “a posteriori identity”, it is typically a priori conceivable that the putative identities don’t hold, even though it is necessary that they do.

Many physicalists, including proponents of the phenomenal concepts strategy, have found Kripke’s notion of a posteriori identities quite attractive for explaining why there is an epistemic gap between the experiential and the physical, but no corresponding ontological gap. However, Kripke himself argues that his notion of a posteriori identity provides no help to those inclined to identify the experiential with the physical. The problem is, he argues, that in many cases of a posteriori identity, there is a sense that the relation between the identified phenomena under question is contingent—a sense which must be explained away on pain of our being forced to conclude that there is merely a contingent relation between them. But in the specific case of the experiential and the

holding that the conceptions provided by our phenomenal concepts are deeply inaccurate and fit nothing in our world. See Jackson (2004), Beisecker (2010), and Pereboom (2009, 2011).
physical, Kripke contends, we cannot explain away the sense that they are merely contingently related in the same way we do in other cases of a posteriori identity.

According to Kripke, the misleading sense of contingency concerning empirically discovered necessary identities comes from the real contingency between the relevant phenomenon and the way we pick it out using one or more of the co-designating concepts. For instance, he argues, although it is necessary that water is H₂O or that heat is mean molecular energy, our methods of picking out the respective phenomena (e.g. by their causing characteristic sensations in us in characteristic circumstances) are only contingently related to them. That is, if the actual world worked differently, we might pick out water or heat by means of very different sensations, or by other means entirely. Kripke contends that when we think we are imagining that water is only contingently related to H₂O or that heat is only contingently related to mean molecular energy, what we actually imagine are circumstances in which our characteristic sensations of water or heat pick out different things. We just confuse these genuine possibilities with the related impossibilities of water not being H₂O or of heat not being mean molecular energy. And, Kripke maintains, modal confusions of just this sort are generally what explain the sense of contingency concerning a posteriori necessities.³⁰⁰

However, Kripke continues, we cannot rely on the same strategy for explaining why the experiential and the physical at least seem to be contingently related. For one thing, we are not aware of our own experiential states by means of their causing in us characteristic sensations in characteristic circumstances. For another, he claims, the standard strategy is to argue that “although the [identity] statement itself is necessary,

³⁰⁰ Kripke does not, however, provide “any general paradigm for the appropriate corresponding qualitative contingent statement”. See page 143.
someone could, *qualitatively* speaking, be in the same epistemic situation as the original, and in such a situation a qualitatively analogous statement could be false". However, Kripke argues, there is no such possibility in the case of the alleged identity of the experiential and the physical.  

This is because, according to Kripke, “to be in the same epistemic situation that would obtain if one had a pain *is* to have a pain; to be in the same epistemic situation that would obtain in the absence of pain *is not* to have pain”. In other words, Kripke takes it that anyone who is qualitatively speaking in the same epistemic situation that one is in when he or she has a pain *just is* for him or her to have a pain, and vice versa (and the same holds with due modifications for other conscious experiences). So there is no real contingency between our epistemic situation with respect to the qualitative character of experiences and the experiences themselves by means of which we can account for the appearance of contingency in psycho-physical identities.

Kripke also thinks that roughly the same point can be made by reflecting on the differences in how we pick out pains and how we pick out heat or water. In the case of heat and water, he argues, we pick them out by their “accidental” properties of causing in us characteristic sensations in characteristic circumstances. But in the case of pain (and other conscious experiences), he contends, “[it] is not picked out by one of its accidental properties; rather it is picked out by the property of being pain itself, by its immediate phenomenological quality”. Indeed, he continues, “If any phenomenon is picked out in

---

301 P. 150.
302 P. 152.
303 P. 152.
304 P. 152.
exactly the same way that we pick out pain, then that phenomenon is pain”. So once again, Kripke concludes, there is no room for any contingency between our epistemic situation with respect to the qualitative character of our experiences and the corresponding experiences themselves which we can use to explain away the appearance of contingency between the experiential and the physical.

There is unsurprisingly an enormous body of philosophical literature which aims at finding some alternative way of explaining away the appearance of contingency between the experiential and the physical, and there are many illuminating insights in it. In my opinion, the most promising strategy along these lines is offered by Grover Maxwell in his 1978 “Rigid-Designators and Mind-Brain Identity”. In this underappreciated paper, Maxwell argues that Kripke’s mistake is to try to locate the physicalist’s needed contingency on the experiential side of the relevant psycho-physical identities. Instead, he contends, the physicalist should explain away the appearance of contingency by appealing to the contingency between our epistemic situation with respect to the physical and the physical itself. On his view, it is physical phenomena that we pick out by means of their accidental properties, by the spatio-temporal and causal-dispositional roles they occupy. Moreover, it is in the case of the physical that we can be in qualitatively the same epistemic situation as in our own case, and in such a situation a qualitatively analogous statement involving our physical theoretical concepts could be false (because different phenomena would occupy the spatio-temporal and causal

305 P. 153.
306 I will only touch on a small sampling of this literature here.
307 Also see Perry (2001a), Chalmers (2002), Papineau (2002), and Stoljar (2001a, 2001b, and 2006).
dispositional roles that the physical does in the actual world). Nevertheless, in the actual world, our phenomenal concepts and our physical-theoretical concepts rigidly pick out one and the same phenomenon, and so they pick it out in every possible world in which it exists. Thus, the relevant psycho-physical identities are necessary, despite any appearances of contingency to the contrary.

I am inclined to think that Maxwell is right in identifying a real contingency between the structural conceptions of the physical provided by our physical-theoretical concepts and the nature of the physical itself. (However, unlike Maxwell, I deny that our structural conceptions of the physical in general play a reference-fixing role in picking out the physical.) But I also think that most physicalists, including Maxwell, already grant Kripke too much. For in addition to locating contingency on the physical side of the relevant psycho-physical identities, physicalists should also allow for contingency on the experiential side of the identities, despite Kripke’s contentions to the contrary.

One reason for attempting to locate contingency on the side of the experiential is that if we accept Kripke’s thesis that what strikes us as such-and-such experience is such-and-such experience (and vice versa), it becomes difficult (and perhaps insurmountably so) for the physicalist to explain how our knowledge of the experiential could have this epistemic intimacy. For as I noted in the previous chapter, Kripke’s thesis demands that the physicalist ground the epistemic relation between our experiences and our thought about them in a metaphysically intimate relation that would rule out certain kinds of introspective error, and that is compatible with the ontological commitments of physicalism. But it is hard to see how the physicalist can successfully take on this

---

This accords with my general stance in holding that our conceptions are typically reference-guided, as opposed to our reference being typically conception-guided. Papineau (in conversation) is inclined to side with Maxwell.
explanatory burden, particularly since, as I’ve already argued, the most obvious candidate relation, *constitution*, is not an epistemically-enabling relation.\(^{310}\)

Another reason why a physicalist might look to locate contingency on the experiential side of the relevant psycho-physical identities is that some will be disinclined to allow for real contingency between the physical and our spatio-temporal and causal-dispositional role-based conceptions of the physical. In particular, physicalists who follow Shoemaker (1980) in holding that spatio-temporal and causal-dispositional profiles are individuative of the physical, will see no room (or precious little) for any contingency between physical phenomena and the structural roles they happen to occupy in the actual world. And even if one rejects Shoemaker’s view, one might still worry that Maxwell’s proposal commits the physicalist to the existence of novel (and arguably unexpected) contingent laws of nature that link the physical phenomena in this world with the spatio-temporal and causal-dispositional roles they happen to occupy in this world, laws which could be very different in different possible worlds. In any case, even if we suppose that these worries can be answered satisfactorily (and I suspect that they can be), we might still find it somewhat surprising that solving Kripke’s puzzle requires that we take a determinate stand on these seemingly independent issues. Hence, it is worth exploring the possibility of contingency of the experiential side of the psycho-physical identities.

I think that there are compelling reasons for thinking that there *is* in fact real contingency on the experiential side of the psycho-physical identities. In particular, the phenomenal continua cases (and other cases of introspective error) we have already considered several times in this dissertation show that our phenomenal judgments can

\(^{310}\) See §5 of the previous chapter.
come apart from the actual subjective, qualitative characters of our experiences.\textsuperscript{311} (We can, for instance, judge that we are having an experience of kind \( A \) when we are in fact having an experience of kind \( B \) if the difference between the characters of \( A \) and \( B \) are too fine-grained for us to discriminate introspectively.) In turn, this shows that there is in fact contingency between our phenomenal judgments and the experiences they are about. So Kripke is wrong in thinking that any phenomena we pick out as such-and-such experience \textit{just is} such-and-such experience, and vice versa.

