Foundations of the Problems of Consciousness

Stevens, Christian

Awarding institution:
King's College London

The copyright of this thesis rests with the author and no quotation from it or information derived from it may be published without proper acknowledgement.

END USER LICENCE AGREEMENT

Unless another licence is stated on the immediately following page this work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International licence. https://creativecommons.org/licenses/by-nc-nd/4.0/

You are free to copy, distribute and transmit the work

Under the following conditions:

- Attribution: You must attribute the work in the manner specified by the author (but not in any way that suggests that they endorse you or your use of the work).
- Non Commercial: You may not use this work for commercial purposes.
- No Derivative Works - You may not alter, transform, or build upon this work.

Any of these conditions can be waived if you receive permission from the author. Your fair dealings and other rights are in no way affected by the above.

Take down policy

If you believe that this document breaches copyright please contact librarypure@kcl.ac.uk providing details, and we will remove access to the work immediately and investigate your claim.
Foundations of the Problems of Consciousness

A thesis submitted to King’s College London for the degree of
Doctor of Philosophy

Christian P.G. Stevens

Faculty of Arts and Humanities,
Department of Philosophy
King’s College London
2021
First and foremost, I would like to thank my advisor Eliot Michaelson for providing emotional support, encouragement, and lots of valuable feedback over the years. Without Eliot’s help and extraordinary patience, I very much doubt that I would have managed to complete this work. Eliot’s strong philosophical background, and remarkable ability to distill complex philosophical arguments and ideas has been a privilege for me to come to know. The support and advice in all matters from my wife, Nadine, has also been invaluable. I have been blessed to have Nadine as my constant companion throughout the ups and downs of this process. Finally, this work would not have been possible without the loving support of my parents, to whom I dedicate this thesis.

I owe a great many debts to fellow PhD candidates and now PhD’s in philosophy from KCL. Their input on my ideas over the years played an important role in my general philosophical development, as well as the views and arguments found in this thesis. I would particularly like to thank Matthew Bodle for his careful consideration of many of my half-baked ideas and arguments.

I would also like to thank David Papineau, Bill Brewer, and Christopher Hughes for feedback on my work, as well as cordial philosophical debate and discussion. Thanks to Benj Hellie for accepting me as an exchange student at U of T for a semester, and taking plenty of time to chat with me about philosophical issues in the philosophy of mind. Benj has a very unique set of philosophical views
spanning a wide range of topics that I would like to think broadened my philosophical horizons, though I can’t say I agree with, or even really understand, most of them.

Finally, I am grateful to both Derek Ball and Keith Hossack for their feedback on the first version of this thesis which ultimately led to quite a different one which, I think, is much improved. Many of the ideas from the first version of the thesis remain off-stage in this one, but they helped broaden and deepen my overall understanding of the philosophical problems of phenomenal consciousness that is hopefully on display in this work.
**Abstract**

Phenomenal consciousness is a central topic in the philosophy of mind. Notoriously contentious, even when it comes to the basic issue of the phenomenon or phenomena to be explained, phenomenal consciousness continues to inspire a broad range of philosophical and empirical work, particularly concerning its place in the natural order. In this thesis, I re-examine some of the foundational issues in this field, and end up with some unorthodox answers—particularly regarding the issue of how best to understand or characterize phenomenal consciousness as a topic for theoretical exploration. In my view, this issue has not received the attention it deserves, and with a variety of subtly different views available for how best to understand phenomenal consciousness as a topic for inquiry, it behooves us to take this question seriously.

In arguing for an unorthodox approach to characterizing ‘phenomenal consciousness’, we also arrive at an important division of conceptual labor between phenomenal consciousness and what is often referred to as ‘phenomenal character’ or ‘phenomenal qualities’. The broad framework for understanding phenomenal consciousness argued for here draws a sharp and principled distinction between phenomenal consciousness and phenomenal character that is put to work in a number of different theoretical contexts, especially the context of the ‘explanatory gap’. Certain major issues related to phenomenal properties such as Gilbert Harman’s (1990) well-known ‘transparency argument/thesis’ are explored in this thesis, and the venerable topic of pain is given a novel treatment.
A central theme developed in this thesis is that of the importance of introspection in the context of theorizing about phenomenal consciousness itself and phenomenal properties. This theme is established early on through my characterization of phenomenal consciousness itself in terms of ‘conditional introspective access’. Unfortunately, theorizing about phenomenal consciousness is (understandably) often done without exploring details concerning its relationship to introspection, and introspection itself. This thesis devotes rather more attention than is customary to the topic of introspection in a thesis about phenomenal consciousness, developing a framework for analyzing introspection’s ‘nuts and bolts’ along fairly traditional ‘normative’ lines.

This framework for understanding introspection is put to work later in the thesis, yielding some new insights in traditional areas such as the explanatory gap. While my focus is not the popular issue of whether or not phenomenal consciousness or certain phenomenal properties are conceptually and/or metaphysically reducible to broadly physical properties or other psychological properties such as representing a content, I do devote some attention to this issue in the last chapter based on my findings up to that point. Certain theses about introspection play an important role in my discussion of the matter, and we arrive at some novel anti-reductionist ideas in this area that are difficult to evaluate in terms of whether they pose more or less of a threat than usual to reductive approaches. The explanatory gap problem and related issues as conceived in this final chapter suggest a fairly broad re-evaluation of these matters may be in order.
Table of Contents

INTRODUCTION

Chapter 1
1.1 A problem of starting points
1.2 A preliminary attempt at characterizing ‘phenomenal consciousness’
1.3 The has-a-phenomenology characterization
1.4 The has-a-phenomenology characterization and the problem of blindsight
1.5 Taking stock and the conditional introspective access characterization
1.6 A loose end

2.1 The higher-order characterization
2.2 Epistemic versus non-epistemic higher-order awareness
2.3 A different argument for an epistemic understanding of higher-order awareness
2.4 The non-epistemic higher-order characterization
2.5 ‘Phenomenal consciousness’ ‘phenomenal qualities’

3.1 The explanatory gap constraint
3.2 Weakening the explanatory gap constraint
3.3 The no-currently-obvious-reductive-analysis constraint and introspective justification
4.1 The kind/type versus predicate views of phenomenal consciousness ........................................41
4.2 The scope of the divide .................................................................43
4.3 Why the kind view of phenomenal consciousness should be rejected ..................................................44

5.1 A summary of the argument thus far .........................................46

6.1 An argument for the conditional introspective access characterization ........................................48
6.2 ‘Access-consciousness’ versus ‘conditional introspective access’ ..................................................49
6.3 A hypothetical characterization and conditional introspective access ..............................................51
6.4 Nothing to lose by accepting the conditional introspective access characterization? ..................52
6.5 Block’s puzzle case for phenomenal consciousness without introspective accessibility ..................53
6.6 The Sperling task doesn’t suggest phenomenal consciousness without introspective access ............54

7.1 Three problems of consciousness .................................................55

Chapter 2

1.1 The details of introspection problem and moving forward ...58
1.2 First-person warrant and the content of introspection ..........58
1.3 An epistemological framework for introspection ...............61
1.4 Epistemic support and propositional versus doxastic justification ..................................................62
1.5 Challenges to first-person warrant ............................................63
1.6 Having epistemic justification as having permission to believe .................................................................64

2.1 Phenomenal consciousness without introspective access ....66
2.2 A minimal account of introspection ........................................67

3.1 The conditional introspective access characterization, introspection, and phenomenal consciousness .................................68
3.2 Major theoretical options for the relationship between introspection and phenomenal consciousness .........................71
3.3 Phenomenal consciousness and the content of introspection ........................................................................74

4.1 A partial summary and some general conclusions ............75

Chapter 3

1.1 The problem of the phenomenal character of phenomenally conscious mentalia ......................................................78
1.2 Phenomenal properties and being phenomenally proprietary .................................................................80

2.1 Phenomenally conscious perceptual events and the contents of introspection ....................................................82
2.2 Can we attend to our own instantiations of perceptual awareness? .................................................................84
2.3 Harmanian Transparency .........................................................86
2.4 Some venerable hard cases: still no inner attention to perceptual awareness itself? ........................................89
2.5 Phenomenal events and the inner sanctum ...........................95
3.1 Defeasible introspection support and Harmanian transparency ........................................................................98
3.2 An argument against Harmanian transparency ..................100

4.1 A problem about pain ............................................................102
4.2 The ‘painfulness’ theory of pain ............................................103
4.3 Pain asymbolia and the painfulness theory .......................105
4.4 Pain asymbolics don’t have pains after all—a defense of the painfulness theory ..................................................107
4.5 Pain asymbolia research and risk of bias .........................109
4.6 Why a headache isn’t a kind of pain: a consequence of the painfulness theory .......................................................112
4.7 A composite view of pain .......................................................114
4.8 Simple CV and the painfulness swamping problem ............117
4.9 Painfulness without pain .......................................................117
4.10 Isn’t it just a verbal issue? A deflationist objection to the painfulness theory .........................................................119
4.11 Pains as sources or causes of a certain kind of displeasure to a certain degree .......................................................120
4.12 Objections to the pains as causes of displeasure view ......122

5.1 Physicalism and reductionist approaches to phenomenal character ......................................................................125

Chapter 4

1.1 Phenomenal consciousness, first-person warrant, and the explanatory gap ............................................................128
1.2 The distinctive source of the explanatory gap problem for phenomenal consciousness .............................................129
2.1 Self-support and awareness accounts of first-person warrant .................................................................131
2.2 Self-support and awareness accounts are the only games in town ........................................................133
2.3 Clarifying the self-support option ........................................133
2.4 Arguments for first-person warrant as non-trivial self-support ................................................................134
2.5 First-person warrant as non-trivial self-support and the explanatory gap ..............................................136
2.6 The awareness account of first-person warrant and the explanatory gap ..................................................137
2.7 The problem of different kinds of explanatory gaps ..........140
2.8 But is this really an ‘explanatory gap’? .........................141

3.1 Phenomenal properties and the explanatory gap ...............142
3.2 Why analytic representationalism can’t succeed ..............145
3.3 Chalmers on analytic representationalism .......................148
3.4 Phenomenal qualities and the Nagelian argument against analytic functionalism ..................................150
3.5 Is there introspective support for ‘phenomenal qualities’ qua qualities? ..................................................154

4.1 Phenomenally proprietary properties and reduction ..........156
4.2 Is painfulness phenomenally proprietary? .......................158
4.3 No good case for secondary qualities being phenomenally proprietary? ............................................159
4.4 No overlap, partial overlap, and full overlap views of phenomenal properties ........................................162

5.1 The two main explanatory gaps in consciousness studies ...164
**Introduction**

This thesis is about phenomenal consciousness and related problems. One of the challenges for theorizing about phenomenal consciousness is that of sorting through the many different ways available for thinking about the relevant phenomena to be explained. That is one of the major tasks of chapter 1. The philosophical literature on phenomenal consciousness has tended to focus on certain (alleged) problems that phenomenal consciousness or associated properties pose for certain prominent theoretical approaches for understanding *mentalia*, such as functionalism and representationalism/intentionalism. Unlike other mental phenomena, it is supposed that phenomenal consciousness or closely associated ‘phenomenal properties’ present a special sort of explanatory burden that doesn’t arise when considering other mental properties such as representing a content. This has resulted in a certain ‘we all know what we’re talking about so let’s get down to business’ approach to phenomenal consciousness, that has, in my view, been detrimental to philosophical progress in this area.

I share the view phenomenal consciousness is ‘wonderful’ in the sense that it provokes a sense of wonder in us that other properties of the mental, such as representing a content, do not. But I also think that the notion of ‘phenomenal consciousness’ is an extremely slippery one that we would do well to carefully mull over before proceeding to more ‘exciting’ topics such as whether or not it is metaphysically basic or fundamental. So, I am not happy to just assume that certain common turns of phrase in the relevant literature can be relied upon without further discussion in tackling the sorts of broad metaphysical issues that most are primarily interested in here.
This makes **chapter 1** of this thesis rather unusual in that we will be agonizing over how to understand phenomenal consciousness and paying close attention to apparently small differences in how we introduce it as a topic for theoretical exploration. My approach here will ultimately be rather unorthodox, and reliant on another controversial notion in the vicinity: introspection. For reasons that should be obvious, though we will make sure they are made so later on, we shouldn’t understand phenomenal consciousness in terms of introspection, but, nevertheless, I will argue that a certain attenuated notion of introspection does provide a satisfying characterization. This entails that there is a closer link than has previously been recognized between phenomenal consciousness and introspection, and we would do well to make clear how we understand the latter in the context of theorizing about the former.

Unfortunately, these two topics have not always been considered together and it is sometimes assumed that we can leave the topic of introspection entirely open while investigating phenomenal consciousness. Part of this tendency, I think, derives from a tendency to elide or conflate the distinction between phenomenally conscious events and certain theoretically interesting and problematic properties of phenomenally conscious events. Clearly separating these two issues is an important agenda in **chapter 1** as well.

Returning to the issue of phenomenal consciousness and introspection, my issue is that I don’t think we can do a very good job in theorizing about phenomenal consciousness without a rather substantial notion of introspection on the table. And if we are unsure or cagey about how to understand introspection, or how to understand a relevant notion of ‘introspection’ in relation to phenomenal consciousness, then we
can’t, I think, get very far when it comes to alighting on the core theoretical issues surrounding phenomenal consciousness.

So, in chapter 2, I develop an epistemological framework for thinking about introspection according to which introspection is a kind of non-inferential normative epistemic support. This is a pretty standard way of understanding introspection, but the details are important since, among other things, there is a tendency in the literature to conflate the sort of epistemic warrant that phenomenal consciousness confers, with the epistemic status of certain beliefs about phenomenally conscious events.

While it is often assumed that our introspectively formed beliefs about phenomenal consciousness have a certain unusual epistemic strength or force, such that they are infallible or close to, or are ‘incorrigible’, I will not be making any such assumption here. That we can develop a theoretically potent notion of introspection which does not assume introspectively formed beliefs are infallible or close to, is, I think, a virtue of my account. There are also other important issues concerning the relationship between phenomenal consciousness and introspection, as well as introspection more generally, that are discussed chapter 2 as well.

From this point, we will turn to what has often been treated as the core of the problem of consciousness: ‘phenomenal qualities’. ‘Phenomenal qualities’ are certain properties we have introspective access such as ‘sensory colors’. Theoretical problems concerning colors as understood/known through our introspective access to perceptual awareness, as well as other ‘phenomenal properties’ such as pain, will occupy us in chapter 3, though the focus here is selective so as to make certain points of contact with topical debates in the literature.
In the fourth and final chapter, we finally turn our attention to the problem of whether or not phenomenal consciousness and certain phenomenal properties are such that they pose an ‘explanatory gap’. **Chapter 4** is rather unusual in that it does not engage with many of the usual sorts of problems in this area, and part of the reason for this is that my understanding of the ‘explanatory gap’ posed by phenomenal consciousness itself is novel. Normally, the explanatory gap is understood as the problem of explaining certain ‘phenomenal qualities’ like colors as understood/known through introspection. But I will have already argued by this point that we need to carefully distinguish between these sorts of properties and phenomenal consciousness. Phenomenal consciousness itself is understood in this thesis as presenting a rather unorthodox explanatory challenge to reductionist approaches.

My understanding of the explanatory gap posed by properties such as ‘sensory redness’ is more traditional, but, again, I don’t go in for the usual back and forth in this area—partly since the ground is so well-trodden at this point that I don’t have much to add to it. That said, I do provide at least one line of argument against ‘conceptual reductionist’ approaches to ‘secondary qualities’ such as colors that I think is helpful in that it avoids the sort of deadlock described by David Chalmers (1996) as ‘the great divide’ in consciousness studies literature. While it’s true that most theorists these days have abandoned ‘a priori physicalism’ and other related ‘conceptual reductionist’ approaches to the relevant phenomenal properties, I still think it is of significant interest to consider a different tack when it comes to arguing against these reductionist views.

My conclusion is that conceptual reductionist approaches don’t fare well when it comes to certain phenomenal properties such as colors, but I leave it open
that they may have more promise when it comes to phenomenal consciousness itself. If the explanatory gap can be bridged for phenomenal consciousness itself but not phenomenal properties, this may seem a hollow victory, but I think it would still be of considerable interest, and serve to sharpen the terms of the debate for future work in this area. That said, I am unsure whether the explanatory gap in the case of phenomenal consciousness itself can be bridged, so I leave it as an important question for future work in this area whether our new understanding of the explanatory gap in this area should be viewed as a reason for optimism or pessimism about reductionist approaches here.

Either way, as I’ve already suggested, I think that this conceptual division of labor here between the relevant explanatory gaps helps to clarify and advance the debate in this area. So, I think that the conclusions arrived at in chapter 4, while modest, provides us with materials that have the potential to stimulate novel and innovative future work. Hopefully, much of the work throughout has this potential as well!
CHAPTER 1

Understanding Phenomenal Consciousness

1.1: A problem of starting points

The availability of multiple starting points for theorizing about phenomenal consciousness is, in my view, a significant obstacle to progress in this area. Theorists will often work from one of these starting points without explaining or justifying their decision. This is understandable, and to some extent inevitable, but one is then left to wonder whether the theory developed from one of these starting points would hold up in the context of another.

A further problem lies in properly identifying and distinguishing between different starting points for theorizing about phenomenal consciousness. Subtle differences in theoretical starting points in this area are often neglected or assumed to be of little relevance—for all intents and purposes, it is supposed, they may be collapsed. I find this widespread practice problematic, and one of the goals of this chapter is to reveal how some of these subtle differences in preliminaries place different constraints on the adequacy of a theory of phenomenal consciousness. The overall goal of this chapter, however, is to separate the wheat from the chafe as much as possible when it comes to different starting points for theorizing about phenomenal consciousness.

1.2: A preliminary attempt at characterizing ‘phenomenal consciousness’

Phenomenal consciousness, we are often told, is the most obvious thing in the world: you are phenomenally conscious if you are reading this right now, and the
psychological state you are in while you read it is also phenomenally conscious—if your comprehension of the sentence just read was not phenomenally conscious, then you wouldn’t be able to tell me what you had just finished reading—but you, dear reader, will have no problem doing that! Unfortunately, matters aren’t as straightforward as such bombastic remarks would suggest.

First off, do we really want to build-in a tight connection between phenomenally conscious psychological states/events and introspection? If so, are we saying that such a connection is at least partly definitional of phenomenally conscious mental states/events? Do we mean the same thing by ‘phenomenal consciousness’ when attributed to a creature at a time as we do when we call a psychological state/event ‘phenomenally conscious’? Most importantly, we cannot arrive at an adequate understanding of ‘phenomenal consciousness’ with just a positive example or two that gets the extension right. No doubt you are, as a matter of fact, phenomenally conscious as you read this, but you are also a lot of other ways as well—including other psychological ways such as mentally representing this sentence.

As Ned Block (1995) famously argues as well, “...this sort of pointing is flawed, because it points to too many things, too many different consciousnesses.”¹ Phenomenal consciousness is so-called for a reason so as to distinguish it from phenomena denoted by other senses of ‘consciousness’ such as the use of ‘conscious’ as a synonym for ‘awake’ or ‘responsive to stimuli’ and so on. Phenomenal consciousness may be a manifest feature of our lives, but that doesn’t mean it is easy to zero-in on for the purposes of a clear and pointed discussion of it.

¹ Block, 1995, p.230.
This much is widely recognized in the consciousness studies literature. The trouble begins with how to proceed. A relatively ‘theory-neutral’ way of honing-in on the target phenomenon is the name of the game here. We want to limit as much as possible the incursion of controversial philosophical doctrines into our characterization of the target phenomenon. If we find that what we are actually dealing with is a cluster of interrelated but distinct phenomena, then the goal is to make sure that these explananda are distinguished as clearly as possible. With these points in mind, let us now turn to what is undoubtedly the most popular positive definition of ‘phenomenal consciousness’—the ‘something it is like...’ or, as I will call it, the ‘has-a-phenomenology’ characterization.

1.3: The has-a-phenomenology characterization

The has-a-phenomenology characterization of ‘phenomenal consciousness’ is rooted in a longstanding tradition heralded by Thomas Nagel’s (1974) celebrated paper, “What is it like to be a bat?”. According to this tradition: a mental event \( e \) is ‘phenomenally conscious’=df. there is something \( e \) is like for \( e \)’s subject. I follow Benj Hellie (2007) in understanding the frame ‘there is something \( x \) is like’ to be synonymous with ‘\( x \) is some way’. Thus, ‘there is something \( e \) is like for \( e \)’s subject’, means: \( e \) is some way for \( e \)’s subject.\(^2\) Now, I think it’s clear that the success of this characterization hinges on whether or not we can get the right sort of mileage out of ‘for the subject’. As Hellie (2007) points out, ‘for the subject’ may simply be understood as equivalent to ‘as regards/to the subject’ in this context.\(^3\) So, one complete unpacking would run as follows: a mental event \( e \) is ‘phenomenally conscious’=df. \( e \) is some way as regards \( e \)’s subject. Naturally, this seems too vague

---

\(^2\) Hellie, 2007, p.448.
\(^3\) Hellie, 2007, p.461.
to be helpful. To provide an illustrative example: e’s being such that e’s subject is
caused to act in a certain way by e is a way for e to be as regards e’s subject.

But if the above semantic analysis is correct, and assuming it helps us zero-in
on something important, then what is the semantic or pragmatic ‘enrichment’ (I
intend this in a non-committal, non-technical sense) we seem to be tapping into
here? One plausible suggestion: e seems some way to/as regards e’s subject. As long
as we interpret the ‘something it is like’ frame this way, we needn’t worry about
‘spurious’ ways that a mental event can be with respect to its subject, while
upholding a straightforward reading of being ‘for the subject’.

The ‘seems’ reading also has the virtue of according with a great deal of
philosophical work in this area that treats ‘there is something e is like for e’s subject’
and ‘e seems some way to e’s subject’, as interchangeable. A mental event’s seeming
some way or other to its subject is often abbreviated by saying that it ‘has a
phenomenology’. Since I’m inclined to think that the ‘seems’ interpretation or
reading of the ‘something it is like...’ characterization of a ‘phenomenally conscious’
mental state/event is best, that is why I call the ‘something it is like...’
characterization the ‘has a phenomenology’ characterization.

No doubt a number of theorists will want to challenge my interpretation of
the ‘something it is like...’ frame in this context. I examine the issue in greater depth
later on. For now, I am only concerned to argue that, short of adopting both the
‘seems’ interpretation of ‘something it is like...’, and a controversial epistemic
unpacking of that interpretation, the ‘something it is like...’ characterization, taken
on its own, is of no use in zeroing in on the target phenomenon. This is liable to
strike many as a startling and implausible claim, but I think that reflection on cases
involving the absence of phenomenal consciousness, such as blindsight, make it all but inevitable that we should accept it.

1.4: The has-a-phenomenology characterization and the problem of blindsight

An adequate characterization of phenomenal consciousness should help us identify paradigmatic cases in which it is absent. Type1 blindsight (henceforth, simply, ‘blindsight’) is that notorious condition in which individuals who are functionally blind (they pass the usual tests for blindness) in a part(s) of their visual field due to brain damage, nevertheless retain the ability to discriminate, and to a limited extent spontaneously respond to, certain stimuli presented to their ‘blind’ field.

Now, it is widely held that blindsight perception is a paradigmatic example of a mental state/event-type that lacks phenomenal consciousness. Does the has-a-phenomenology characterization of ‘phenomenal consciousness’ render the verdict that blindsight perception is phenomenally unconscious? Well, let’s ask ourselves: is the blindsighter’s perception of items in their ‘blind’ field such that there is something that mental event is like for its subject? It would seem that the has-a-phenomenology characterization, in and of itself, is entirely silent on the matter. If so, then the has-a-phenomenology characterization is inadequate.

I anticipate that at least some proponents of the has-a-phenomenology characterization will respond to this by suggesting that there is a conceptual entailment of some sort between having a phenomenology and introspection. If, for example, a mental event’s having a phenomenology conceptually entails a kind of conditional introspective access to that event, then it may be that the has-a-
phenomenology characterization of phenomenal consciousness can, after all, succeed
in identifying cases in which a mental event lacks phenomenal consciousness.

The blindsighter clearly doesn’t have introspective access to the fact that they
are perceiving items in their blind field. And since there is no reason to suppose that
blindsighters lack or have a degraded capacity for introspection in general, we
conclude that blindsight perception lacks a property others have which is needed for
it to be introspectively accessible. In other words, a mental event $e$’s being
phenomenally conscious, conceptually entails that in an individual with a normally
functioning or intact capacity for introspection, $e$ is introspectively accessible. And
since the blindsighter is an individual with a normally functioning or intact capacity
for introspection, we conclude that the blindsighter’s visual awareness of items in
their blind field is not phenomenally conscious.

A major problem for this strategy, however, is that having a phenomenology
doesn’t obviously conceptually entail conditional introspective access: where’s the
contradiction in supposing that a mental event has a phenomenology, but that its
subject lacks conditional introspective access to it? We could, of course, opt for a
certain reading of the ‘seems’ interpretation of the has-a-phenomenology
characterization of ‘phenomenal consciousness’ according to which a mental event’s
‘seeming some way or other to its subject’ is roughly equivalent in meaning to a
mental event’s being ‘conditionally introspectively accessible’. That is why I
claimed earlier that the has-a-phenomenology characterization of ‘phenomenal
consciousness’ is, on its own, inadequate, unless we adopt a certain epistemic
version of the ‘seems’ interpretation of the has-a-phenomenology characterization.
1.5: Taking stock and the conditional introspective access characterization

Let us take stock at this point. I have briefly argued that the ‘something it is like...’ or has-a-phenomenology characterization of ‘phenomenal consciousness’ is, on its own anyway, unlikely to bring us the sort of clarity we need to begin full-blown theorizing about phenomenal consciousness. My argument was that an adequate characterization of ‘phenomenal consciousness’ should be able to identify paradigmatic instances of phenomenally conscious mental state/event-types, and that the has-a-phenomenology characterization fails to meet this condition—it cannot even render the verdict that blindsight perception is a phenomenally unconscious mental state/event-type.

From there we examined a reply on behalf of the proponent of the has-a-phenomenology characterization, according to which ‘having a phenomenology’ conceptually entails a kind of ‘conditional introspective access’ to the mental state/event in question. I suggested that this reply is problematic as it is by no means obvious that ‘having a phenomenology’ conceptually entails the relevant epistemic notion—unless, trivially, ‘having a phenomenology’ is understood as being roughly semantically equivalent to ‘conditional introspective access’.

‘Conditional introspective access’ itself may be a notion we can use to define or characterize ‘phenomenal consciousness’. What are the psychological enabling conditions of introspective access? Well, at the very least, one must be able to have thoughts with the relevant psychological content, and it is plausible that certain non-human creatures and human beings lack this ability though they plausibly have phenomenally conscious mental states/events. If a mental event’s being ‘phenomenally conscious’ just means that its subject has conditional introspective

23
access to it, then it is wide open how this epistemic property is grounded by certain mental events and not others, which is plausibly a virtue of this approach.

At the very least, this approach has the virtue of clearly demarcating certain mental states/events that are unanimously considered unconscious from those that are clearly phenomenally conscious. On a closely related alternative, ‘conditional introspective access’, semantically determines/fixes the reference of ‘phenomenal consciousness’. In other words, ‘phenomenal consciousness’ refers to the property such that mental events which instantiate it are conditionally introspectively accessible.

Ultimately, I don’t think much of substance hinges on whether we adopt the reference-fixing over the meaning-giving versions of the introspective access characterization. But I will assume the more familiar reference-fixing version in what follows. A substantive question for this way of characterizing ‘phenomenal consciousness’ is whether it is correct to consider all mental states/events that aren’t conditionally introspectively accessible as being phenomenally unconscious. Certain prominent theorists, such as Ned Block, have held that there may be phenomenally conscious states/events that are not conditionally introspectively accessible. We will confront this issue later on.