Even more to the point, Russell taught us that having (and thereby being acquainted with) such-and-such experience is one thing, but knowing \textit{that} it is such-and-such experience (or physical phenomena) is quite another. And while we might be in a position to acquire relatively reliable, partial knowledge of some limited aspects of our occurrent experience, our special first-personal ways of being acquainted with the experiential do not provide us with infallible, exhaustive, or revelatory knowledge of its essential nature.\textsuperscript{312} However, Kripke’s modal argument seemingly requires the outright denial of Russell’s crucial insights here. On this issue, I believe that the physicalist should follow Russell rather than Kripke.

In fact, I think that the intuitive force backing Kripke’s claims about our epistemic situation with respect to the subjective, qualitative characters of our experiences is the result of an ambiguity in Kripke’s notion of “being in the same epistemic situation, qualitatively speaking”. In my view, we must be careful to distinguish between two distinct elements which are present in what Kripke labels as our “epistemic situation”.\textsuperscript{313} First, there are the subjective, qualitative characters of our experiential states (together

\textsuperscript{311} See §3 of Chapter 1 of this dissertation and §5 of the previous chapter.
\textsuperscript{312} Schwitzgebel (2008) considers additional examples of the unreliability of our introspective capacities.
\textsuperscript{313} Sellars (1956) deserves a lion’s share of credit for drawing attention to the distinction.
with how such subjective, qualitative experiences “intentionally” present the world as being). And second, there are our—sometimes hedged—judgments (or dispositions to so judge) about how things are with respect to the subjective, qualitative characters of our experiences or with respect to the wider world with which our experiences put us into epistemic contact.

Following Schwitzgebel (2008), we might call the first aspect of what Kripke labels our “epistemic situation” the “phenomenal sense” of how things “look”, “seem”, or “appear” to be for one, and we might call the second aspect the “epistemic sense” of how things “look”, “seem”, or “appear” to be for one. I believe that the intuitive force of Kripke’s modal argument results from the fact that it is part (but only part) of the essential nature of pains that they seem painful to the subjects who have them in the phenomenal sense of “seeming”. In other words, Kripke is intuitively correct that it is an essential feature of instances of pains that there is a specific “something it is like” for the subjects who have them to have them. But it simply does not follow that whenever it seems to us that we are in pain (in the epistemic sense of “seeming”), we are in pain; nor that whenever we are in pain it will seem to us that we are (again, in the epistemic sense of “seeming”). As we’ve seen, it is quite the contrary. Moreover, in making judgments (including modal ones) about our experiences, we may only legitimately hold fixed our epistemic situation in the epistemic sense of how things seem to us to be. But then there will be contingency between our epistemic “take” on the subjective, qualitative characters of our experiences and how things actually are with respect to our conscious experiences. And so the physicalist has plenty of additional material for explaining why psycho-physical identities, though necessary if true, at least seem to be contingent.

Pp. 262-3.
There are, however, some who will find this entire line of argument unconvincing. For instance, Stoljar (2006) has recently argued that the attempt to locate contingency between our phenomenal judgments and the experiences they are about more-or-less leads one to contentious “perceptual or observational models of self-knowledge or introspection”. But, he thinks, it is not plausible that such “psychological intermediaries” are needed for our introspective knowledge of our experiential states.

There is a sense in which Stoljar is right here, and a sense in which he is clearly wrong. What he is right about is that our fundamental, non-reflective phenomenal awareness of our own experiences does not require any psychological intermediaries; we are acquainted with the subjective, qualitative characters of our experiences simply by having them. But he gives us no reason to give up the plausible view that our reflective judgments about our experiences are in an important sense psychologically mediated.

Moreover, holding that our reflective judgments about our experiences are psychologically mediated is not tantamount to endorsing a perceptual or observational model of introspection. All that is required is that the vehicle of our phenomenal judgment, and our epistemic “take” it is based on, are distinct from the represented experience. What’s more, adopting this alternative does not force us to hold that the relation between experiences and our judgments about them are merely causal in character. For we can perfectly well hold that our phenomenal judgments are in fact psychologically mediated by direct informational relations.

We also needn’t worry about Stoljar’s (2006) recent objections to Williamson’s (2000) slightly different use of phenomenal continua cases for putting pressure on the alleged “self-intimating” character of experiences: that if we have an experience and

---

315 P. 184.
consider whether we are having one, we cannot fail to know that we are.\textsuperscript{316} According to Stoljar, Williamson’s argument is basically that “there might be cases in which a particular experience is so faint that you cannot reliably tell whether you have it or not. If so, although it is true that you are having the experience, it is not true that you are in a position to know that you are”.\textsuperscript{317} But, Stoljar argues, this does not significantly undermine Kripke’s modal argument since we can simply reformulate it in terms of “extreme” (or vivid) experiences. In other words, we can say that if we have a vivid experience and consider whether we are having one, we cannot fail to know that we are. So, Stoljar continues, at least in some cases (i.e. the vivid ones) we can have knowledge of our experiences without the need for any psychological intermediaries. But if we can have psychologically unmediated knowledge of our experiences in some cases, Stoljar reasons, what reason would we have for then supposing that our introspective knowledge is ever psychologically mediated? And even more to the point, why can’t we simply argue that if we pick out anything in the same way that we pick out such-and-such vivid experience then it \textit{is} such-and-such vivid experience, and vice versa? So we can, Stoljar contends, simply run Kripke’s modal argument using vivid experiences and get around Williamson’s challenge.

There are a number of problems with Stoljar’s argument here, including his running together of the thesis that vivid experiences are self-intimating with the thesis that our introspective knowledge of \textit{their occurrence} is psychologically unmediated, and of the latter with Kripke’s claim that what strikes us as such-and-such experience \textit{is} such-

\textsuperscript{316} Stoljar (2006), pp. 184-5. Note that Stoljar himself uses the term “self-presenting” rather than the term “self-intimating” to designate this alleged epistemic feature of experiences. I prefer my own terminology here.

\textsuperscript{317} P. 185.
and-such experience (and vice versa). But even if we grant that Stoljar’s objections to Williamson are satisfactory, they do not call into question my Russell-inspired use of phenomenal continua cases against Kripke’s modal argument. For even in cases in which $A$ and $B$ are both vivid experiences, we cannot know on the basis of our experience alone whether they are instances of the same kind of experience, or of two different kinds. Knowledge of identity and diversity is knowledge of truths about a thing, and is logically distinct from our more fundamental knowledge of them by acquaintance. And there is always the possibility of error in moving from our knowledge of a thing by acquaintance to knowledge of truths about it. Hence, it is not the case that whenever we pick out an experience in the way we pick out experiences of such-and-such kind, it is of that kind (and vice versa). And restricting our attention to vivid experiences alone does nothing to undermine this crucial point.

There is however a possible line of objection to my arguments that does have some degree of merit: that I am already simply granting Kripke too much. Indeed, there is a very good case to be made that the lingering “sense of contingency” concerning some a posteriori identities is a purely psychological matter, rather than a metaphysical one. And if this is correct, then explaining it away does not evidently require that the physicalist identify some real contingency with which we are confusing it. Rather, the physicalist need only explain why the relevant identities seem to us contingent in a purely epistemic sense of “seem”. And, as Papineau has recently pointed out, there are many possible explanations about why we would mistakenly take there to be contingency where there is in fact none, including cultural-historical ones, purely psychological ones, phylogenetic or ontogenic developmental ones, etc. In fact, I more-or-less agree that I

---

318 Papineau has recently taken up this objection to Kripke’s modal argument.
needn’t grant Kripke as much as I do. Still, I think it is worthwhile to demonstrate that
the key epistemic premise on which Kripke rests his argument, a premise which far too
many physicalists readily accept, is seemingly false. Moreover, once we see that Kripke’s
crucial premise is false, we have more powerful resources for defending the claim that
the lingering “sense of contingency” is a purely psychological matter.

Hence, I think that locating a source of contingency on the experiential side of the
relevant psycho-physical identities is only one important step in defending physicalism
against Kripke’s modal argument. But all by itself, it does not provide us with sufficient
materials to explain fully our intuitive feeling that we could be in such-and-such
neurobiological state without having such-and-such experience, and vice versa. Even
still, it does is teach us an important lesson about our pre-theoretical modal intuitions
regarding psycho-physical identities: that they aren’t particularly reliable.

Indeed, prior to reflecting carefully on phenomenal continua cases (and other
cases of introspective error), we have the strong intuition that there is simply no room for
contingency on the experiential side of psycho-physical identities. Similarly, prior to
Maxwell-style (i.e. Russellian) reflections on the relationship between physical theories
and the physical phenomena they represent, we have fairly robust intuitions that the
essential nature of the physical is revealed by an ideal physical explanation. But if
Maxwell and/or I are correct, then our pre-theoretical intuitions are simply wrong about
the modal status of the relations between our ways of picking out the relevant phenomena
and the phenomena themselves. Likewise, the causal arguments in favor of physicalism
give us good grounds at least to question whether there are necessary relations where we
pre-theoretically take there to be merely contingent ones. So we have strong reason to distrust our modal intuitions here.