1.6: A loose end

Before turning to examine a prominent alternative characterization of ‘phenomenal consciousness’—what I will call (in line with the literature) the ‘higher-order’ characterization, there is a loose end I’d like to tie up here. A proponent of the has-a-phenomenology characterization might object that, with minimal supplementation by way of some positive examples, the has-a-phenomenology characterization is in much better shape than I have claimed.
Consider the following passage:

A mental state is conscious when there is something it is like to be in that state. Conscious states include states of perceptual experience, bodily sensation, mental imagery, emotional experience, occurrent thought, and more. There is something it is like to see a vivid green, to feel a sharp pain, to visualize the Eiffel tower, to feel a deep regret, and to think that one is late.\(^4\)

This passage is from David Chalmers’ (2010a) well-known paper, “Consciousness and its place in nature”. It represents a common strategy for characterizing phenomenal consciousness: the ‘something it is like...’ descriptive content supplemented by some positive illustrative examples. How does such an attempt at zeroing in on phenomenal consciousness fare? No better than without the examples, I think. A major reason is that ‘seeing a vivid green’ or ‘thinking that one is late’ needn’t be phenomenally conscious mental states/events: one can unconsciously see a vivid green and unconsciously think that one is late. Thus, such cases cannot serve as illustrative examples of phenomenally conscious mental states/events.

Bodily sensations and mental imagery are tougher to assess: can there be an unconscious pain or an unconscious mental image? Theorists have clashed on this issue, and we can’t settle the matter here. But even if we grant that bodily sensations and mental imagery are examples of phenomenally conscious mental states/events, it won’t do to have only these in our arsenal of examples, as they represent only a highly circumscribed class of mental states/events that we want to be able to confidently ascribe phenomenal consciousness to.

What about talk of ‘perceptual experience’ and ‘emotional experience’ in the above passage? Do these notions give us any further purchase on our target phenomenon? No, ‘experience’ is just roughly synonymous with ‘phenomenal consciousness’, and it raises the same questions as ‘phenomenal consciousness’ does

\(^4\) Chalmers, 2010a, p.104.
in terms of how we are to understand it. Seeing something is an example of an
‘experience’ as far as ordinary English is concerned, but clearly this is not how
‘experience’ is to be understood in the consciousness studies literature, since, in this
context, seeing something needn’t be an ‘experience’ at all, as in the case of
blindsight perception.

One might have better luck in this area by appealing to perceptual
appearances as examples of phenomenally conscious mental states. Looking round to
a subject S, in the relevant (non-comparative and non-epistemic) sense of ‘looks’,
might be such that some of us will have a hard time envisaging this relation or
pseudo-relation being instantiated in the case of blindsight perception, for example.
Again, I cannot settle the issue here, but this is a more committal approach to the
sorts of mental states/events that can be phenomenally conscious—and that’s
probably not a good thing when it comes to finding a theory-neutral starting point for
theorizing in this area.

What’s more, if we hold that perceptual appearances, bodily sensations, and
mental imagery are all examples of phenomenally conscious mental states/events,
then it begins to seem as though the ‘something it is like...’ descriptive content is an
idle wheel: we seem to simply be on the lookout for all those mental state/event-
types whose tokens are plausibly always such that they are conditionally
introspectively accessible, and that will complete our inventory of ‘phenomenally
conscious’ mental events.

Perhaps some will find this particular sort of ‘kind/type’ view of phenomenal
consciousness an adequate starting point for theorizing about phenomenal
consciousness, but it is obviously controversial (I will argue later that such views are
to be rejected). Aside from, in effect, construing phenomenal consciousness as a
determinable mental event-type, this understanding of ‘phenomenal consciousness’ apparently leaves ‘occurrent propositional attitudes’ out in the cold. Both commitments are highly substantive, and so a characterization of ‘phenomenal consciousness’ along such lines is unlikely to be one we can rest content with at this early stage.

Assuming that these mental event-types are just examples of phenomenal consciousness, and not all we should be interested in when it comes to theorizing about ‘phenomenal consciousness’, then not only are they still fairly circumscribed, but also potentially misleading. Tokens of the mental event-type ‘seeing a vivid green’, for example, shouldn’t be ruled out as candidates for being phenomenally conscious, even though it is highly plausible that not all tokens of this event-type will be phenomenally conscious mental events. Boyer, Harrison, and Ro (2005), for example, used transcranial magnetic stimulation to produce a visual condition akin to (type 1) blindsight in which both figure orientation and color plausibly were unconsciously perceived—based on results from a forced-choice experimental paradigm.

Ultimately, then, I think that there aren’t any particularly helpful illustrative examples we can provide of phenomenal consciousness. This is obviously a rather surprising claim, but it seems to be borne out by the fact that we usually end up referring to mental state/event-types as examples, and these are simply unhelpful here. If the event-type we refer to is such that all of its tokens (at the actual world anyway) are phenomenally conscious, then this suggests an understanding of phenomenal consciousness that is arguably too narrow. But if we refer to a mental event-type that is not this way, such as ‘seeing a vivid green’, then this is also of no
use because we need to draw attention to only some tokens of this type, not the type itself.

Perhaps this is why a relatively common informal tactic in this area is to try to get the audience to attend to a phenomenally conscious token mental event like my current visual awareness of the sun from my office window. Aside from substantive indeterminacy and under-determination worries for this approach, it’s unclear whether we should agree that phenomenal consciousness is a property of a mental event that we can pay ‘inner’ attention to in the first place.

Moreover, even if we assume that phenomenal consciousness is a manifest feature of our mental lives, not merely an enabling feature of attention to our mental lives, being able to recognize this feature from one mental event token to another is of little help to us here unless we can devise some sort of reasonably informative and determinate publicly accessible means of pointing to it. And our discussion thus far may only serve to reinforce the impression that if we can, indeed, pay attention to the phenomenally conscious aspect of the relevant token mental events, then the upshot of this is just a Wittgensteinian ‘beetle in a box’ that resists all attempts at direct communication. This loose end pulled us off into some substantially new territory, and we haven’t seen the last of it. For now, we press on to examine the ‘higher-order’ characterization of phenomenal consciousness.

2.1: The higher-order characterization

Higher-order approaches to phenomenal consciousness have a long and distinguished history. Here I am only concerned with a subset of these approaches, namely, those that offer a higher-order characterization or definition of ‘phenomenal consciousness’. Many who advocate for a higher-order theory of phenomenal consciousness are less than forthcoming about the issue of whether it is a kind of
conceptual truth that a ‘phenomenally conscious’ mental state/event is a mental state/event of which its subject is aware.

William Lycan is an important exception in this regard. As long as we are careful to distinguish between the ‘phenomenological’ aspects or ‘phenomenal character’ of a phenomenally conscious mental state/event, and a mental state/event’s being ‘phenomenally conscious’ at all, Lycan defines a ‘phenomenally conscious’ mental state/event as a mental state/event its subject is aware of. Lycan (2001) provides an instructive example of how such a premise can figure in an argument for a higher-order representational theory:

(1) A conscious state is a mental state whose subject is aware of being in it. [Definition]
(2) The ‘of’ in (1) is the ‘of’ of intentionality; what one is aware of is an intentional object of the awareness.
(3) Intentionality is representation; a state has a thing as its intentional object only if it represents that thing.
Therefore,
(4) Awareness of a mental state is a representation of that state. [2,3]
And therefore,
(5) A conscious mental state is a state that is itself represented by another of the subject’s mental states. [1,4] QED

An obvious problem with the above argument has been noted by Mikkel Gerken (2008), namely, that (5) simply does not follow from the stated premises; nothing rules out a self-representing mental event in contrast to two distinct mental states/events with the one representing the other. But if we stick with (4) as the sole conclusion, then we have a valid argument. An interesting issue with the argument, so understood, concerns the status of (2). When Lycan (2001) claims that ‘The ‘of’ in (1) is the ‘of’ of intentionality’, it’s not clear whether he intends this as a claim about the meaning of ‘aware of’ in (1) or not. If it’s a claim about the meaning of

---

‘aware of’ in (1) then it seems to be dubious: The intentional sense of ‘aware of’ isn’t clearly the only candidate here—what about a non-intentional ‘acquaintance’ sense of ‘aware of’? If it’s not a claim about the meaning of ‘aware of’ in (1), then (2) amounts to the substantive claim that, as a matter of fact anyway, we aren’t acquainted with any of our own mental states/events. Given all this, it is somewhat surprising that Lycan (2001) writes, “Premiss (2) seems hard to deny; I can think of no very plausible argument against it.”

2.2: Epistemic versus non-epistemic higher-order awareness

Finally, there is a larger issue or set of issues facing Lycan’s (2001) first premise. Should a variant of (1) framed in terms of fact-awareness be considered interchangeable with it? Perhaps not, but we then need a reason for privileging ‘thing-awareness’ versions over fact-awareness ones. What’s more, fact-awareness versions raise the possibility an epistemic definition or characterization of ‘phenomenal consciousness’ in terms of conditional introspective access. Let’s spell this out. A mental event $e$ is ‘phenomenally conscious’=df. $e$’s subject is aware of the fact that they have $e$. If we restrict to those with the ability or capacity to

---

6 Lycan, 2001, p.4.
have introspective knowledge, then this unpacking is just a version of a conditional introspective access *definiens* of a ‘phenomenally conscious’ mental event.

**2.3: A different argument for an epistemic understanding of higher-order awareness**

David Chalmers (1996) and Charles Siewert (1998) both argue for a close relative of the above point, namely, that ‘a phenomenally conscious mental event is one that its subject is aware of’ can be interpreted as saying that conscious mental events are such that they are, in our terminology, conditionally introspectively accessible. Both Chalmers (1996) and Siewert (1998) assume that we have an independent grip on ‘phenomenal consciousness’ and that this notion is such that there is a conceptual entailment from ‘being a phenomenally conscious mental event’ to ‘being conditionally introspectively accessible’. We examined this idea in the context of the has-a-phenomenology characterization of a ‘phenomenally conscious’ mental event, and found this purported conceptual entailment to be lacking support.

The point I wish to draw attention to here, however, is that the relevant notion of ‘awareness’ in this context is ‘awareness of’, and both Chalmers (1996) and Siewert (1998) still find it plausible to understand this epistemically. This suggests that, whether stated in terms of ‘awareness that’ or ‘awareness of’, it is natural to understand the higher-order ‘slogan’ in epistemic terms. We have been interested in the higher-order slogan as a definition of a ‘phenomenally conscious’ mental event. Neither Chalmers (1996) nor Siewert (1998) consider that the slogan might be interpreted as a definition of ‘phenomenal consciousness’ in terms of conditional introspective accessibility. This is likely, as noted earlier, because they both think we have an adequate grip on the notion of a
‘phenomenally conscious’ mental event that is independent of such epistemic notions. For this reason, they think that the ‘analytic ring’ to the higher-order slogan can be explained by the fact that there is a conceptual entailment from ‘phenomenal consciousness’ to ‘conditional introspective accessibility’.

I would argue, instead, that the ‘analytic ring’ to the higher-order slogan is plausibly accounted for by the conditional introspective access understanding of ‘phenomenal consciousness’. Moreover, the fact that both the has-a-phenomenology and higher-order characterizations of ‘phenomenal consciousness’ naturally admit of this sort of epistemic interpretation provides further support for it. Having said that, we should now consider the prospects of a non-epistemic interpretation of the higher-order characterization of a ‘phenomenally conscious’ mental state/event.

2.4: The non-epistemic higher-order characterization

A non-epistemic understanding of the higher-order characterization is clearly available. It is also clear, however, that such a definition or characterization of a ‘phenomenally conscious’ mental event would be quite committal—especially in light of the alternative epistemic understanding. If we could rule out the conditional introspective access understanding in this context, then the force of the higher-order slogan might begin to push us in the direction of such a controversial starting point for theorizing about phenomenal consciousness. Suppose, for the sake of argument, that the epistemic interpretation fails, and that the related ‘deflationary’ understanding of the slogan suggested by Chalmers (1996) and Siewert (1998) doesn’t hold up—perhaps because their background assumption that we grasp the meaning of a ‘phenomenally conscious’ mental event independently of the higher-order slogan is incorrect. Then it seems we might, indeed, be left to wonder about
non-epistemic readings of ‘aware of’ in the *definiens* of phenomenal consciousness—on pain of failing to address the force of the higher-order slogan.

To summarize the issue in a slightly different way, if we reject the Chalmers (1996) and Siewert (1998) line that the ‘is’ in the slogan is the ‘is’ of predication and hold, instead, that it is the ‘is’ of definition (as seems fairly plausible if the Chalmers (1996) and Siewert (1998) line fails), and we can rule out the conditional introspective access reading of the right-hand side of the slogan, then, given the need to account for the slogan’s apparently platitudinous status, we arrive at a non-epistemic higher-order awareness *definiens* of a ‘phenomenally conscious’ mental event: a mental event $e$ is ‘phenomenally conscious’ =df. $e$’s subject is occurrently and non-doxastically aware of $e$. But we aren’t forced into either disputing the apparently platitudinous force of the higher-order slogan or accepting a non-epistemic higher-order awareness characterization of ‘phenomenal consciousness’. There is no compelling reason, I think, for rejecting the conditional introspective access interpretation in this context.

An alternative strategy would be to look for another working definition of ‘phenomenal consciousness’ which could then be used to run the strategy Chalmers (1996) and Siewert (1998) suggest for deflating the force of the slogan. I ultimately don’t think that this strategy will work, but we haven’t yet exhausted all major options for characterizing a ‘phenomenally conscious’ mental state/event. In any case, I won’t dwell any further on the issue of characterizing ‘phenomenal consciousness’ in higher-order terms, since the choice in this context is between the relevant epistemic interpretation, which is somewhat controversial as a definition of ‘phenomenal consciousness’, and a non-epistemic interpretation, which is obviously highly controversial as a definition of ‘phenomenal consciousness’. The topic of
higher-order approaches to phenomenal consciousness will come up again later, but I think we can safely move on from higher-order characterizations of ‘phenomenal consciousness’ at this point.

2.5: ‘Phenomenal consciousness’ and ‘phenomenal qualities’

Those familiar with the consciousness studies literature may be wondering at this point: why can’t we understand a ‘phenomenally conscious’ mental state/event in terms of those familiar properties of conscious mental events such as what it’s like to experience a red tomato? That is, why can’t we use examples such as ‘phenomenal redness’ to zero-in on the relevant mental states/events? One problem is that the notion of ‘phenomenal redness’ is contentious if we take it to refer to a property that cannot be had by an unconscious mental state/event. And we need to assume this if ‘phenomenal redness’ or the like is to be of any help to us in understanding phenomenal consciousness.

But it would seem that even if there are properties that are the preserve of phenomenally conscious events, we are still dealing here with phenomena that are importantly distinct from our target phenomenon. So, a stronger sort of claim is needed here if the idea is that these properties are good enough to introduce phenomenal consciousness itself. If we assume that *phenomenal consciousness a kind of mental event*, then ‘phenomenal redness’ will be something like a determinable of phenomenal consciousness and so we get some traction here in claiming that such phenomenal properties may be used to introduce phenomenal consciousness itself. I will argue later on, however, that this sort of view is to be rejected.

Having now examined a number of different strategies for characterizing ‘phenomenal consciousness’, I hope it is clear that finding an appropriate starting
point for theorizing in this area is quite challenging—more so, anyway, than many seem to have appreciated. It would be a lot easier to get started if we were to simply focus on what we take to be certain interesting aspects of phenomenally conscious mental states/events. But we shouldn’t abandon the project of characterizing a ‘phenomenally conscious’ mental state/event just because it is onerous and some or most of us feel confident that we are talking about the same thing, or occupying a similar ‘logical space’ in terms of what the notion entails, and so on.

‘Phenomenal consciousness’ may, broadly speaking, play the same ‘conceptual-role’ for most who theorize about it, but there are certainly some relatively minor differences here between theorists that can make a relatively major difference viz. theorizing about phenomenal consciousness. Having said that, let us now consider an important area of widespread agreement in the ‘conceptual-role’ of ‘phenomenal consciousness’: the ‘explanatory gap’ that it (allegedly) gives rise to.

Can the ‘explanatory gap’ phenomenon further elucidate or help us zero-in on ‘phenomenal consciousness’ in a way we have yet to consider?

3.1: The explanatory gap constraint

It is widely held that phenomenal consciousness gives rise to an ‘explanatory gap’ (see Levine, 1983, and Chalmers, 2010b) that a physical *explanans* cannot cross. The exact details of this claim are somewhat contentious, but the basic idea is that narrowly and broadly physical truths are either structural or functional truths, and truths about structure and function can only ever explain other structural or functional truths. Thus, if ‘phenomenal consciousness’ is not a structural or functional notion, then truths about phenomenal consciousness will not be explicable in physical terms—hence the ‘explanatory gap’ between the physical and the phenomenal. There is obviously a good deal to clarify and discuss when it comes to
this line of argument, but for now the focus is on the claim that ‘phenomenal consciousness’ is not a structural or functional notion. We are primarily interested here in whether a claim like this can further elucidate or help us zero-in on our target phenomenon in a new way.

Though unpopular nowadays, some maintain ‘phenomenal consciousness’ is ultimately a functional or dispositional notion. A rough example of such a view would be that a ‘phenomenally conscious’ mental event simply means a mental event that tends to cause certain other mental phenomena and/or behavior, and may itself tend to be caused by certain mental phenomena and/or the environment, depending on the type of mental event in question (i.e. perceptual or cognitive and so on). In short, ‘phenomenally conscious’ simply means ‘causal-role R’ where ‘R’ is schematic for the broad sort of ‘causal-role’ canvassed above. ‘Analytic functionalist’ views along such lines are apparently ruled out by what we can call the ‘explanatory gap constraint’ on understanding ‘phenomenal consciousness’, according to which ‘phenomenal consciousness’ is not structural or functional notion—we cannot define ‘phenomenal consciousness’ in structural or causal-role terms.

The explanatory gap constraint tells us that ‘phenomenal consciousness’ is not a structural or functional notion, which is to say that for any structural or functional notion one might propose here, this is not what we mean by a ‘phenomenally conscious’ mental state/event. Why think this is so? Well, it has been suggested that for any functional/causal-role definition of ‘phenomenal consciousness’ one will always be in a position to coherently ask: is the mental event with the relevant functional/causal-role phenomenally conscious? And if this is correct, then we can’t define ‘phenomenal consciousness’ in functional terms, since
a successful definition is one where the *definiens* conceptually entails the *definiendum*.

But why think that *for any current or future* proposed functional/causal-role characterization of ‘phenomenal consciousness’ we can coherently ask the relevant question? And why think that our understanding of ‘phenomenal consciousness’ can even properly bear the weight of such a question? In short, what reason is there to even accept the explanatory gap constraint on a characterization of ‘phenomenal consciousness’?

No doubt many theorists find such a constraint plausible, but to secure the claim that all extant and future functional analyses of ‘phenomenal consciousness’ will fail, however, it seems plausible that one must point to some positive aspect of our understanding of ‘phenomenal consciousness’ that conceptually entails this claim. But this leads into troubled waters. For example, even though it would secure the relevant entailment, we don’t want to hold that it is *definitional* of ‘phenomenal consciousness’ that it refers to an intrinsic, non-relational, property. I’m inclined to think, then, that the explanatory gap constraint is *too strong*.

3.2: Weakening the explanatory gap constraint

There are weaker versions of the explanatory gap available. On one such version, while all current or extant functional analyses of ‘phenomenal consciousness’ fail, some future functional analysis may yet succeed. This is also quite strong: perhaps there is some extant functional analysis such that ‘functional-role R’ and not ‘phenomenal consciousness’ only *seems* coherent to us for now; further rational reflection will ultimately reveal that it is not. As non-ideal reasoners, this possibility may not be one we are in a position to foreclose. One might suggest, then, that the constraint be understood along the following lines: all extant functional
analyses of ‘phenomenal consciousness’ currently seem to fail. This even weaker version, however, may still be too strong for our purposes.

All we really need commit to, I think, is something along the following lines: all extant functional analyses of ‘phenomenal consciousness’ are such that it is currently unclear or inconclusive whether the relevant open-question has application. If, for any and every extant functional analysis, we are currently unable to tell whether or not ‘functional-role R’ and not ‘phenomenal consciousness’ is coherent, then this is still quite problematic for the analytic functionalist claim that there is some functional analysis R, such that ‘R’ and not ‘phenomenal consciousness’ is incoherent. The suggestion that there is some as of yet unconceived ‘R’ that will do the trick here seems forlorn. If, instead, it is suggested that some extant ‘R’ succeeds, then either the inconclusiveness claim must be challenged, or the analytic functionalist must resort to the equally forlorn idea that further rational reflection will eventually reveal that ‘R’ succeeds. Thus, this weakest version of the ‘explanatory gap constraint’ still has considerable bite.

The weakest version also seems to offer some insight into the epistemic situation many of us seem to find ourselves in with respect to currently available functional analyses of ‘phenomenal consciousness’. A functional analysis of ‘phenomenal consciousness’ is advanced, and while some immediately decry it as clearly inadequate, many of us, I think, simply don’t find it plausible because it is supposed to be a conceptual truth—and it is anything but clear to us whether this is so. And the burden of argument is supposed to lie with those who advance such an analysis—it is their job to persuade the rest of us of its cogency. So, if a proposed analysis in this regard doesn’t have an ‘analytic ring’ to it, and we find ourselves continuing to wonder whether or not the relevant sort of ‘open-question’ applies to
it, then that should be reason enough to reject the proposed analysis—at least until some further consideration(s) in its favor is advanced.

3.3: The no-currently-obvious-reductive-analysis constraint and introspective justification

On balance, then, I’m inclined to think we should only demand that the weakest of the explanatory gap constraints we’ve examined be satisfied by a characterization of ‘phenomenal consciousness’. On this point, I seem to be broadly in agreement with Eric Schwitzgebel (2016) who writes that,

...phenomenal consciousness should retain at least a superficial air of mystery and epistemic difficulty, rather than collapsing immediately into something as straightforwardly deflationary as dispositions to verbal report, or functional ‘access-consciousness’ in Block’s (1995/2007) sense, or an ‘easy’ problem in Chalmers’ (1995) sense. If the reduction of phenomenal consciousness to something physical or functional or ‘easy’ is possible, it should take some work. It should not be obviously so, just on the surface of the definition.  

Schwitzgebel (2016) calls this the ‘wonderfulness condition’ on a satisfactory characterization of ‘phenomenal consciousness’. I will call my weakest explanatory gap constraint the ‘no currently obvious reductive analysis constraint’. Wordier, I know, but I think it’s more important, for our purposes anyway, to be specific as to the nature of this constraint. Now, does the no-currently-obvious-reductive-analysis constraint suggest any novel way of characterizing ‘phenomenal consciousness’? Not that I can discern. Does it help narrow our field of characterizations of ‘phenomenal consciousness’? I don’t think so, though certain epistemological accounts of introspection may not satisfy this constraint.

The issue is somewhat complex, but so long as ‘introspective access’ is understood to involve epistemic justification, and we understand ‘epistemic

---

7 Schwitzgebel, 2016, p.225.
justification’ along normative lines, then ‘conditional introspective access’ will plausibly meet the relevant constraint. Irrespective of one’s views concerning the nature of knowledge (e.g. one might hold that knowledge is just the most general factive attitude and doesn’t involve epistemic justification), it is quite plausible that judgements about our own current mental states/events are often epistemically justified in a distinctly first-personal way. What this ‘distinctly first-personal way’ amounts to is the subject of considerable disagreement, but there is widespread agreement that there is such a way(s).

Regarding ‘epistemic justification’, it is not particularly controversial to assume that this is a normative notion, though it is sometimes suggested that it is actually an evaluative notion. In maintaining that ‘introspective justification’ is a normative notion, however, I do not commit to a deontic view according to which the normative is not reducible to the evaluative. Perhaps a notion such as ‘introspective justification’ or ‘epistemic justification’ in general, can ultimately be analyzed in terms of promoting some epistemic value(s) such as true belief. I am doubtful of this, but, in any case, this will be a highly substantial piece of analysis that is no threat to my suggestion that ‘introspective justification’ is a normative notion. All I intend here is that ‘epistemic justification’ and ‘introspective justification’ are, at least, prima facie normative.

As stated above, if we agree that there is such a thing as introspective justification, and that ‘epistemic justification’ in general seems to be a normative notion, then I think it’s plausible that there will be no ‘easy’ or obvious functionalist analysis available for the notion of ‘conditional introspective access’. We will discuss the issue in more depth in the next chapter. In what follows, I return to a major fault-line in understanding ‘phenomenal consciousness’, and draw some
morals for theorizing. After this, I build on the points and arguments developed in this chapter to render a final verdict on how we should understand the explananda of a ‘theory of phenomenal consciousness’.

4.1: The kind/type versus predicate view of phenomenal consciousness

As we saw earlier on, the phrase, ‘phenomenally conscious mental event’ may be understood as referring to a type or kind of mental event, or as referring to a mental event and the property of being phenomenally conscious. The view that ‘phenomenal consciousness’ refers to a property, whether inessential or not, differs from the view that ‘phenomenal consciousness’ refers to a (psychological) kind. Even assuming that kinds are universals, it may be argued that they, nevertheless, differ from properties. And if they are ultimately reducible to properties, then we still need to bear in mind the important difference between holding that phenomenal consciousness is an essential property of a mental event, and holding that phenomenal consciousness is the essence of a mental event. Only latter captures the modal force of kind view, though it is also, perhaps, stronger than the kind view needs to be.

The kind view arguably need only entail that, at the actual world, a mental event that lacks the property of being phenomenally conscious can’t subsequently acquire it and be numerically the same mental event, and vice-versa. Only a ‘sortal essentialist’ view here entails that a mental event that is phenomenally conscious is phenomenally conscious at every counterfactual world at which it exists. The metaphysical issue of the distinction between kinds and properties is certainly fraught, and I can’t hope to adequately address it here. But I don’t think we need to make any major decisions in this regard so long as we note that any ‘reductionist’
view of kinds to properties must be to properties (or instantiated properties?), that have the relevant modal force.

In his important paper, “On a confusion about a function of consciousness”, Ned Block (1995) suggests that a ‘phenomenally conscious mental event’ refers to a mental event-type/kind. He contrasts this with his now well-known ‘access-conscious’ sense of ‘conscious’, claiming that this refers to an inessential property of a mental event. Block writes,

A third difference [between ‘phenomenally conscious’ and ‘access-conscious’ mental states/events] is that there is such a thing as a P-conscious [phenomenally conscious] type or kind of state. For example, the feel of pain is a P-conscious type — every pain must have that feel. But any particular thought that is A-conscious [access-conscious] at a given time, could fail to be accessible at some other time, just as my car is accessible now but will not be so later when my wife has it. A state whose content is informationally promiscuous now may not be so later.\(^8\)

Block directs us to a particular type of experience, namely, pain experience, to make his point, but he presumably also holds that ‘experience’ itself refers to a general type or kind of psychological event—and this is really what is at issue here.\(^9\) Could a token experience \(e\) actually persist without being an experience? And could a token psychological event \(m\) lack the property of being an experience at a time \(t_1\) and then have the property of being an experience at \(t_2\)? The matter seems far from clear, and even if we are inclined to side with Block and others in holding that experiences are a distinctive kind or type of psychological event, this is certainly not a claim that we can reasonably just accept without argument.

\(^8\) Block, 1995, p.232.

\(^9\) It’s true that Block only says, ‘\textbf{there is such a thing} as a P-conscious \textit{type or kind} of state’, which technically leaves it open that phenomenal consciousness is sometimes a type or kind and sometimes not, but this seems incoherent. Using a bit of charitable interpretive license, then, I attribute to Block (1995) the view that phenomenal consciousness, in general, is a kind/type.
For one thing, it should not be a trivial consequence of the very notion of ‘experience’ or ‘phenomenal event’ that functionalism about phenomenal consciousness is false. But if it is analytic that ‘experience’ or ‘phenomenally conscious mental event’ refers to a distinctive psychological event-type, then functionalist views of phenomenal consciousness are ruled out \textit{a priori} as resting on the ‘category mistake’ of understanding phenomenality to be an \textit{inessential} property that a psychological event could have at a time $t_1$ and subsequently lack at $t_2$, and vice-versa.