This does not mean that we should simply ignore our intuitions of contingency wholesale, however. After all, we still have to make sense of our psychologies when we are making these likely errors about the modal status of the relevant relations. Nevertheless, once we have shifted our focus to what's going on in our epistemic lives when we make these modal judgments rather than looking for real contingencies in the world, Kripke’s modal argument loses its bite.

In fact, I think that our intuitive “sense of contingency” isn’t doing any real work in generating our intuitions of the distinctness of the experiential from the physical. On the contrary, I think that, psychologically speaking, it is our intuitions of distinctness that are responsible for our sense of contingency about psycho-physical identities. In other words, I think that our sense of contingency is the result of the special epistemic gap between our conceptions of the experiential and the physical. They seem to be contingently related because we can imagine zombies, inverted spectra, etc., and because there is a distinctive explanatory gap.

Furthermore, if I am right, the special epistemic gap between the experiential and the physical is the result of the radical differences between our phenomenal concepts and our physical theoretical concepts and the conceptions of reality that they provide. When we reflect on how different the conceptions of reality provided by these distinct kinds of concepts are, and how there is no conceptual link between them, we will be inclined to wonder about how they could possibly be of the same thing, even if we have good reason to think that they are. And this explains why many physicalists still feel the sense of
contingency, even though they are convinced that the relevant psycho-physical relations are actually necessary.

So basically, I think that we need to focus on the psychological phenomenon of our sense of contingency, and when we do so, we see that Kripke’s modal argument stands or falls with other dualist arguments such as the zombie argument and the explanatory gap argument. I simply don’t think that the antecedent physicalist owes any additional explanation of the sense of contingency concerning psycho-physical identities. Nor must he or she explain how we can stand in an epistemic relation to the experiential such that how things seem to be is how they are (and vice versa). We simply don’t stand in any such relation, either to the physical, or to the experiential. But if the dualist still demands that the antecedent physicalist locate some real contingency, we’ve seen that there is plenty to offer—on both sides of the relevant psycho-physical identities.

§6 THE EXPLANATORY GAP

In recent years, many have pointed out that there is a deep explanatory gap between the experiential and the physical which is seemingly not to be found elsewhere in the physical domain. In most cases of a posteriori theoretical identities, the argument runs, once the identification is successfully carried out, there is no longer any sensible question about how the identified phenomena could possibly be one and the same. For instance, once we have identified water with H₂O or heat with mean kinetic energy, there is no further question about why we find water whenever we find H₂O or find heat whenever we find mean kinetic energy. However, in the case of putative psycho-physical identities, there does at least appear to be a further sensible question about how such-and-such an
experience could possibly be such-and-such a physical phenomenon. Seemingly, even after we have otherwise compelling grounds for identifying the experiential with some part of the physical domain, there is an “open question” about why “what it’s like” to be in such-and-such physical state is like this rather than like that, or like anything at all. This suggests for many that our physical explanations of reality simply leave out the subjective, qualitative characters of our experiences. And one straightforward explanation for why physical theory leaves out the subjective, qualitative characters of the experiential is that they simply aren’t physical in nature.

Though Levine (1983, 2001) is most often associated with the explanatory gap argument, he does not himself draw any ontological conclusions about the nature of consciousness from the fact that it seemingly eludes structural-functional characterization and hence cannot be explained fully by means of the standard explanatory models of the physical sciences. However, a number of recent dualists have used Levine’s contention that there is a distinctive explanatory gap between the experiential and the physical to construct arguments that do have ontological import. For instance, Chalmers (2002, 2010) makes the following case against physicalism: The physical domain is exhausted by what can be fully characterized by means of a complete and ideal physical theory, and by their very nature, physical theories explain at most how the physical is distributed in space and time (or space-time) and how the physical behaves in various circumstances. But we cannot explain the subjective, qualitative characters of conscious experiences in terms of such structural-functional characterizations of the world. Thus, the subjective, qualitative characters of our experiences are distinct from the physical domain.
Naturally, physicalists have objected to arguments like Chalmers’ on a number of different grounds. One immediate worry that many have is that Chalmers is asking of physicalism something that doesn’t even make sense.319 After all, if the experiential is identical with some part of the physical domain, then asking for an explanation of the identity seems wrongheaded from the start. Identities just aren’t the sort of things which require, or can be given, explanations. If we want to know why Mark Twain is identical with Samuel Clemens, for instance, all we can say is that “they” are one and the same person. So, many have contended, there is simply no room for any explanatory gap here, let alone a deep one.

On the other hand, Papineau (2002) has argued, there is a sensible question to be asked when we make standard theoretical “identifications” in the physical sciences. True, explaining why “two” things are one and the same thing makes no more sense in the standard physical cases than trying to explain psycho-physical identities, or the identity of Mark Twain and Samuel Clemens. But we are not really making genuine identity claims in standard cases of theoretical identification. Rather, what we are doing is explaining how such-and-such physical phenomenon plays the structural-functional role associated with our pre-theoretical concept of some natural phenomenon or with a different theoretical concept of the same thing.320

According to Papineau, this is possible precisely because our theoretical concepts and our pre-theoretical concepts of natural phenomena both have associated descriptive materials which characterize the spatio-temporal and causal-dispositional roles that their purported referents occupy (even if these descriptive materials play no reference-fixing

---

319 Examples include Block and Stalnaker (1999) and Papineau (2002).
320 Papineau (2002), Chapter 5.
role). But, he argues, there are no role-characterizing descriptive materials associated with our genuine phenomenal concepts; they simply pick out the subjective, qualitative characters of our experiences directly by exemplifying them in our phenomenal concepts themselves. And since our phenomenal concepts do not have associated role-characterizing descriptive materials, we cannot (and need not) explain how they occupy such-and-such spatio-temporal causal-dispositional roles in the way we do in standard cases. All we can (and need) do is identify them with such-and-such physical phenomena and say how this physical phenomena occupies such and-such spatio-temporal causal-dispositional role.

I think there is something right about this line of argument. It does seem senseless to demand an explanation for the identity of “two” phenomena. On the other hand, I don’t think that these objections concerning the explanatory potential of identity really get to the heart of the issue. What the dualist (or Levine) is after is not an explanation of why the relation of identity holds or doesn’t hold of “several” things. Rather, what they are after is an explanation of how the distinct conceptions associated with our concepts could possibly be conceptions of the same thing. They think that in the case of standard identities, we have explanatory models that make intelligible how our distinct conceptions can be of one and the same physical phenomenon. But in the case of psychophysical identities, they argue, we do not have explanatory models that make intelligible any non-contingent link between our conceptions of the experiential and the physical. If we then take the further step of thinking that our distinct conceptions are accurately getting at aspects of the world (as Chalmers does), then the explanatory gap between our conceptions of the experiential and the physical starts to suggest that there is a

321 Papineau (2002), Chapter 5.
corresponding ontological gap between them. For this reason, Chalmers (but not Levine) concludes that our conceptions must not be of one and the same aspects of the world—the experiential must therefore be distinct from the physical.

Furthermore, I think that Papineau is mistaken about our phenomenal concepts not having any associated role-characterizing descriptive materials (though his is absolutely right that they don’t play any reference-fixing role, and don’t ordinarily figure into the content of our phenomenal thought and talk). On the contrary, our phenomenal concepts quite obviously have associated conceptions of the causal-dispositional roles that our experiences play in our mental and behavioral lives. That is an important reason why phenomenal concepts are useful in the first place; for without such causal-dispositional conceptions, we could not use them to explain and predict the behavior of ourselves or others. Moreover, if we flat-out deny that our phenomenal concepts have any associated conceptions of the causal-dispositional roles that our experiences play, we undercut one of the principal grounds for embracing physicalism in the first place: the causal argument for psycho-physical identities. If we are to take issue with Chalmers’ argument from the explanatory gap to an ontological gap, we cannot do so by simply denying that one or the other of our concepts has any associated role-characterizing descriptive materials.

One possible way of going at this point is to contend that either one or both of the conceptions associated with our concepts is inaccurate. For example, Pereboom (2009, 2011) has recently contended that arguments like Chalmers’ are unsound because they

---

322 I say “ordinarily” here because we always can use our associated descriptive conceptions of phenomenal experience to designate things indirectly, even if we don’t typically do so.
323 Even most epiphenomenalists will grant this much; they will simply say that our causal conceptions are erroneous.
fail to take into account the possibility that our phenomenal concepts falsely represent our experiences as having subjective, qualitative characters which they in fact lack. And if our phenomenal concepts turn out to systematically misrepresent our experiences in this way, he plausibly argues, it is no mark against the truth of physicalism that we don’t have explanatory models that make intelligible any non-contingent link between our conceptions of the experiential and the physical. Of course, it is a daunting task to explain satisfactorily how, and why, our phenomenal concepts systematically misrepresent our experiences as having subjective, qualitative characters which they in fact lack. But despite its sheer implausibility and counter-intuitiveness, I do think that it is one potential way of defending physicalism against Chalmers’ explanatory gap argument.

In my estimation, a more promising strategy is to focus attention on the incompleteness and difference in character of the conceptions provided by of our phenomenal concepts and physical-theoretical concepts rather than alleging that one or the other is systematically inaccurate. Physical-theoretical concepts provide us with conceptions of what things there are in the physical domain and what spatio-temporal and causal-dispositional roles these things occupy. Phenomenal concepts provide us with conceptions of the causal-dispositional roles of our experiences and of their subjective, qualitative characters. While there is certainly some overlap between these two kinds of conceptions, there are also radical differences in them where they do not overlap. Dualists see the differences in the non-overlapping portions of our conceptions as characterizing different aspects of the world (either different things or different features). But that simply does not follow. Indeed, there is always the alternative possibility that the

\[324\] Jackson (2004) and Beisecker (2010) advance similar arguments.
non-overlapping portions of our conceptions are characterizing *one and the same aspects of the world differently* (whether the aspects concern things or features).

Part of what is driving dualists to ignore this alternative is the (sometimes implicit) assumption that a complete and ideal physical theory will exhaustively *reveal* the nature of the physical.\(^\text{325}\) In fact, Chalmers makes explicit appeal to this assumption as one of the premises for his explanatory gap argument against physicalism. But the question is this: Is there any compelling reason why the antecedent physicalist should assume this? Does one’s commitment to physicalism ipso facto commit one to the view that a complete and ideal physical theory will exhaustively reveal the nature of the physical?

I am inclined to think not. One reason for being skeptical of wedding together the ontological view that everything (concrete) is physical to the view that a complete and ideal physical theory would exhaustively reveal the nature of the physical is “Hempel’s Dilemma”.\(^\text{326}\) In a nutshell, Hempel’s (1980) well-known dilemma for defining ontological physicalism in terms of physical theory is this: If we rely on what our current best physical theories tell us about the world, then physicalism will surely turn out to be false. After all, not only does the history of scientific theorizing suggest that our current best physical theories will eventually be replaced by others, but we *know* that our current best physical theories are radically incomplete (e.g. we have little understanding of dark matter) and seemingly internally inconsistent (e.g. quantum theories and general relativity seem to be in conflict). But if we instead define ontological physicalism in terms of a

---

\(^\text{325}\) And they hold the same view about phenomenal concepts and the experiential, one which phenomenal continua cases give us good cause to reject.

\(^\text{326}\) Levine (2001), pp. 18-21 provides a good introductory discussion of Hempel’s Dilemma.
complete and ideal physical theory, then we run into a number of serious worries. First, some worry that the ontological thesis of physicalism would then lack determinate content. After all, the worry goes, if we don’t know what will be included in the hypothesized complete and ideal physical theory, how can we evaluate whether or not it is true? Second, some worry that physicalism will turn out to be trivially true; for if the “physical” just is what a complete and ideal physical theory says is the case, then physicalism seemingly will come out true no matter what course complete and ideal scientific theorizing ends up taking. And third, some worry that in the attempt to explain the workings of the world, ideal physical theorizing could turn out to appeal to objects and features which are intuitively physically unacceptable (such as immaterial souls, conscious particles, emergent mental forces, etc.).

Unsurprisingly, there is a great deal of literature on how the physicalist might avoid Hempel’s Dilemma, which I cannot hope to survey here. But one thing shared by the majority of the attempts to evade it is the (often partial) divorcing of ontological physicalism from physical theory, either by placing additional constraints on ontological physicalism or by characterizing it in some entirely different fashion. Either way, avoiding Hempel’s Dilemma seems to require a rejection of the view that the physical just is what is exhaustively revealed by a complete and ideal physical theory.

Another reason for questioning the dualist assumption that a complete and ideal physical theory would exhaustively reveal the nature of the physical is what Alter and Howell (2009) have recently called “the problem of emptiness”. The problem of

---

327 Wilson (2006) gives a helpful discussion of these worries.
329 Wilson (2006) dubs this the “inappropriate extension worry”.
emptiness seemingly arises due to the general character of current physical theorizing, which we have seen at several different points in this dissertation. We have seen, for instance, that Chalmers thinks that it is the nature of physical theory that it explains at most the structure and dynamics of concrete reality (where “structure” characterizes how the various things in reality relate to one another and “dynamics” characterizes how those relations change over time). Similarly, we’ve seen that both Maxwell (1978) and Papineau (2002) hold that current physical theories characterize the physical purely in terms of their spatio-temporal and causal-dispositional roles. Finally, my own account of our (current) physical-theoretical concepts has it that they provide us with conceptions of what things there are in the world and what role they play in its workings.

What all of these (ultimately Russell-inspired) views have in common is the belief that at its core, current physical theorizing the essentially role-based or relational in character. It tells us what things there are, how they are related to one another in space and time (or space-time), and what kinds of causal-dispositional powers they have. But this, the worry goes, provides us with a merely formal conception of the world that is empty of concrete “content”. Intuitively, there must be more to the nature of the physical than what current physical theory tells us about. Intuitively, there must be something occupying (or “filling in”) the roles; there must be relata standing in the various relations. What’s more, there must be something in the nature of the physical role-bearers such that they have the causal-dispositional powers that they do given how they are distributed in space and time (or space-time) and related to one another. In other

331 Alter and Howell (2009), p. 90.
333 Alter and Howell (2009), p. 90.
words, the physical phenomena in this world must have some underlying non-relational “categorical nature” that (at least in part) grounds their causal-dispositional powers. The idea that the physical domain simply consists in “brute”, ungrounded dispositions all the way down (and all the way up) simply leaves us with a strikingly insubstantial (and intuitively unsatisfying) picture of concrete reality.\footnote{However, there are some who are quite content with the view that concrete reality is composed entirely of such “brute dispositions” or “causal powers”.} This is more-or-less the core intuition behind the problem of emptiness.

There are a number of attitudes that an antecedent physicalist might take in response to the problem of emptiness. Some, such as Ney (2007), simply deny that all current physical theories leave out characterizations of the underlying categorical nature of physical phenomena (she thinks that “gauge theories” are one interesting exception). Others, such as McGinn (1989), accept that there is a deep problem of emptiness and contend that we are so cognitively constituted as to never hope to achieve an understanding of the underlying categorical nature of the physical. And still others, such as Stoljar (2006) and Strawson (1998, 1999), leave open the possibility that radical changes in the nature of our physical theorizing might open up for us the possibility of a scientific understanding to the underlying categorical nature of the physical.

I won’t here take a stand on which attitude is the right one to have on these issues. What is important for us here is simply this: If we assume along with Chalmers that physical theories can explain at most how the physical is distributed in space and time (or space-time) and how the physical behaves in various circumstances, the problem of emptiness should give us serious pause in thereafter following him in accepting that a complete and ideal physical theory would exhaustively reveal the nature of the physical.
For if it were indeed a problem, then ontological physicalism would be shown to be false long before, and completely independent of, whether or not the experiential is part of the physical domain. Physicalism would be falsified simply on the grounds that there would be more to concrete reality than physical theory could characterize structurally or dynamically, and thus more to concrete reality than physical reality. But this suggests that something has gone terribly wrong. After all, why should an ontological view concerning the mind-body problem be straightforwardly falsified on the basis of the seemingly independent issues about whether or not there are brute dispositions, and whether or not physical theories are essentially limited to causal-dispositional characterizations of concrete reality?

This is all simply to say that Chalmers’ assumption that ontological physicalism is committed to the view that a complete and ideal physical theory would exhaustively reveal the nature of the physical is seemingly false.\textsuperscript{336} The relationship between physical theory (and physical explanation) and physicalism is simply not that straightforward. So even if a complete and ideal physical theory would not provide us with explanatory models which make intelligible some non-contingent link between our conceptions of the experiential and our conceptions of the physical, that would not all by itself demonstrate that physicalism is false. Indeed, physicalism is compatible with an explanatory gap even at the limits of physical theorizing. And in a case like ours, where we know outright that our current physical theories are neither complete nor ideal, we should be even less concerned about the existence of an explanatory gap between the experiential and the physical.