If the kind view of phenomenal consciousness only ruled out \textit{analytic} functionalist views of phenomenal consciousness, that might be acceptable, but it should not rule out functionalism about phenomenal consciousness in general. This is one reason for thinking that the kind view of phenomenal consciousness is, at the very least, not a good starting point for our discussion. I will argue shortly that, in conjunction with some additional reasons against it, the case against the kind view is sufficient for us to reject it here.

\textbf{4.2: The scope of the divide.}

The problem generalizes to higher-order approaches to phenomenal consciousness, according to which a mental event $e$’s being phenomenally conscious consists in $e$ being the object of a certain sort of awareness by $e$’s subject. If a mental event’s being the object of a certain sort of awareness by the subject of the event types that event, then, as David Rosenthal (2005) has often noted, we get little to no explanatory purchase on phenomenal consciousness. Instead, we are left to wonder why only some mental events are tokens of the type ‘object of meta-mental awareness’.
Understandably, then, higher-order theories of phenomenal consciousness reject the view that phenomenal consciousness is a mental event-type. And higher-order theorists such as Rosenthal (2005) typically do, in a way, argue against the kind/type view of phenomenal consciousness. They usually focus on a paradigmatic ‘hard case’ such as pain, and argue that a token pain may be unconscious at time t₁ and phenomenally conscious at a time t₂ and vice-versa. This is admirable—though, of course, highly contentious. Higher-order theorists such as Lycan (2001) who adopt a definitional higher-order approach are in danger of assuming ‘the predicate view’ without argument: the view that phenomenal consciousness is an inessential property of mental events.

4.3: Why the kind view of phenomenal consciousness should be rejected.

If we assume that phenomenal consciousness is a kind/type of mental event, then it follows that familiar psychological event-types such as visual awareness and pains aren’t actually unified mental event-types at all such that a phenomenally conscious episode of visual awareness could subsequently be phenomenally unconscious or vice-versa: if phenomenal consciousness types mental event tokens, then when a token ceases to be phenomenally conscious it thereby ceases to exist and thus does not remain a token of some independent mental event-type such as visual awareness.

So, on the kind view of phenomenal consciousness, all familiar mental event-types are actually divided into phenomenal and non-phenomenal counterparts, such that there is no metaphysical category of ‘visual awareness’ per se that may be phenomenally conscious or not. Instead, there is phenomenal visual awareness vs.
unconscious visual awareness, *qua* opposing types of mental events. This sort of view may have some merit, but it is problematic.

First, these ‘phenomenal event-types’ must be ‘phenomenally conscious’ in the sense that they are conditionally introspectively accessible. In virtue of what are these phenomenal event-types such as ‘phenomenal judgement’, such that their tokens are all conditionally introspectively accessible? There does not seem to be any good answer to this question now that we have foreclosed the possibility of phenomenal consciousness itself being an independent property capable of playing the relevant epistemic role.

Second, it is plausible that when it comes to at least some familiar mental event-types such as judgements, a token judgement may be unconscious and then conscious, and vice-versa. But on this view, judgements, as we come to learn about them through introspection, are all tokens of the kind phenomenal consciousness, such that when suppose ourselves to be thinking about these events happening without being conditionally introspectively accessible, we aren’t actually thinking about the relevant events after all—and this seems wrong.

So, for at least a couple of reasons (besides the fact that this sort of view basically rules out functionalist and higher-order views of phenomenal consciousness), I think that the kind view should be rejected. This is not to say that pains, for example, aren’t essentially conscious, though I will argue against this view later on as well. All I am saying here is that pain, as we understand it through introspection, is *not* a sub-kind or determinate of the kind, phenomenal consciousness; there plausibly is no such ‘determinable’ kind.
5.1: A summary of the argument thus far.

Let us take a moment to summarize the major conclusions we have arrived at thus far and some of the reasoning behind them. It was argued that the commonplace ‘something it is like...’ or has-a-phenomenology characterization of a ‘phenomenally conscious’ mental state/event is inadequate unless it is understood in terms of ‘conditional introspective access’. The driving force behind this argument was the notion that only such an epistemic reading of the ‘something it is like...’ frame clearly renders the verdict that (type 1) blindsight perception is, indeed, a paradigmatic example of an unconscious mental state/event-type.

After considering a typical, but unsuccessful, attempt from the literature that combines ‘positive examples’ of phenomenal consciousness with the ‘something it is like...’ frame, we then turned to examine the prospects of characterizing ‘phenomenal consciousness’ in terms of higher-order awareness. It was argued that the higher-order definition of ‘phenomenal consciousness’ is best interpreted as an epistemic characterization of ‘phenomenal consciousness’ along the lines of conditional introspective accessibility. It was also pointed out that the higher-order approach to characterizing ‘phenomenal consciousness’ has the weakness that the higher-order slogan needn’t be understood as providing a definition. But in the absence of a satisfactory alternative definition of ‘phenomenal consciousness’, the interpretation of the slogan as providing a definition is difficult to resist.

The final popular alternative we considered for characterizing ‘phenomenal consciousness’ was what we might call the ‘phenomenal qualities’ characterization. We ultimately rejected this attempt to define ‘phenomenal consciousness’ in terms of certain properties of such states/events since it requires the controversial theoretical commitment that such properties are proprietary to those mental states/events that
are phenomenally conscious. In addition, I argued that even if there are properties that can only be had by phenomenally conscious mental states/events, these properties do not plausibly furnish us with an adequate grip on phenomenal consciousness itself.

After considering all of these strategies for providing a direct positive characterization of ‘phenomenal consciousness’, we discussed an adequacy constraint on such definitions/characterizations of ‘phenomenal consciousness’: the ‘explanatory gap constraint’. It was argued that, at least in the context of searching for an adequate characterization of ‘phenomenal consciousness’, the explanatory gap constraint is open to serious challenge and there is good cause for thinking it too strong to serve as an adequacy constraint here. We then explored a number of ways of weakening the explanatory gap constraint and found that the weakest version, the no-currently-obvious-reductive-analysis constraint, is fairly plausible. The no-currently-obvious-reductive-analysis constraint did not suggest any additional way or ways of characterizing ‘phenomenal consciousness’, nor did any of our previously discussed characterizations of ‘phenomenal consciousness’ seem to violate it.

We then moved on to a largely unacknowledged issue in the consciousness studies literature: kind versus property characterizations of phenomenal consciousness. I argued that kind characterizations are a poor choice for a starting point in examining phenomenal consciousness as they are basically incompatible with popular theories of phenomenal consciousness such as functionalist and higher-order theories. Rounding out the discussion, I then argued that kind views of phenomenal consciousness should simply be rejected altogether for some additional reasons. Let us now proceed to the final tasks of this chapter.
6.1: An argument for the conditional introspective access characterization

The only characterization of ‘phenomenal consciousness’ that has stood up to scrutiny thus far is the conditional introspective access characterization. To be clear, the conditional introspective access characterization has it that a mental event $m$’s being ‘phenomenally conscious’, refers, as a matter of semantic rule, to $m$ having a certain property such that mental events that have this property are conditionally introspectively accessible.

‘Introspection’, as I understand it, is a relatively ‘thin’ notion that can’t be straightforwardly analyzed in functional and/or dispositional terms. And many thick understandings of ‘introspection’, such as those along ‘evidentialist’ lines, very plausibly meet the no-currently-obvious-reductive-analysis constraint. At this point, I want to generalize an earlier line of argument for the conditional introspective access characterization.

Aside from certain other problems for the has-a-phenomenology characterization, such as the plausible sounding suggestion that only conscious mental states/events have a phenomenology, I argued earlier that it struggles to deliver the desired verdict that blindsight perception is a clear case of a phenomenally unconscious mental state/event. This stands in contrast to the conditional introspective access characterization which clearly delivers the desired verdict. Though I only raised the issue in the context of the has-a-phenomenology characterization, there is reason to think that the problem generalizes to other characterizations we have considered as well.

Consider, for example, a non-epistemic higher-order characterization of ‘phenomenal consciousness’. Where’s the contradiction in supposing that the subject
of a blindsight perception is non-epistemically aware of their being in or having that mental state/event? It might be suggested that ‘non-epistemic higher-order awareness’ (or some form of it, such as higher-order ‘acquaintance’ awareness) conceptually entails something like ‘conditional introspective accessibility’, but here we run into the same difficulty we encountered earlier for the has-a-phenomenology characterization. If there is a conceptual entailment from the relevant notion of ‘phenomenal consciousness’ to ‘conditional introspective accessibility’ it is far from obvious.

6.2: ‘Access-consciousness’ versus ‘conditional introspective access’

To round out our argument here, I think we should touch on a ‘consciousness’ notion that, like ‘conditional introspective access’, also delivers the verdict that blindsight perception is phenomenally unconscious: Block’s (1995) notion of ‘access-consciousness’. If ‘access-consciousness’ is a purely dispositional notion, then it is not a suitable candidate for characterizing ‘phenomenal consciousness’ as it will violate the no-currently-obvious-reductive-analysis constraint. One might wonder whether one or more of the notions of ‘phenomenal consciousness’ we have discussed thus far conceptually entails ‘access-consciousness’, but I think it will be obvious that none plausibly do.

Here is a rough unpacking of ‘access-consciousness’: a mental state/event \( e \) is ‘access-conscious’=df. \( e \) is poised/available “...free use in reasoning and for direct “rational” control of action (including reporting).”\(^\text{10}\) Since the notions of ‘free-use’ and ‘rational’, the latter of which Block (2002) puts in scare quotes, give rise to certain troubling questions, we should probably prefer Christopher Hill’s (2009) alternative formulation that does away with them. Hill’s (2009) characterization runs

\(^{10}\) Block, 2002, p.208.
as follows: a mental state/event $e$ is ‘access-conscious’=df. $e$ “...is, potentially, at least, a maximally proximal causal trigger for several of the high level cognitive agencies that are recognized by folk psychology.”\textsuperscript{11} He goes on to helpfully clarify the details of this proposal as follows:

These agencies include the ones that are responsible for producing speech, forming belief and other propositional attitudes, making choices, exercising on-line control of intentional actions, creating memories, monitoring mental states, and producing introspective judgements. In saying that a mental event must be a \textit{maximally proximal} causal trigger for the agencies in question, I mean that it must be capable of doing causal work without being retrieved from memory, “refreshed,” or converted into a different form. In short, it must be capable of doing causal work without additional processing. Further, in saying that a mental event need only be capable of serving as a causal trigger for \textit{several} of the relevant agencies in order to count as an experience, [being ‘access-conscious’ in our terminology] I mean to allow for the fact that creatures can enjoy conscious experiences even though they lack one or more of them.\textsuperscript{12}

An important difference between Hill’s (2009) ‘access-consciousness’ and ‘conditional introspective accessibility’ is that the former does not privilege introspective access over other forms of cognitive access. This suggests that, even in a creature with a fully intact faculty of introspection, a mental state/event may be access-conscious without being introspectively accessible, so long as that mental state is available to several other cognitive agencies (not necessarily including any of these agencies in particular).

To my mind, this effectively rules out Hill’s notion of ‘access-consciousness’ as a way of characterizing ‘phenomenal consciousness’: a creature $C$ with a fully intact faculty of introspection will be such that phenomally conscious mental states/events in $C$ are introspectively accessible—this is a kind of ‘conceptual truth’ about ‘phenomenally conscious’ mental state/events at the actual world. ‘Access-consciousness’, then, is ultimately more grist for our mill that turns in favor of

\textsuperscript{11} Hill, 2009, p.12.
\textsuperscript{12} Hill, 2009, p.12.
‘conditional introspective access’ being the best of the lot when it comes to characterizing ‘phenomenal consciousness’.

6.3: A hypothetical characterization and conditional introspective access

Having said all this, I can’t rule out the possibility that there is some characterization of ‘phenomenal consciousness’ that clearly renders the correct verdict on blindsight perception and meets our two adequacy constraints. Since ‘rendering the correct verdict on blindsight perception’ is plausibly just another adequacy constraint, we can summarize the challenge for characterizing ‘phenomenal consciousness’ as one of meeting our two constraints (the no-obvious-reductive-analysis constraint, and this one). My argument thus far has been that only the ‘conditional introspective access’ characterization plausibly does so.

If there were a novel or overlooked characterization of ‘phenomenal consciousness’ that meets these two adequacy constraints, it seems likely from all the work that has transpired in this field so far that it would be a notion that bears a tight conceptual connection to ‘conditional introspective accessibility’. If the relevant notion were to conceptually entail and be entailed by ‘conditional introspective accessibility’, then working with ‘conditional introspective access’ as our definition of ‘phenomenal consciousness’ would be entirely satisfactory.

Even if only a conceptual entailment in one direction between ‘conditional introspective access’ and our hypothetical notion of ‘phenomenal consciousness’ were to hold, our understanding of ‘phenomenal consciousness’ in terms of conditional introspective access might still be largely satisfactory. Moreover, as I will now argue, proponents of extant alternatives to the conditional introspective access characterization plausibly don’t have anything to lose by adopting it. Chances are, this would also be true of our ‘hypothetical’ characterization.
6.4: Nothing to lose by accepting the conditional introspective access characterization?

As stated earlier, I think that the conditional introspective access characterization is quite nice in so far as it doesn’t prejudice the issue what sort of mental property might ground the relevant epistemic property. It is a point of agreement among most proponents of different approaches to characterizing ‘phenomenal consciousness’ that a mental state/event’s being phenomenally conscious does, in some way, ground an epistemic property along the lines of conditional introspective access. The fact that the conditional introspective access characterization accommodates this and is silent on the sort of mental property that does the grounding seems a virtue of this approach.