\textsuperscript{336} And phenomenal continua cases (and other cases of introspective error) show us that we should draw corresponding conclusions about what our phenomenal concepts tell us about the experiential.
One might, however, worry that in saying this I am basically arguing that physicalism is compatible with the possibility of the brute emergence of the experiential from the physical, a view which typically is considered quite incompatible with the truth of physicalism (and rightly so). In my view, this worry is misguided and trades on an ambiguity between two different kinds of “brute emergence”. What is incompatible with physicalism is the ontologically brute emergence of the experiential from the physical. If there is nothing in the nature of the physical which makes it the case that conscious experience comes into the world as a natural metaphysical consequence, then it does seem like there is something more to concrete reality than physical reality, and so some kind of dualism is true.

On the other hand, physicalism is not incompatible with epistemically brute emergence. There is nothing in the thesis of physicalism itself which demands that in addition to there being something in the nature of the physical such that the experiential naturally emerges from it our physical theories must also provide us with transparent knowledge of how and why this happens. The antecedent physicalist needn’t (and shouldn’t) take on any epistemic obligation quite so demanding as that. So once again, the existence of an explanatory gap between the experiential and the physical, even a novel one, does not all by itself give us any compelling reason to reject physicalism. In fact, if I am correct that our phenomenal concepts and physical-theoretical concepts are cognitively isolated concepts which are based on radically different ways of being acquainted with one and the same objects and features, and which provide very different kinds of conceptions of them, we should expect there to be a novel explanatory gap between them.
The Two-Dimensional Argument

In the previous chapter, we saw that Chalmers holds that linguistic expressions and concepts have two distinct patterns of evaluation. They have a primary intension which characterizes how their reference or extension depends on how the world turns out to be. They also have a secondary intension which characterizes their reference or extension given the way the actual world is.

On Chalmers’ view, the primary intension of a concept captures its “epistemic character”; it specifies what the concept picks out were various epistemically possible circumstances to obtain. When we evaluate the semantic value of a concept using its primary intension, we are evaluating what it pick out in what Chalmers calls “epistemic scenarios” or “counteractual possibilities”. Counteractual possibilities are ways the world could be which on ideal rational reflection the subject either cannot rule out a priori or can positively imagine to be the case.

The secondary intension of a concept captures how it behaves across the space of metaphysical possibilities given its reference or extension in the actual world. When we evaluate the semantic value of a concept using its secondary intension, we are evaluating its reference or extension “subjunctively” across counterfactual possibilities. Counterfactual possibilities are, Chalmers argues, genuine metaphysical possibilities.

Chalmers thinks that this two-dimensional framework has several important features. For one thing, it preserves what he sees as a crucial link between conceivability and metaphysical possibility. Because the primary intension of a concept is constitutively tied to what on ideal rational reflection a subject cannot rule out a priori or can positively imagine.
imagine to be the case, Chalmers takes subjects to have a priori access to what it picks out across epistemic possibilities, at least as long as subjects take sufficient care to rule out any potential “a priori incoherence, contradiction, or unimaginability”. Thus, he contends, there is a constitutive connection between what on ideal rational reflection a subject can conceive to be the case and what might be the case, epistemically speaking.

Furthermore, since the space of metaphysical possibilities turns on what a concept’s reference or extension is in the actual world, there is a crucial link between its primary and secondary intensions. In particular, for many concepts, we can consider what they refer to counterfactually given how their reference varies across epistemically possible scenarios. What we end up with, Chalmers argues, is a priori knowledge of special semantic conditionals: we know that if the world turns out to be one where such-and-such is picked out by the primary intension of some concept of ours in the actual world, then that concept will pick out the same such-and-such across all metaphysically possible scenarios. Consequently, subjects have a priori access to what conceivably might have been the case had the actual world turned out to be such-and-such (centered) logically possible world. Of course, drawing definite conclusions about the metaphysical nature of the world still requires empirical knowledge about the actual world. But if we have enough empirical knowledge to determine whether such-and-such epistemically possible scenario is in fact an actual one, we can use our a priori accessible knowledge of what, if anything, a concept pick out in it to determine a priori what it thereby picks out across the space of metaphysical possibilities.

341 Descriptive and rigidified descriptive concepts require a somewhat different two-dimensional treatment.
Hence, Chalmers’ notions of “epistemic possibility” and “subjunctive (or metaphysical) possibility” are “ultimately grounded in what it is rationally coherent to suppose”. As such, Chalmers argues, his two-dimensional semantic analysis over logical space is “perfectly suited to analyze such rational and psychological matters as counterfactual thought, rational inference, and the contents of thought and language”, which he contends are “precisely the purposes for which possible worlds are needed in the first place”.

In addition, Chalmers argues that his two-dimensional framework has the explanatory resources to accommodate all of the key insights about the semantic behavior of our concepts across metaphysical possibilities which Kripke has brought to our attention. For instance, it can explain standard Kripkean a posteriori necessary identities as ones in which the concepts flanking the identity sign have the same secondary intensions but different primary intensions. In other words, their primary intensions will pick out different things across the space of epistemically possible circumstances, but given the way the actual world is, they will pick out the same thing across the space of metaphysically possible circumstances. On Chalmers’ view, this explains why it is a priori (primarily) conceivable that the designated things are only contingently related even though there is a metaphysically necessary relation between them.

On the other hand, he thinks that his two-dimensional framework shows why we cannot appeal to such standard Kripkean a posteriori identities to block the inference from the conceivability of zombies (or inverts) to their metaphysical possibility (and hence the falsity of physicalism). This is because the concepts used to express psycho-

---

physical identities plausibly have primary and secondary intensions that do not come apart across the space of epistemic possibilities and subjunctive possibilities (though we’ll see that he is less committed to this in the case of physical-theoretical concepts). For instance, he argues, the concept “hydrogen molecule” most plausibly rigidly picks out whatever plays the causal role of hydrogen molecules in any world considered as actual. This is because all there is to being a hydrogen molecule is to play the relevant causal-dispositional role. So given any appropriately described epistemic scenario, “hydrogen molecule” will pick out the same phenomenon (i.e. whatever plays the hydrogen role in that scenario). What’s more, given widely held views about rigid counterfactual behavior of natural kind terms, “hydrogen molecule” will also pick out the same phenomenon across all metaphysically possible worlds. Chalmers takes it that this lesson plausibly generalizes for many (and perhaps all) other physical-theoretical concepts.

Similarly, Chalmers claims that our “direct phenomenal concepts” (and other “pure phenomenal concepts”), which pick out phenomenal properties directly in terms of their intrinsic phenomenal natures, have the same primary and secondary intensions across modal space. For Chalmers, direct phenomenal concepts are special concepts we use to think about conscious experiences when (and only when) we have them and attend to them. One way in which they are special is that they are in part constituted by the very phenomenal properties for which they stand. Thus, there is an especially intimate manner in which direct phenomenal concepts pick out their referents; their referents are picked out by their being exemplified in the concept itself.

---

345 We should notice right away that Chalmers’ account of direct phenomenal concepts is very close to physicalist constitutional accounts of phenomenal concepts.
They are also special in that there is an especially intimate manner in which we are aware of phenomenal properties when we think about them using direct phenomenal concepts. On Chalmers’ view, we are acquainted with them in a way that provides us with a substantive grasp of their essential natures. Together, the intimate constitution of direct phenomenal concepts and the intimate epistemic relation that makes our acquisition of them possible lends itself to a special sort of incorrigibility in our phenomenal beliefs about our occurrent conscious experiences (or at least so long as only one act of attention is involved). In particular, we cannot be mistaken in forming beliefs of the form “this experience is such-and-such experience”, where the such-and-such experience is thought of under a direct phenomenal concept.\(^{346}\)

Chalmers takes it that these special features of our direct phenomenal concepts (and our derivative standing pure phenomenal concepts) are what make them have same primary and secondary intensions. Indeed, he argues, given these special features, direct phenomenal concepts allow us to pick out the same referents in any (centered) world considered as actual. In fact, he sometimes puts the point by saying that phenomenal properties are part of the “epistemic core” that we take with us whenever we consider epistemic scenarios; they are not at a distance from our epistemic perspective on things, but an essential part of it. Consequently, like our physical-theoretical concepts, our direct phenomenal concepts (and derivative standing ones) do not have shifting counteractual patterns of evaluation and are both epistemically and subjunctively rigid.

\(^{346}\) Chalmers (2004), section 4. Pure phenomenal concepts differ from direct phenomenal concepts in that the former are standing concepts which are derived from, and inherit many epistemological and semantic features of, the latter. Pure phenomenal concepts also tend to be more “course-grained” and are not partly constituted by the phenomenal properties they designate. Instead, they are constituted by what Perry (2001a) calls a “Humean Core” together with dispositions to have and to recognize their designated phenomenal properties. Presumably, Chalmers allows for some fallibility in the case of judgments of the form “this experience is such-and-such experience” where the such-and-such experience is thought of under a standing pure phenomenal concept.
But given that our physical-theoretical concepts and our direct phenomenal concepts (and standing ones) both have primary and secondary intensions which align, Chalmers argues, there is a deep problem for physicalism. The problem is that physicalists cannot explain away why it is conceivable that there be scenarios in which the physical facts are just the same as in our world but the experiential facts are different. Whereas in standard Kripkean cases of true a posteriori identities, our concepts have different primary intensions but the same secondary intensions, in the case of putative psycho-physical identities, we cannot appeal to their distinctness of primary intensions but sameness of secondary intensions, and for a straightforward reason. Their primary and secondary intensions are the same in all (centered) worlds (whether considered as counteractual or counterfactual). So it follows that if their primary intensions are distinct, then so are their secondary intensions. But then if zombie worlds are so much as epistemically possible, then they are metaphysically possible as well. So physicalism would be false. And of course, Chalmers contends, zombie worlds are at least epistemically possible. Hence, he concludes, physicalism is evidently false.

Chalmers grants, however, that there is one reasonable strategy a physicalist might take in order to resist dualism: one could follow Maxwell (1978) and deny that our physical-theoretical concepts have aligning primary and secondary intensions across all (centered) possible worlds. As we saw in our previous discussion of Kripke’s modal argument and the explanatory gap, Maxwell makes the case that our physical-theoretical concepts provide us with conceptions of the spatio-temporal and causal-dispositional roles that physical phenomena occupy in concrete reality without characterizing their underlying categorical nature in virtue of which they occupy those roles (given their

distributions in space and time and the laws of nature of this world). Nevertheless, he argues, our physical-theoretical concepts rigidly designate the underlying categorical nature of the physical, and so if our phenomenal concepts rigidly co-designate the same underlying categorical phenomena, then our phenomenal experiences are physical and there is no metaphysical possibility of the experiential and the physical coming apart. What is possible, on the other hand, and what we presumably conceive of when we coherently imagine the possibility of zombies (and inverts), is that there are circumstances in which the spatio-temporal and causal-dispositional structure is the same as in the actual world but different categorical phenomena occupy the relevant roles.

In Chalmers’ vocabulary, such circumstances will verify the primary intensions of our physical-theoretical concepts, but they will not satisfy their secondary intensions. Consequently, such circumstances will be epistemically possible, but they will not be subjunctively (or metaphysically) possible. So the (ideal) conceivability of such circumstances will in no way impugn the truth of physicalism. On the other hand, Chalmers thinks that the ontological view we end up with, when worked out in detail, “is a highly distinctive form of physicalism that has much in common with property dualism and that many physicalists will want to reject.”

While I agree with Chalmers that Maxwell’s “Russellian Physicalism” is a “highly distinctive form of physicalism”, I see nothing in it which should bother a clear-thinking, card-carrying physicalist. And in fact, I plan to develop and defend Russellian Physicalism as a genuinely physicalist ontological view in my future work.

But I do agree that some physicalists will see things otherwise.\textsuperscript{351} Hence, it is worth considering whether or not there are any additional shortcomings in Chalmers two-dimensional argument against physicalism.\textsuperscript{352}

In my view, another problem with Chalmers’ argument is that Russell’s discussions of phenomenal continua cases show us that the primary and secondary intensions seemingly do not align either. Since I have discussed Russell’s thoughts on phenomenal continua cases numerous times already, I will here simply focus on how his remarks bear on the issue of the two-dimensional semantic behavior of our phenomenal concepts.\textsuperscript{353} The crucial point of Russell’s discussions is simply this: In cases of synchronic or diachronic phenomenal continua, we can attend to the qualitative character of some region of our “experiential field” (at a particular moment or over some duration) and not be in a position to determine on the basis of our experience (and selective introspective attention) alone whether or not we are acquainted with one quality or many distinct qualities. On Russell’s view, our knowledge of the numerical identity or diversity of the objects of our acquaintance is knowledge of truths about them, and is logically distinct from our more basic knowledge of things (by acquaintance) of them. Moreover, Russell’s maintains that all of the epistemic methods and capacities we have for arriving at knowledge of truths from our more basic knowledge of things are fallible to some degree or other. In fact, he argues, all of our knowledge of truths is to some degree infected with uncertainty, including our knowledge of truths about the existence,

\textsuperscript{351} For instance, see our earlier discussion in §5 about Shoemaker’s (1980) views on physical phenomena being individuated in terms of their (necessary) spatio-temporal and causal-dispositional profiles.

\textsuperscript{352} In full disclosure, I think that two-dimensional semantic theories themselves are deeply problematic, and so I don’t think Chalmers’ argument even gets off the ground. But this is not the place for me to advance a systematic case against two-dimensionality, and so I will here simply assume its adequacy and focus on problems specific to Chalmers’ use of it against physicalism.

\textsuperscript{353} For our previous discussions of phenomenal continua cases, see Chapter 1, §2, Chapter 3, §5, and Chapter 4, §2 and §5 of this dissertation.
characters, and identities of our own occurrent conscious experiences (though he
acknowledges that these constitute our most certain items of knowledge).

If Russell is correct regarding the consequences of phenomenal continua cases
(and I’ve argued that there is good reason for thinking so), then our phenomenal concepts
seemingly do not have primary and secondary intensions which align. For as we saw in
our discussions of Kripke’s modal argument, we can be in an epistemic situation in which
we take ourselves to be presented with an experiential region with such-and-such a
subjective, qualitative character when in fact we are presented with one with a different
subjective, qualitative character (perhaps a complex one whose distinct proper parts have
varying subjective, qualitative characters). Consequently, when we rationally reflect
(even ideally) on what our phenomenal concepts pick out across the space of epistemic
possibilities, our modal judgments will not align with what our phenomenal concepts
pick out across the space of metaphysical possibilities (given the actual character of our
introspectively attended-to experience). Indeed, given the fallibility (and partiality) of

---

354 As we’ve seen, Russell also opens the door for a species of reference-failure regarding our reflective
phenomenal judgments. For more on these thorny issues, see Chapter 1, §4 and Chapter 3, §5 of this
dissertation.

355 This raises a question about what to say about Chalmers’ notion of “direct phenomenal concepts”. One
might say that we have direct phenomenal concepts which lack the robust epistemic and semantic features
that Chalmers attributes to them. However, Chalmers defines direct phenomenal concepts in terms of such
features (and in terms of ideal rational reflection), so one might instead say that we just don’t have any
direct phenomenal concepts. If we take this second route, then we should say that we instead have what
Chalmers calls “quasi-direct phenomenal concepts”.

Another issue is whether Russell’s remarks about phenomenal continua generalize to all of our
phenomenal concepts. One response available to Chalmers is to say that in cases of phenomenal continua
(and other instances of introspective error) our phenomenal judgments involve quasi-direct phenomenal
concepts, but in unproblematic circumstances our phenomenal judgments involve full-blown direct
phenomenal concepts. So he could still maintain a restricted version of the two-dimensional argument and
his infallibility claim.

The problem with this response is that we would never be able to tell on rational reflection whether any
particular phenomenal concept is direct or merely quasi-direct. So we would not be in a position to know a
priori which of our phenomenal concept-based modal judgments accurately reflect metaphysical
possibility. Of course, if Chalmers could make the case that at least one of our phenomenal concepts is a
direct one, he could still construct a modified two-dimensional argument against physicalism. But it is
unclear how such an argument would go, or at least it is unclear why any physicalist should accept it. After
our attention-based introspective capacities, our phenomenal concepts will presumably lead us astray in our modal reasoning across the space of metaphysical possibilities in any number of ways. So even if we find the existence of zombies (or inverts) epistemically possible, this seemingly tells us nothing about their metaphysical possibility. Hence, Chalmers’ two-dimensional argument against physicalism looks to be in serious trouble.