Thus, proponents of existing alternative characterizations, such as the has-a-phenomenology characterization, plausibly aren’t giving up anything substantial by opting instead for the conditional introspective access characterizing—as long as they hold that there is an entailment from phenomenal consciousness to conditional introspective accessibility (or some close relative). I mentioned in passing earlier on that Ned Block, in a series of papers, has come out against an entailment from phenomenal consciousness to various forms of cognitive access, including conditional introspective access. Though a full assessment of the quasi-empirical case that Block (2007) makes for his view is beyond the scope of our discussion, I also think it is not called for here, as we are in a position to ‘jump in’ at an early stage in his argument to offer a satisfactory rebuttal.
6.5: Block’s puzzle case for phenomenal consciousness without introspective accessibility

Block’s (1995; 2007) argument for phenomenal consciousness without ‘cognitive access’ takes off from Sperling’s (1960) partial-report paradigm experiment and recent variants thereof. Sperling (1960) presented subjects with a visual array consisting of three rows of four letters for a very brief duration (usually around 50 milliseconds). In the so-called ‘whole report’ condition, Sperling had subjects report as many items from the array as they could in their correct spatial positions. Under this condition, he found that subjects could accurately report around 4 letters. Sperling noted, however, that subjects typically reported having seen more letters than they could recall.

In an effort to determine whether or not this was so, Sperling devised his famous ‘partial-report’ condition in which one row of the array was cued for report after a short delay. Sperling found that subjects could, on average, report nearly all of the letters from the cued row. He inferred from this that, after array offset, subjects must have had roughly 9 letters of the array available to them, if only for a very short period of time. After a delay of about 1 second, Sperling found that the partial-report cue resulted in a level of task performance that calculated out to only around 4 letters being available for report—the same number of letters reported under the whole-report condition. Thus, Sperling had apparently discovered a high capacity rapidly decaying visual memory store.

Block (1995; 2007) argues on the basis of examples such as the Sperling task that phenomenal conscious ‘overflows’ cognitive accessibility. The Sperling task and others using the partial-report procedure show that we are often phenomenally aware of more than we can report at a given time. Block (1995; 2007) thinks that
they also suggest at least a *prima facie* case for the view that phenomenally conscious awareness can occur without cognitive access/accessibility. Though he doesn’t discuss *introspective* access specifically, it is clear that he takes cognitive access to include introspective access. I will argue that, on reflection, the Sperling task does not suggest that phenomenal consciousness can occur independently of (conditional) introspective access.

**6.6: The Sperling task doesn’t suggest phenomenal consciousness without introspective access**

Now, the mere fact that roughly 9 of the 12 letters are stored in a volatile and extremely short-lived memory system obviously does not establish that subjects in the Sperling task are phenomenally aware of 9 of the 12 letters. Nevertheless, I think it is rather plausible that subjects are, in fact, phenomenally aware of all 12 letters—in all their detail as well. I find it difficult to suppose otherwise as one who has ‘tried on’ the experiment. Is this a problem for the view that phenomenal consciousness entails conditional introspective access? I don’t think so. Nor do I think it is a problem for the conditional introspective access characterization of ‘phenomenal consciousness’.

In my view, there is no problem with the claim that all the letters—in all of their detail, are introspectively accessible to a subject in the Sperling task. For a mental state/event $m$, such as ‘seeing all twelve letters in the Sperling task’, to be introspectively accessible is for the subject of $m$ to know, or be in a position to know, that they are *currently* in $m$. The fact that subjects in the Sperling task can’t remember all of the letters after the array has been switched off in no way threatens this claim. Once the array of letters is switched off, the subject is no longer in the relevant visual state/event, so the only potential threat from this sort of case derives
from that short period of time when the subject is in the relevant visual state/event. Clearly, subjects can’t report all the details of the visual state/event they are in during that short time frame.

But does this suggest that the relevant visual state/event isn’t introspectively accessible? No, introspective knowledge doesn’t entail being able to express that knowledge—especially when one has less than a second to do so. Oftentimes we can robustly express the things we know through introspection, but when we are only in a certain mental state/event for a very short amount of time, this can be quite difficult, and we will then have to rely on our short-term memory. In this case, by the time we can issue a robust verbal report, we have already forgotten a good deal of what we saw when the array was flashed on screen for us for a short time.

At this point, I hope to have done enough to convince the reader that the conditional introspective access characterization of ‘phenomenal consciousness’ is one we can safely deploy as a starting point in our theorizing about phenomenal consciousness. In the upcoming last section of this chapter, I want to map out the logical geography of the ‘problem of consciousness’ in light of this starting point.

7.1: Three problems of consciousness

A major part of the problem of consciousness, or a ‘problem of consciousness’ in its own right, is the problem of ‘phenomenal character’: the problem of accounting for the many manifest properties of phenomenally conscious mental states. The problem is particularly acute if we think that there are at least some properties of phenomenally conscious mental states/events that are proprietary. For example, if we think that there is such a thing as ‘phenomenal intentionality’; a kind of intentionality that is only had by phenomenally conscious
mental states/events, then an account of ‘phenomenally unconscious’ intentionality will likely be of little help to us in accounting for its ‘phenomenal’ counterpart.

Certain properties of phenomenally conscious mental states/events such as ‘sensory redness’ are highly problematic regardless of whether or not we take them to be proprietary of phenomenally conscious mental states/events. If we assume that they aren’t ‘phenomenally proprietary’ so to speak, then we are really just dealing with the general metaphysical problem of ‘secondary qualities’. If we assume that ‘sensory redness’ or a ‘red experience’ is phenomenally proprietary, then we get a more traditional ‘problem of consciousness’, and it is a very confusing one to say the least.

The explanatory gap problem plausibly arises with respect to sensory qualities, irrespective of whether or not we take them to be phenomenally proprietary—an issue we will discuss in more detail in chapter 4. Bodily sensations are a class of mental event-types that raise unique problems of their own and addressing them is a must for any complete theory of phenomenal consciousness. I will discuss pain in some depth in chapter 3—though I will not be attempting anything like a complete theory of phenomenal consciousness in this thesis.

Certain sub-problems within the problem of the manifest properties of phenomenally conscious mental events tend to lead us back to the problem of phenomenal consciousness itself. We want to know what this property is such that the relevant mental events are conditionally introspectively accessible. An account of what phenomenal consciousness is that doesn’t explain why such mental events are conditionally introspectively accessible is bound to have a certain air of mystery surrounding it. It would be neat if an account of the nature of phenomenal
consciousness made this connection intelligible, but I will argue in chapter 4 that this is probably forlorn.

The final significant component of ‘the problem of consciousness’ is the problem of conditional introspective accessibility itself. I have been rather vague about the details of conditional introspective accessibility, as I wanted to operate with a notion that most in this area would find acceptable. Filling-in at least some of the many details here in a way that is both substantive and compelling is an important task for adequate theory of consciousness—especially one such as ours that makes such heavy use of the notion. Call this the ‘details of introspection problem.’

These three problems, and their many sub-problems, will occupy us for the remainder of this thesis. Since, as noted earlier, I am not attempting anything like a comprehensive theory of consciousness, some problems will, somewhat arbitrarily, receive more attention than others. This is a function of both my individual interests and ability to (hopefully) bring something new to the table that can advance the literature. That said, I will do my best not to jump around too much, and address certain familiar disputes that most would consider to be of major interest and importance in this field. Let the games begin!
CHAPTER 2

Developing a Framework for Introspection

1.1: The details of introspection problem and moving forward

In this chapter, we will address the details of introspection problem. Since the ‘details of introspection problem’ is the problem of developing a plausible ‘minimal notion’ of introspection that leaves open a good deal of room as to the form of a correct answer to the problem of phenomenal consciousness, it makes sense to tackle this first. With this minimal notion of introspection on the table, we will then discuss the relationship between phenomenal consciousness and introspective access.

1.2: First-person warrant and the content of introspection.

‘Introspection’, to a first approximation, refers to the distinctive first-person form of epistemic access that creatures such as ourselves have to some of our own current mental states/events. It is widely held that this form of epistemic access to some of our own current mental states/events is epistemically superior to third-person or ‘publicly accessible’ forms of epistemic access to the relevant subject matter.

Indeed, one common debate in this area concerns whether or not the form of epistemic access we have here is certain or infallible. For now, we need not even assume that introspective access is privileged epistemic access. We will assume, however, that introspective access to some of our own current mental states/events involves a distinctive sort of epistemic warrant—a warrant that only, as a matter of fact at least, supports certain propositions about the current mental state/event one is in/having. Let’s refer to this as ‘first-person warrant’.
I am inclined to think that first-person warrant is a distinctive sort of epistemic support for propositions such as ‘mental event \( m \) is occurring’. Even though the content here is not subject-involving, the relevant sort of epistemic support for it is such that it must involve the subject of the mental event that figures in the content, and not involve any other subject of experience. This constraint is designed to capture the notion that the distinctive sort of epistemic warrant here is first-person.

So, for example, if the relevant form of epistemic support here is just subject \( S \)’s having mental event \( m/\)mental event \( m \)’s occurring in \( S \), then it is obvious how this account of first-person warrant meets the above ‘first-person’ constraint. To take a more interesting example, suppose that first-person warrant is some form of awareness of a mental event. In order for this account to meet the first-person constraint, the form of awareness in question must be such that, at least as a matter of fact, only the subject of the mental event can stand in this ‘awareness’ relation it.

It may also be held that there is a distinctively first-person epistemic warrant for subject-involving contents such as ‘this mental event is occurring in me’. Moreover, one might think that this content is ‘strongly first-person’ in that the relevant epistemic warrant (in and of itself) only favors the truth of such de se contents. On a natural development of this view, the relevant first-person warrant here is ‘mode of presentation sensitive’, so even though I am Christian, this first-person warrant, in and of itself, never supports a proposition such as, ‘Christian is currently thinking about his chess match’.

Christopher Peacocke (1999) argues that while the content of introspection is not subject-involving, a rational and conceptually competent subject with first-person warrant for a relevant subject-free psychological proposition is thereby also
in a position to know the relevant *de se* content.\textsuperscript{13} Since this is a fully distinctive way of knowing a proposition such as ‘*this* pain is occurring in *me* now’—one that can only be had by being the subject of the token pain, we will likely want to distinguish here between something like ‘primary’ and ‘secondary’ introspective access, or ‘original’ and ‘derived’ introspective access.

Indeed, so long as it is held that we have a distinctive first-personal sort of epistemic support for the relevant subject-involving or, *a fortiori*, *de se* propositions, and that first-person-only epistemic support also exists for the relevant subject-free propositions, then the distinction between primary and secondary introspective access will plausibly be needed.

Some may deny such a distinction on the grounds that there is only introspective access to subject-involving propositions. This view, however, seems hard to motivate in the context of phenomenally conscious *mental events*. Granting that there is a kind of first-person-only epistemic support for the relevant subject-involving propositions, it is hard to deny in this context that there is also a kind of first-person-only epistemic support for their subject-free counterparts: where there is first-person warrant for the proposition ‘*this* pain is occurring in *me* now’, there is also plausibly first-person warrant for the proposition ‘*this* pain is occurring now’. Moreover, on the surface anyway, it seems wrongheaded to suppose that the latter is somehow *epistemically parasitical* on the former—if anything, it seems to be the other way around. In any case, we needn’t settle the issue here.

Finally, it is often claimed that first-person-only epistemic support for the relevant *de se* propositions is such that it exhibits a certain kind of *partial* infallibility/incorrigibility, namely, it is *immune to error through misidentification*

\textsuperscript{13} Peacocke, 1999, chapter 6.
relative to the first-person. This controversial but widely accepted idea traces back to certain remarks by Wittgenstein (1958), and to Sydney Shoemaker’s (1968) celebrated paper, ‘Self-reference and Self-awareness’. I will also set this controversial issue aside in what follows as it lies beyond the scope of our discussion, and is peripheral to our focus in this thesis on phenomenal consciousness and (certain) phenomenal properties.

1.3: An epistemological framework for introspection.

We need to distinguish between first-person warrant, having first-person warrant, introspective access, and introspectively justified beliefs. ‘First-person warrant’, as I have already claimed, is a kind of epistemic support for certain propositions. ‘Having first-person warrant’, to a first approximation, is some sort of ‘awareness of’ or ‘psychological responsiveness’ to the relevant epistemic support for a relevant proposition. ‘Introspective access’, roughly speaking, is being in a position to form a belief/judgement with the relevant content and with first-person warrant.

There is no obvious entailment from first-person warrant to having first-person warrant, but this may depend on what one ultimately takes first-person warrant to be. ‘Having/possessing first-person warrant’ and ‘introspective access’ are similar notions, but, on the face of it, it seems coherent that one may have awareness of first-person warrant while lacking the ability to have thoughts with the relevant content. This may depend, however, on how first-person warrant and having first-person warrant are ultimately developed. If, for example, first-person warrant is the fact that one is in the relevant mental state, and having first-person warrant is a kind of propositional awareness of this fact, then this entails that one is capable of
having thoughts with the relevant content, and the gap between having first-person warrant and introspective access would seem to be closed.

Finally, ‘introspectively justified beliefs’ are just those beliefs/j judgements with the relevant content that are had with first-person warrant. For a relevant belief/judgement to be had with first-person warrant, that belief must be based on the subject/believer’s first-person warrant. What this amounts to exactly a vexed issue, but it is also a general problem we needn’t resolve here. There may also be a special answer to the ‘basing problem’ in the case of introspection. In what follows we will be more interested in introspective access than introspectively justified beliefs.

1.4: Epistemic support and propositional versus doxastic justification.

Now, first-person warrant is understood here as providing at least pro tanto support for a relevant proposition. In other words, at the very least, first-person warrant favors the truth of a relevant ‘introspective’ proposition. And it is plausible that ‘favoring the truth’ cannot simply be analyzed in terms of ‘increasing the likelihood of a proposition’s being true’ or some such. Evidence, a paradigmatic, if not only, source of epistemic support for propositions, can favor the truth of a proposition $p$, but it is implausible to suppose that this simply amounts to its increasing the likelihood that $p$ is true—though presumably it does do something along such lines. In terms of a core notion of ‘epistemic justification’, I think we should say that taking a (certain) doxastic attitude to a proposition is ultima facie permissible if and only if one has all-things-considered epistemic support for the relevant proposition. This is sometimes called ‘propositional justification’ as opposed to ‘doxastic justification’.

In cases where we take there to be a kind of epistemic support for propositions of a certain kind, our core notion of ‘epistemic justification’ is
concerned with the epistemic status of taking or forming certain doxastic attitudes, not with the epistemic status of those doxastic attitudes themselves. The epistemic status of doxastic attitudes in this sort of case is parasitic upon the epistemic status of taking or forming such doxastic attitudes. In other words, ‘propositional justification’ here, is conceptually prior to ‘doxastic justification’—which is not to say that ‘doxastic justification’ doesn’t present challenges of its own, or that we can rule out cases of doxastic justification without propositional justification (if there are such cases, then the notion of ‘doxastic justification’ that applies to them will likely differ from the notion we are concerned with here and be a non-normative one).

There is a rather unfortunate tendency in the literature to conflate all-things-considered epistemic support for a proposition, with being epistemically justified in taking a (certain) doxastic attitude to a proposition. One cannot be epistemically justified in taking a doxastic attitude to a proposition unless one has epistemic support for that proposition. In other words, the mere existence of all-things-considered epistemic support for a proposition \( p \) doesn’t make it rationally permissible for a subject \( S \) to take or form a certain doxastic attitude towards \( p \); \( S \) may be completely unaware of the relevant epistemic support for \( p \). With these points in place, let us now consider certain objections/challenges to first-person warrant.

1.5: Challenges to first-person warrant.

Those who favor ‘process reliabilism’ or some other purely truth-conduciveness account of at least ‘pro tanto doxastic justification’ when it comes to introspection, will no doubt reject first-person warrant. Others who think that knowledge does not

---

14 See, e.g., the papers in Conee and Feldman (2004).
15 See, e.g., Goldman (1986).
entail justification may also reject first-person warrant on the grounds that ‘introspective access’ is best understood in terms of introspective knowledge, or ‘being in a position to (introspectively) know’. In response, all I can really say is that it seems hard to deny that there is at least a sense of ‘introspective epistemic access’ that involves first-person warrant. Moreover, in the case of the latter view, even if we agree that knowledge is best understood as, say, the most general factive attitude, an understanding of ‘introspective epistemic access’ in these terms needn’t be the only one available in this context.

Ultimately, it just seems highly counterintuitive to deny the claim that there is a sense of ‘introspective epistemic access’ according to which subjects such as ourselves often have distinctively first-person grounds for judging, e.g. ‘I am in pain!’. The best way to argue against this claim, I think, would be to attempt to show that the different ways of developing first-person warrant all face grave difficulties. This strategy, however, does not dispute the prima facie plausibility of first-person warrant, and requires an examination of the different approaches that might be taken here. We will come to this issue later on, and it should be clear that there are at least some ways of developing first-person warrant that do not seem to face any crushing objections. In any case, we will now return to a foundational issue in the vicinity that we glossed over in 1.4 of this chapter.

1.6: Having epistemic justification as having permission to believe.

Understanding ‘epistemic justification’ in terms of ‘doxastic permissibility’ is fairly common nowadays,\(^{16}\) but some continue to think of ‘epistemic justification’ in terms of right belief/what one would be right to believe. In my view, there is no such thing as right belief from a purely epistemic perspective, as there is no such

\(^{16}\) See, e.g., Littlejohn (2012).
thing as being *wrong not to believe* something from a purely epistemic perspective. Sherry has abundant evidence that Terry strongly dislikes her, and she has no defeaters for this, but since Sherry doesn’t care in the least how her nodding acquaintance Terry might feel about her, she doesn’t have any relevant doxastic attitude here. If we are to convict Sherry of failing to do something that she rationally ought to do, then we are all *frequently and unavoidably* guilty of doing the same. If ought implies can, then we get the false result that we can, in fact, do better than Sherry when it comes to such cases. Perhaps an ideal epistemic agent could do better, but this is not an adequate reply here since we are concerned with what non-ideal epistemic agents such as ourselves are rationally required to do. And assuming ought implies can, this can’t be something that *only* an ideal epistemic agent can do.

In general, our evaluations of what an agent *should or ought* to believe seem keyed to practical and moral relevance. For example, suppose Sheila has abundant evidence that her cat is very ill and in need of veterinary care, and she has no defeaters, then (setting aside forgotten evidence) it is quite plausible that Sheila really *should or ought* to believe this; she should take enough of an interest (assuming here that she is so able) in her pet’s physical well-being to have arrived at this belief given her evidence concerning her cat’s bodily integrity.

The point here though is just that from a purely epistemic perspective, there is nothing obviously *wrong* with Sheila *not* forming the relevant doxastic attitude. If this is correct, then we can credit non-humans and small children with *having* first-person warrant, but lacking the ability to think about their own mental states. In other words, we can credit them with having first-person warrant in the absence of introspective access. Let me explain.
2.1: Phenomenal consciousness without introspective access.

The guiding idea behind the conditional introspective access characterization of ‘phenomenal consciousness’ is that certain non-humans and humans may lack the psychological enabling condition(s) needed for introspective access to their own current phenomenally conscious mental states/events. Plausibly, certain non-humans and (human) children share the epistemic component of introspective access with us, but they lack the ability to think about their own mental states/events or, perhaps, mental states/events in general.

I think that the main source of resistance to the idea that non-humans and small children can have first-person warrant without introspective access derives from the concern that if these individuals have first-person warrant, then, other things being equal, they should or ought to believe a relevant mental proposition, and since ought implies can, it therefore (incorrectly) follows that they can, indeed, form beliefs/judgements about the relevant propositions.

I have already argued, however, that ‘epistemic justification’ is best understood in terms of what one is allowed or permitted to believe/what beliefs would be permissibly held. If that is correct, then there is no problem here since the fact that non-human animals and small children can’t form the relevant beliefs is quite compatible with their being permitted to do so: just because one is allowed to do x does not entail that one can, in fact, x.

Alternatively, if first-person warrant does not entail having first-person warrant, and having first-person warrant is understood as a psychological enabling condition of introspective access, then we can arguably bypass the issue of how best to understand ‘epistemic justification’ by holding that certain non-human animals and small children can instantiate first-person warrant without also having it (in the
relevant sense of ‘having’, of course—I will always italicize this use of ‘having’ to make this clear). Depending on one’s account of first-person warrant, it may be more or less plausible to suppose that there is an entailment from first-person warrant to having it, and it may be more or less plausible to suppose that first-person warrant is shared between us and certain non-humans (and humans).

2.2: A minimal account of introspection.

Though our ‘minimal account’ of introspection arguably isn’t so minimal after all, it seems independently well-motivated. It wasn’t built around trying to accommodate the no-currently-obvious-reductive-analysis constraint, and it doesn’t even presuppose that introspective access is privileged access or the like. Our understanding of the sort of ‘epistemic justification’ that is built-into our notion of ‘conditional introspective access’, however, is clearly prima facie normative, and so it respects the no-currently-obvious-reductive analysis constraint.

Moreover, even if there is no entailment from first-person warrant to having first-person warrant, it is plausible that first-person warrant itself, as a kind of epistemic support, meets the no-currently-obvious-reductive-analysis constraint as well. Favoring the truth of a proposition/providing epistemic support for a proposition is a relation that is at least prima facie resistant to a reductive analysis—especially a reductive analysis in purely structural/functional terms. In any case, I will now show that this framework, in conjunction with our understanding of ‘phenomenal consciousness’ as having its reference semantically determined by ‘conditional introspective access’, yields some further constraints on the correct form of an account of introspection as well as the relationship between phenomenal consciousness and introspection.
3.1: The conditional introspective access characterization, introspection, and phenomenal consciousness.

A phenomenally conscious mental event is one that is conditionally introspectively accessible. Thus, in a subject $S$ who has the relevant psychological enabling condition(s) for introspective access, phenomenally conscious mental events are introspectively accessible. This means that $S$ is in a position to have an introspectively justified belief about a relevant psychological proposition, simply by virtue of having a phenomenally conscious mental event, and having the relevant psychological enabling condition(s). In our framework, this means that there is first-person warrant for a relevant psychological proposition, and $S$ has this warrant. So, just by virtue of having a phenomenally conscious mental event, and having the relevant psychological enabling condition(s) for introspective access, $S$ has first-person warrant for a relevant psychological proposition—this is what $S$’s having introspective access to the relevant proposition presupposes in our framework.

Moreover, first-person warrant here is either just having a phenomenally conscious mental state/event, or is entailed by having a phenomenally conscious mental state/event. That’s the only way for it to be the case that $S$ has introspective access to a relevant psychological proposition simply by virtue of having a phenomenally conscious mental event, and having the relevant psychological enabling condition(s) for introspective access. It should be noted that unless it is assumed that having first-person warrant is a psychological enabling condition for introspective access, we get an entailment from first-person warrant to having first-person warrant. If having first-person warrant is not a psychological enabling condition for introspective access, then the fact that a subject $S$ who has a phenomenally conscious mental event and the relevant psychological enabling
condition has introspective access to that mental event (strictly, to the relevant proposition) means that first-person warrant entails having first-person warrant.

It is clear, then, that once it is granted that subjects with the relevant psychological enabling condition(s) are such that they are in a position to have first-personally supported beliefs about the presence of their phenomenally conscious mental states/events, certain additional constraints on the shape of a correct account of introspection emerge. One might worry that it was hasty to understand ‘phenomenal consciousness’ as having its reference (semantically) determined by ‘conditional introspective access’, and wonder whether some looser semantic connection to introspection might be more appropriate. I think it is independently plausible, however, that phenomenal consciousness is related to conditional introspective access in the manner entailed by the conditional introspective access characterization of ‘phenomenal consciousness’ and our epistemological framework for introspection. Consideration of Timothy Williamson’s (2000) well-known ‘anti-luminosity’ argument, in my view, provides indirect support for this claim. Let me explain.

Much of the interest behind Williamson’s (2000) well-known ‘anti-luminosity’ argument is that some claim in the vicinity of the ‘luminosity’ of phenomenally conscious mental states/events is fairly plausible. The claim that phenomenally conscious mental events are ‘luminous’ is the claim that when a subject $S$ has a phenomenally conscious mental event $e$, $S$ is in a position to know that they have $e$. I don’t think that Williamson’s (2000) phenomenal sorites argument against luminosity succeeds, but I also think that luminosity, as it stands, is clearly inadequate.
My main issue with the claim that phenomenally conscious mental events are luminous is that infants and (many) non-human animals have phenomenally conscious mental events, but they are also clearly not in a position to know that they are having such events; they lack the ability to have the relevant thoughts. Moreover, even in those who possess the relevant psychological enabling condition, the claim that such individuals are in a position to know that they are having the relevant mental event is still too strong for our purposes, unless, perhaps, knowledge is understood as a primitive factive attitude. In any case, our understanding of introspective access incorporates epistemic justification, namely, having first-person warrant, and so our counterpart to a (phenomenal) luminosity thesis is weaker in so far as it only involves being in a position to have an epistemically justified (first-personally warranted) belief concerning a relevant content, not being in a position to know a relevant content.

Our view of the relationship between phenomenal consciousness and introspection, then, capture something of the intuitively plausible spirit of luminosity, without its major flaw viz. infants and non-humans having phenomenally conscious mental events without being in a position to know them, and without the unnecessarily strong commitment to knowledge when it comes to being in a certain epistemic position by virtue of having a phenomenally conscious mental event. Thus, the details of the relationship between phenomenal consciousness and introspection that are entailed by the conditional introspective access characterization of ‘phenomenal consciousness’ seem independently well-motivated. With these important points in place, we will now turn to explore some of the ways in which the relationship between phenomenal consciousness and introspection may be developed within this framework.
3.2: Major theoretical options for the relationship between phenomenal consciousness and introspective access

In the previous section, it was shown that our theoretical framework for introspection, in conjunction with the conditional introspective access characterization of ‘phenomenal consciousness’, leads to the conclusion that having a phenomenally conscious mental event either is first-person warrant, or entails a psychological phenomenon that is/constitutes first-person warrant. If we take the first option, then we must consider the relationship between first-person warrant and having first-person warrant. If having a phenomenally conscious mental event entails being acquainted with that mental event, then it is highly plausible that first-person warrant entails having first-person warrant. The question, then, is whether or not having a phenomenally conscious mental event entails some sort of awareness/acquaintance with that mental event.

In my view, it is plausible that we can attend to our phenomenally conscious mental events, and attending to a phenomenally conscious mental event is a psychological enabling condition for introspective access. So, if having a phenomenally conscious mental event constitutes first-person warrant, then having first-person warrant is plausibly just attending to that mental event. Thus, having first-person warrant turns out to be a psychological enabling condition for introspective access, and first-person warrant does not entail having first-person warrant. As we saw in chapter 1, the notion that phenomenally conscious mental events are mental events that we are, in some sense, ‘aware of’, may plausibly be cashed out in terms of conditional introspective accessibility.

Setting this sense of ‘aware of’ aside here, it is far from clear whether there is some further sense in which we are always ‘aware of’ our own phenomenally
conscious mental events; I find this suggestion to be rather dubious. The only further sense in which we are ‘aware of’ our own phenomenally conscious mental events is an *attentional* sense, and this ‘awareness of’ our own phenomenally conscious mental events is not one that always or perhaps even typically accompanies such events. It is not strenuous, of course, to attend to our own phenomenally conscious mental events, but this does not mean that we usually bother attending to them. Having said that, there may be a focus of attention that typically zeroes in on at least some of our phenomenally conscious mental states/events from one moment to the next, but this is *not* to say that all our currently occurring phenomenally conscious events are always attended to—that seems quite implausible.

It may be wondered whether the above account of first-person warrant and *having* first-person warrant can apply to ‘higher-order awareness’ views of phenomenal consciousness itself. Suppose that phenomenal consciousness constitutes first-person warrant, and that a phenomenally conscious mental event, roughly speaking, is a mental event *e* that has the (relational) property of being the target of a certain sort of awareness by the subject of *e*. Is it still plausible here to suggest that *having* first-person warrant is attending to the relevant mental event? I think so. At the very least, I can discern no obvious reason for rejecting the view that on the relevant ‘higher-order’ metaphysics of phenomenal consciousness, we can still attend to a relevant mental event, and this form of awareness of a mental event substantially differs from that which constitutes a mental event’s being phenomenally conscious. In any case, I will set this issue aside.