However, Chalmers (2010) has recently advanced a modified version of the two-dimensional argument which, he claims, goes through even if the primary and secondary intensions of our phenomenal concepts do not align. The first premise of Chalmers’ revised version of the two-dimensional argument is that it is (ideally) conceivable that there be a minimal physical duplicate of the actual world in which there are zombies (or inverts). The second premise is that if it is (ideally) conceivable that there be a minimal physical duplicate of the actual world in which there are zombies (or inverts), then such circumstances are epistemically possible. In other words, Chalmers contends, there is a logically possible world which is a minimal duplicate of ours but which does not verify the primary intensions of one or more of our (direct or pure) phenomenal concepts and corresponding phenomenal beliefs. But, he argues:

This world must differ from our world because the primary intension of [one or more of our direct and pure phenomenal concepts and corresponding phenomenal beliefs] is true in our world… It follows there is a minimal [physical duplicate of our world] that is not a duplicate of our world, so that physicalism is false with regard to our world.357

In fleshing-out this possibility, Chalmers continues:

---

It could be that, strictly speaking, physicalism will be true of *consciousness* because [the minimal physical truths of our world necessitate the phenomenal truths of this world], but physicalism will be false of properties closely associated with consciousness, namely those associated with the primary intensions of [our direct and pure phenomenal concepts and corresponding phenomenal beliefs]. We might think of this sort of view as one in which phenomenal properties are physical properties that have nonphysical properties as modes of presentation.\footnote{Chalmers (2010), p. 153. Chalmers then notes again that a physicalist can avoid dualism by taking the primary and secondary intensions of our physical-theoretical concepts to come apart. The resulting view, he argues, “can be considered another form of Russellian monism in that the intrinsic properties of physics in our world are crucial for constituting the properties associated with the modes of presentation of consciousness”. Once again, I think there is nothing in such a view that should trouble a physicalist, but I will set it aside and focus on whether Chalmers’ revised argument works against other versions of physicalism.}

To sum up Chalmers’ argument here (as I understand it), the idea is that if there is a logically possible minimal physical duplicate of the actual world such that our phenomenal judgments about it come apart from the phenomenal beliefs we have about the actual world, then there must be *some* difference (if only one involving our epistemic access to consciousness there) between it and the actual world which accounts for the difference in our phenomenal judgments.\footnote{I am somewhat hesitant to attribute this argument to Chalmers since I find it very problematic. On the other hand, it is a straightforward interpretation of his remarks, and I have been unable to come up with a more charitable interpretation. So I will proceed as if my reconstruction is accurate.} But physicalism is committed to the view that any minimal physical duplicate of the actual world is a duplicate of our world *simpliciter*. Hence, the difference which explains why our phenomenal judgments about the logically possible minimal physical duplicate world come apart from those about the actual world must involve something nonphysical in *our* world (either regarding consciousness itself or our epistemic access to it). So the two-dimensional argument shows that physicalism is false with respect to the actual world even if the primary and secondary intensions of our phenomenal concepts fail to align.
In my view, there is little reason for the antecedent physicalist, even one who accepts a two-dimensional semantic theory, to accept Chalmers’ modified argument against physicalism. In particular, I think that the antecedent physicalist has good grounds for questioning Chalmers’ inference from the epistemic possibility of a minimal physical duplicate world about which our phenomenal judgments diverge from those concerning the actual world to the metaphysical possibility that there be such a world. That is, the antecedent physicalist ought to deny that for every epistemically possible scenario of which we can ideally and positively conceive, there is a corresponding, qualitatively matching (centered) metaphysically possible world.

Unsurprisingly, Chalmers considers this strategy and offers a number of reasons for questioning whether such a strategy can be successfully carried out. For one thing, he argues, positing a modality that is not tied to what it is rationally coherent for us to suppose opens the possibility of problematic epistemically “brute necessities” which we must take “as primitive in the same way that we take the fundamental laws of nature as primitive”. But, he plausibly maintains, it is doubtful that there are any a posteriori fundamental metaphysical laws which are analogous to fundamental natural laws.

For another thing, Chalmers contends, if we introduce a notion of the space of metaphysically possible worlds that is distinct from (and smaller than) the space of epistemically possible worlds, we must embrace a problematic “modal dualism, with distinct and independent metaphysical modalities and distinct and independent spaces of worlds that answer to them”. On one hand, we will have Chalmers’ notion of “metaphysical modality”, which is grounded in the a priori subjunctive evaluation of our

---

concepts across the space of (centered) logically possible worlds. But on the other hand, we will also have a distinct notion of “metaphysical modality”, which is grounded in some way independent of (ideal) rational reflection and which ranges over a completely distinct (and smaller) space of metaphysically possible worlds. Chalmers argues that this not only creates a fractured picture of modal space and reasoning, it also raises substantive questions about why, and how, epistemic possibility could be any sort of guide to this second species of “metaphysical modality”. In contrast, if we stick to his notion of “metaphysical modality” alone, we can maintain a simpler and epistemically more straightforward picture of modal space and modal reasoning.

Moreover, Chalmers argues that there is simply no reason to postulate any additional notion of “metaphysical modality” or separate space of “metaphysically possible worlds”. First of all, he contends, we do not plausibly have a grip on any pre-theoretical notion of this kind of metaphysical modality, some evidence for which is the lack of clear examples of it being invoked in ordinary non-philosophical discourse.\(^{362}\) Secondly, the most common alleged counterexamples to a notion of metaphysical modality that is constitutively linked to rational reflection, Kripkean a posteriori necessities, can be accounted for on Chalmers’ two-dimensional semantic theory, and so are arguably not actual counterexamples. In fact, Chalmers argues, his account of metaphysical modality as grounded in the a priori subjunctive evaluation of our concepts across the space of (centered) logically possible worlds can account for all of the modal phenomena which needs explaining, such as counterfactual reasoning, inference, the content of thought and talk, etc.\(^{363}\) Together, these considerations lead Chalmers to


conclude that the proposed alternative notion of metaphysical modality “is an invention; nothing in our conceptual system requires it. It is a primitive that answers to no one and does no work”.

Chalmers case against a distinct and independent notion of metaphysical modality is quite powerful. Nevertheless, I do not think that the antecedent physicalist should be swayed by it. On the contrary, I think that we do have some intuitive grip on the relevant notion of metaphysical modality, which we have reason for thinking is real, which is useful for our modal reasoning, and which accords with plausible views about the epistemology of modality. Furthermore, I think that we can construct a plausible case for this notion of metaphysical modality which is general in character, and so does not rely on particular (potentially contentious) counterexamples to Chalmers’ two-dimensional theory.

Unfortunately, I cannot hope to construct such a case here. But what I will try to accomplish here is to provide some evidence which at least suggests that such a project can be successfully carried out. Hopefully, this will be enough to lend serious weight to my claim that the physicalist should reject Chalmers’ revised two-dimensional argument against physicalism by denying that the mere epistemic possibility of a (centered) minimal physical duplicate of our world in which there are zombies (or inverts) entails the metaphysical possibility of such a world.

The first thing I want to note is that Chalmers himself doesn’t hold that we only have grip on modal notions which are explicitly tied to what it is rationally coherent to

---

365 I intend to do so in future work, however.
366 I also think that an antecedent physicalist should adopt some version of Russelian Physicalism, but that is another matter.
suppose. For in addition, he thinks that we also have a grip on the modal notion of 
natural necessity, which is tied to how things could be in our world given the particular 
laws of nature it happens to have. His real objection is to any modal notion that is wholly 
distinct from any rational modality or natural modality. In a sense, I agree with him on 
this point. On the other hand, I don’t think that such agreement forces us to deny the 
reality of a metaphysical modality which is divorced from what is rationally coherent to 
suppose. This because there is a plausible case to be made that our pre-theoretical notion 
of natural necessity is simply a restricted version of a broader pre-theoretical modal 
notion (one which arguably corresponds perfectly with the proposed theoretical notion of 
metaphysical modality), and so is not wholly distinct from it.

In particular, I think that our modal notion of “natural necessity” is grounded in a 
broader notion of the ways the world could be independently of our ways of thinking and 
talking about it. On my view, we have a powerful pre-theoretical intuition that there is 
something in the nature of the world itself (and of its various parts) such that given how 
its parts are arranged, and given the laws of nature which happen to govern it, things 
work in such-and-such a manner. (In fact, I think that this is the very pre-theoretical 
notion that ultimately lends intuitive weight to the Russellian possibility that the 
underlying “categorical” physical phenomena of the actual world could occupy different 
spatio-temporal and causal-dispositional roles in other possible worlds, a possibility that 
Chalmers himself finds fairly plausible.) What’s more, I believe that we have an equally 
powerful pre-theoretical intuition that our own powers of (ideal) rational reflection and 
reasoning are only imperfect guides to the ways in which the underlying nature of the 
world (and of its various parts), together with the arrangement of its parts and the laws of
nature which govern it, determine its workings. This is because we are cognizant of (or at least sensitive to) the fact that many of the concepts we rely on to represent and characterize the “categorical” nature of the world (and of its parts), its arrangement, and the laws of nature which govern it, are partial, opaque, and non-revelatory.

I suspect that our distinct notion of “metaphysical modality”, of the ways the world could be independently of our ways of thinking and talking about it, is a more-or-less straightforward consequence of combining these two powerful pre-theoretical intuitions. Moreover, as I’ve said, I think that our notion of “natural necessity” is simply a restricted version of this broader notion in which we hold fixed what we take the laws of nature in the actual world to be, while allowing our conceptions of the facts about the nature of the world (and of its parts) and/or its arrangement to vary. However, given that such laws of nature are intuitively contingent, and given our cognizance of (or sensitivity to) the fact that our conceptions of them are partial, opaque, and non-revelatory in any case, it is plausible that we should have some grip on a modal notion in which the laws of nature are free to vary as well.

Of course, if there is such a modal notion of “metaphysical modality” which corresponds to the ways the world could be independently of our ways of thinking and talking about it, Chalmers is right to note that it will generate epistemically brute necessities. However, the antecedent physicalist should not be worried about the possibility of brute necessities, so long as they are merely epistemically brute ones. For as we noted in the previous section on the explanatory gap, there is arguably a more-or-less clear distinction to be drawn between epistemically brute necessities and ontologically brute necessities. And the physicalist who is inclined to endorse a notion of metaphysical
modality which is divorced from (ideal) rational reflection should have no additional
qualms about accepting necessities which, epistemically speaking, appear brute, while at
the same time maintaining the belief that there must be something in the metaphysical
nature of things such that the relevant relation is a straightforward, natural consequence
of how things are.

Moreover, the antecedent physicalist needn’t suppose that such epistemically
brute necessities will involve a posteriori fundamental metaphysical laws which are in
any way analogous to fundamental natural laws. To see this, we can simply consider the
metaphysical necessities that many physicalists take to result from the nature of
numerical identity. Indeed, such identity-involving metaphysical necessities are
intuitively quite different in character from any sort of fundamental natural law. So if the
antecedent physicalist is right in offering them as paradigmatic examples of the kinds of
metaphysical necessities which are epistemically brute, we have little reason to suppose
that Chalmers is correct in asserting that a notion of “metaphysical modality” will even
typically involve problematic a posteriori fundamental metaphysical laws.

What’s more, once we begin reflecting on the many issues raised by the nature
of numerical identity, it becomes far less difficult to see why a notion of “metaphysical
modality” which is divorced from (ideal) rational reflection might be very useful for our
modal reasoning. For if we are cognizant of (or at least sensitive to) the fact that the mere
epistemic possibility of such-and-such circumstances on (ideal) rational reflection is no
sure guide to its metaphysical possibility where numerical identity (and the resultant
possibility of misidentification) is involved, we will want some means of taking this into
account in our modal reasoning about the world. And of course, one way of taking this
issue into account, and perhaps the only way, is to posit two distinct spaces of possibility
(where one is smaller than the other) with distinct and independent modal notions
answering to each of them.

However, it is far from clear that anything I have said here in motivating an
alternative notion of “metaphysical modality” undermines the link between (ideal)
conceivability and metaphysical possibility completely. For instance, an antecedent
physicalist might hold that (ideal) conceivability of such-and-such circumstances at least
provides prima facie evidence of its metaphysical possibility in much the same way that
having a perceptual experience as of such-and-such circumstances (arguably) provides
prima facie evidence that such-and-such circumstances are the case. Of course, in the
case of perceptual experience, we often disregard the prima facie evidence it provides,
such as when we have reason to believe we are in circumstances in which our perception
is unreliable. But in the absence of such reasons for supposing that we are in
circumstances in which our perception is unreliable, we are quite within our rights in
taking such-and-such to be the case on the basis of our perceptual experience as of such-
and-such.

Similarly, an antecedent physicalist might argue, if we have no reason to suppose
that we are in circumstances in which our (ideal) conceivability is an unreliable guide to
metaphysical possibility, we are within our rights to take such-and-such circumstances to
be metaphysically possible given we can (ideally) conceive of such-and-such
circumstances. On the other hand, where we have reason to believe that the
circumstances render the evidence provided by (ideal) conceivability to be unreliable (as

367 I am not here endorsing this view of the epistemology of perception; I am simply using it as one example of how an antecedent physical might hold on to some link between (ideal) conceivability and metaphysical possibility.
when we think of one and the same thing or feature via cognitively and logically distinct physical-theoretical concepts and phenomenal concepts), we should disregard the prima facie evidence provided by the (ideal) conceivability of such-and-such circumstances.\textsuperscript{368}

Whether or not this particular proposal has sufficient merit, I think there is good reason to suppose that an antecedent physicalist can maintain some more-or-less reliable link between (ideal) conceivability and metaphysical possibility. On the other hand, I think that he or she has very good grounds for denying that (ideal) conceivability is any kind of guide to metaphysical possibility in the specific case of psycho-physical identities. For not only does the antecedent physicalist have good reason for thinking numerical identity is involved in such cases, he or she can appeal to my acquaintance-based account of phenomenal concepts (and my contrasting account of physical-theoretical concepts) to explain why the (ideal) conceivability of zombies (and inverts) provides no reason for supposing that they are metaphysically possible. Thus, while I have certainly not here provided a knock-down defense of physicalism, I suspect that the antecedent physicalist has little to fear from either Chalmers’ original two-dimensional argument against physicalism, or his revised version.

§8 CONCLUSION

In this chapter, I have presented my own acquaintance-based account of phenomenal concepts and used it to say why there is a distinctive epistemic gap between the

\textsuperscript{368} Moreover, an antecedent physicalist who (unlike me) embraces some version of a two-dimensional semantic theory can still maintain that in cases in which the primary and secondary intensions of their concepts align, conceivability is always a reliable guide to possibility. So there would remain some link between (ideal) conceivability and metaphysical possibility. But he or she could at the same time deny that the (ideal) conceivability of zombies (or inverts) entails their metaphysical possibility, since we have seen that there is good reason for thinking that neither our physical-theoretical concepts, nor our phenomenal concepts, have primary and secondary intensions which align.
experiential and the physical, but no corresponding ontological gap. I have argued that there are five features which, when conjoined, distinguish our phenomenal concepts from our physical-theoretical concepts and familiar demonstratives. Firstly, our phenomenal concepts are genuinely singular, directly referential devices that pick out experiential states, properties, or kinds. Secondly, their reference is fixed in a role-based manner by an inner demonstration on the basis of our selective introspective attention. Thirdly, they are for this reason ‘irreducibly perspectival’ and hence not reducible to any discursive or descriptive “objective” concepts. Fourthly, their inner reference-fixing is ultimately made on the basis of our standing in special informational relations (i.e. “acquaintance relations”) to our own experiential states. In particular, the reference of genuine phenomenal concepts is fixed by an inner demonstration (on the basis of selective introspective attention) to our self-presenting experiential states of which we are “phenomenally” aware simply by having them. And fifthly, because the reference of our phenomenal concepts is grounded in our standing in special informational relations to our own experiential states, relations which are tied to special epistemic methods and capacities for picking up information about those states, they provide us with distinctive conceptions of them: conceptions of their subjective, qualitative characters as subjective and qualitative.

I have also argued that my acquaintance-based account of phenomenal concepts is more promising than the leading alternative strategies for defending physicalism without flat-footedly denying the existence of a distinctive epistemic gap. Unlike standard acquaintance-based defenses of physicalism, it can explain the special features of our thought and talk about the experiential. Unlike the demonstrative-recognitional model,
my acquaintance-based model has the resources to say what is distinctive about the epistemic gap between the experiential and the physical. And unlike the constitutional model, my acquaintance-based model has the resources to explain the special epistemic relation between experiences and our thoughts about them without committing us to the infallibility of our phenomenal judgments.

Finally, I have argued that my acquaintance-based model undercuts the leading dualist arguments in a more direct way by challenging the crucial dualist assumption that acquaintance provides us with transparent or revelatory knowledge of the essence of the experiential. More specifically, I have shown, in turn, how my account deals with Chalmers’ zombie argument, Jackson’s knowledge argument, Kripke’s modal argument, Levine’s explanatory gap argument, and Chalmers’ two-dimensional semantic argument against physicalism. Certainly, there is much more to say about each of these dualist arguments. And a full defense of my account of phenomenal concepts would require considering a number of recent arguments against the viability of any version of the phenomenal concepts strategy. But by showing how an acquaintance-based account of phenomenal concepts at least promises to defuse all of the leading dualist arguments, I hope to have pointed out a new and fruitful path for defending physicalism. For if my acquaintance-based account of phenomenal concepts is correct, an antecedent physicalist can grant that there is a distinctive epistemic gap between the physical and the experiential without being forced to conclude that there is a corresponding ontological gap.
REFERENCES


269


----------. (In Draft). Disjunctivism and the Intentionality of Perception.


--------- (Forthcoming). Referring to the World.