Let us now turn to the second option: having a phenomenally conscious mental event entails some psychological phenomenon that constitutes first-person warrant. It is difficult to envisage any development of such a view except along the
lines of having a phenomenally conscious mental event entailing acquaintance with that mental event, and acquaintance, or whatever form of ‘higher-order’ awareness one prefers, being/constituting first-person warrant. An immediate challenge to this sort of view is that it appears to lack any clear motivation: if one thinks that we are acquainted with our phenomenally conscious mental events, then why not just hold that having a phenomenally conscious mental event constitutes first-person warrant, and acquaintance is having first-person warrant?

Well, one might have concerns about the notion that having a phenomenally conscious mental event is first-person warrant for a relevant proposition about that mental event: how can something provide evidence or epistemic support for itself? This concern does not arise on a higher-order metaphysics of phenomenal consciousness, but one might think that such a metaphysics of phenomenal consciousness should be rejected. If so, then it does seem fairly plausible that we are left with the rather puzzling question of how the fact that one has a phenomenally conscious mental event provides non-trivial epistemic support for its own occurrence. So, there does seem to be a broad strategy available for meeting the challenge from motivation.

Another challenge for the view, however, derives from the need to address the issue of having first-person warrant. If we suppose for the sake of argument that having first-person warrant is something along the lines of attending to first-person warrant here, then the idea is that we can attend to our acquaintance with phenomenally conscious mental events. But this seems, on the surface anyway, to be dubious. We (many of us) may be able to attend to our phenomenally conscious mental events, but we hardly seem able to attend to our acquaintance with our phenomenally conscious mental events.
This seems to exhaust the major theoretical options as regards the relationship between phenomenal consciousness and introspective access given our background assumptions and arguments. I will not attempt to settle the question of which approach is to be preferred over others, but I do think that, other things being equal, we should take phenomenal consciousness itself to constitute first-person warrant, not some property supposedly entailed by its instantiation such as ‘acquaintance’.

If this is right, then we still have to decide between views according to which phenomenal consciousness is something like higher-order awareness, and those according to which phenomenal consciousness is a ‘first-order’ phenomenon. I will not attempt to settle the matter here as I think that a good answer to this question can only emerge from an extended discussion of certain metaphysical pros and cons associated with the relevant views which is beyond the scope of this thesis. That said, by the end of this thesis, we will be in a better positioned to fruitfully address the matter in future work. In what follows, we will briefly address the related question in this area of whether or not phenomenal consciousness itself figures in the content of introspection.

3.3: Phenomenal consciousness and the content of introspection

Though we have already noted certain important issues concerning the content of introspection, we have yet to address the fundamental issue of whether or not the content of introspection includes phenomenal consciousness. The view that phenomenal consciousness does not figure in the content of introspection is more in line with our understanding of ‘phenomenal consciousness’ in terms of conditional introspective accessibility. If phenomenal consciousness itself figures in the content
of introspection, then one would expect that we would have a grip on phenomenal consciousness independently of its relationship to introspection.

We needn’t deny, however, that the correct theory of phenomenal consciousness and/or introspection may have such a consequence. So long as it is not readily apparent that we do have such a grip on phenomenal consciousness, however, the conditional introspective access characterization earns its keep. That said, it seems plausible that, at best, phenomenal consciousness qua property that grounds conditional introspective accessibility is indirectly or secondarily introspectively accessible—and even this much is far from clear. So, I’m ultimately inclined to think that the content of introspection just involves phenomenal properties, that is, properties had by phenomenally conscious mental events.

4.1: A partial summary and some general conclusions.

In this chapter, we addressed the details of introspection problem: the problem of developing a minimal notion of introspection that can be put to use in the context of theorizing about phenomenal consciousness. I claimed that the epistemic core of introspection is ‘first-person warrant’ a distinctive first-personal form of epistemic support that is often present for certain present-tense psychological propositions. Though it is difficult to directly argue for this claim, it clearly has prima facie support, and we did not encounter any compelling reasons for abandoning this intuitively plausible starting point.

Having said that, there are a number of important alternative understandings of introspection which I have not discussed, such as so-called ‘constitutivist’ views according to which there is some sort of constitutive connection between phenomenally conscious mental states/events and beliefs about such states/events.\footnote{See, e.g., Shoemaker, 1994.}
I’m inclined to think, however, that constitutivist approaches to introspection are, at best, only a serious contender when it comes to a subset of phenomenally conscious mental states/events. In other words, I don’t think they hold much promise as a general approach to introspection.

Moreover, all I am really committed to here is that first-person warrant is a legitimate epistemic phenomenon that is closely connected to all phenomenally conscious mental state/events. If certain phenomenally conscious mental states/events or even all such states/events have some further epistemic feature or features, or are closely related to such a feature(s), then as long as they are compatible with first-person warrant, there is no problem for my account. In other words, my account seems compatible with epistemic pluralism about introspection, though this would presumably need to be assessed on a case-by-case basis.

Some general epistemological background claims were also made in this chapter, many of which are familiar, but, as one might expect, few are entirely uncontroversial. I offered a brief defense of some of these claims, but others were more or less assumed for the sake of theory-construction. Having said that, those general epistemological claims that were largely assumed in this context also have a certain amount of prima facie plausibility, so I don’t feel too guilty in advancing them without much discussion.

Upon arriving at a general framework for understanding introspection, we then proceeded to examine some of the ways in which introspection, so understood, might be related to phenomenal consciousness. It quickly became obvious, however, that without making certain further assumptions about phenomenal consciousness itself, it is difficult to make significant headway when it comes to the precise relationship between the two. This, however, is what one would expect from an
appropriately minimal account of introspection: it shouldn’t place such constraints on theorizing about phenomenal consciousness so as to seriously prejudge the issue of its metaphysical nature or its exact relationship to introspection. With these epistemological foundations in place, the next chapter addresses ‘phenomenal character’ and, in particular, some ‘phenomenal properties’ that are of major interest to theorists, such as colors and pain.
CHAPTER 3

Phenomenal Properties and their Problems

1.1: The problem of the phenomenal character of phenomenally conscious mentalia

Having clarified and developed our framework for thinking about introspection in the previous chapter, we are now in a position to put it to use in the context of the problem of ‘phenomenal character’, or the problem that certain ‘phenomenal properties’ pose for theorizing in this area. In this chapter, we will examine certain phenomenal properties in more depth than others, but a number of the points developed here plausibly have a broad application and/or significance for this area.

So, while a good deal of the focus of this chapter is on color as we understand/know it through introspection, much of what is said here applies to ‘secondary qualities’ in general, or is at least relevant to such properties. The other phenomenal property we will concentrate on in this chapter is pain, and much of what is said here about pain will be relevant to theorizing about bodily sensations in general. In any case, let us begin with some important preliminaries and general issues before proceeding on to more specific topic areas here.

Phenomenally conscious mental states/events have a phenomenal character. This much, at least, is typically agreed by experts in the field. There is considerable controversy surrounding how exactly to understand the notion of ‘phenomenal character’. Let us start, as is customary, with the suggestion that the ‘phenomenal character’ of (a token) phenomenally conscious mental event, refers to the way that
(token) mental event seems or is like, to/as regards, its subject. It is rather dubious, perhaps, to casually suggest that there is just one/the way that a token phenomenally conscious mental event will seem or be, to/as regards, its subject: presumably there are many relevant ways that a token mental event will seem or be, to/as regards, its subject.

This raises the controversial issue of whether or not there is some total or ‘unified’ way that any token phenomenally conscious mental event at a time $t$, will seem or be, to/as regards its subject at $t$. This does seem fairly plausible to me, but I will not pursue the matter here. My goal here is just, as advertised, to arrive at a solid working understanding of ‘phenomenal character’ for the purposes of our discussion of the matter/related matters, in this chapter. So, let’s grant the point that there are multiple relevant ways that token phenomenally conscious mental events can seem or be, to/as regards, its subject. Now, we can remain neutral on the further issue of whether or not there is, on top of this, some further total or complete way that a relevant token mental event will seem or be, to/as regards, its subject.

Trouble immediately arises for this suggestion, however, in terms of how to make either ‘seems to the subject’ or ‘as regards its subject’, more concrete without begging any major questions along the way. For example, as we noted in chapter 1, the ‘seems’ characterization is at least ambiguous or polysemous between an introspective/conditional introspective interpretation, and what we might call a neither doxastic nor conditional doxastic interpretation. On the former interpretation, ‘phenomenal character’ just amounts to phenomena of phenomenally conscious mental events that are conditionally introspectively accessible.

Plausibly, there are a good many properties of phenomenally conscious events that are conditionally introspectively accessible (beyond just a token pain
event being a pain, *mutatis mutandis*, for other phenomenally conscious events). Such properties will be counted among the various *ways* that such (phenomenally conscious) mental events seem to their subjects in either sense of ‘seeming’, though on some views there may be certain ways these mental events non-doxastically seem to their subjects that aren’t ways that are also conditionally introspectively accessible: what Derek Shiller (2017) has suggestively called ‘hidden qualia’. In effect, I argued against this sort of view in 7.5 & 7.6 of chapter 1. Thus, I will be assuming the ‘conditional introspective’ interpretation of ‘*way* a mental event *e* seems to its subject’ in what follows.

1.2: Phenomenal properties and being phenomenally proprietary

An important issue in the vicinity is whether these various properties of phenomenally conscious mental events can be instantiated by phenomenally unconscious mental events. In other words, are these various ways that phenomenally conscious mental events seem to their subjects the preserve of phenomenally conscious *mentalia*? The answer to this question might initially seem to be an obvious ‘yes’: since only phenomenally conscious mental events are such that they *seem* some way or other to their subjects in the first place, these various *seemings* will, of course, only be the preserve of the phenomenally conscious.

This answer assumes, however, that these ways that phenomenally conscious mental events seem to their subjects are *ways of seeming* as opposed to *ways things seem*. And, thus far, we have only committed to the latter. With this distinction in mind, there is no obvious answer to the question of whether or not these ways that phenomenally conscious mental events seem to their subjects are the preserve of phenomenally conscious *mentalia*. For example, suppose that a phenomenally conscious mental event *e* seems unpleasant to its subject. Being unpleasant to its
subject might be a way that only a phenomenally conscious event can be, but then again, perhaps not.

The issue can be brought more sharply into focus if we stick to the conditional introspective access understanding of ‘seeming to its subject’. The question then becomes: are all conditionally introspectively accessible properties of phenomenally conscious mental events the preserve of such-wise mental events (setting aside the property ‘phenomenal consciousness’ itself)? The ‘egalitarian’ view holds that for all properties that are conditionally introspectively accessible/figure in introspective contents, these are such that they can be instantiated by phenomenally unconscious mental events.

As I pointed out in chapter 1, certain families of mental event-types such as bodily sensations seem to be such that they are the preserve of phenomenally conscious mental events. But ‘cognitive’ mental events, such as phenomenally conscious judgements or affirmations, seem to be clear cases of introspectively accessible properties that can be instantiated by phenomenally unconscious mental events as well. We will consider the matter later on (chapter 4).

There is little doubt that many of the properties of phenomenally conscious mental events that are conditionally introspectively accessible raise major puzzles and problems for theorizing in this area—irrespective of whether or not such properties are the preserve of the phenomenally conscious. There is also little doubt, however, that if there are such properties (aside from phenomenal consciousness itself), then this makes the task of explaining these properties, and their phenomenally unconscious ‘counterparts’, so to speak, all the more difficult.

The egalitarian view would definitely make theorists’ lives easier in this regard—especially those with reductive aspirations. For the moment, however, we
can remain neutral on the issue by simply concentrating on those properties of phenomenally conscious mental events that are conditionally introspectively accessible. This is difficult enough, as there is considerable controversy surrounding which properties of phenomenally conscious mental events are conditionally introspectively accessible at all, and how exactly those that are conditionally introspectively accessible should be characterized. Phenomenally conscious perceptual events are an obvious problem case in this respect. I will discuss a number of important aspects of this ‘problem case’ in what follows, before turning to examine other phenomenal properties such as pain.

2.1: Phenomenally conscious perceptual events and the contents of introspection.

One of the initial challenges for theorizing about phenomenally conscious perceptual events is that of sorting out which properties are properties of these events themselves, and which are properties of the mind-independent world, or are represented as such. For example, the phenomenal character of a phenomenally conscious visual event may be partially constituted by color properties such as ‘bluish’, but it is notoriously contentious to hold that the relevant phenomenal events instantiate such a property.

Having said that, it does seem plausible that color properties are such that we can have introspective access to them. On a simple picture of how this might work, when a subject like myself has a phenomenally conscious visual event that involves some sort of ‘visual awareness’ of a blue object, they have introspective access to their standing in the relevant awareness relation to that blue object and (often) its blueness. Thus, even though blueness is, let’s suppose for now, a property of
externalia, I can have introspective access to it because I can have introspective access to my standing in the relevant awareness relation to it.

In my view, a crucial question here is whether or not we can attend to perceptual awareness itself, or only to one of its relata, namely, its ‘object’ or content, depending on one’s views concerning the nature of phenomenally conscious perception. It seems reasonably plausible, as I argued in chapter 2 concerning introspective access to phenomenally conscious mental events, that inner attention is a psychological enabling condition of introspective access, since ‘inner attention’ is the right way of cashing out having first-person warrant, and having first-person warrant is required for a subject to have introspective access to one of their own phenomenally conscious mental events.

In this framework, it is also natural to suggest that attention to a property of a phenomenally conscious event is also required in order for one to have introspective access to that property. In other words, for one to be in a position to form a first-personally warranted belief/judgement that ascribes a certain property of a phenomenally conscious mental event to that event, one must first have registered the presence of that property in some way—according to the view suggested here, by way of an act of ‘internal’ attention. Now, the trouble is that it is quite difficult to say whether or not perceptual awareness, that is, the relevant relational property of phenomenally conscious perceptual events, is such that we can attend to it.

Some, such as Benj Hellie (2013), have suggested that we can’t attend to perceptual awareness, whereas others, such as Chalmers (2013), have suggested that we can do so. If we can’t attend to it, then we require a different sort of account as to how our introspective access to such properties, broadly speaking, works. Such an account needn’t be incompatible with the account of introspection offered in chapter
as our account of introspection in the previous chapter was mainly concerned with introspective access to (token) phenomenally conscious mental events, and not properties of such events beyond subject-involvement, and what we might call ‘ontic category’, such as our introspective access to a token pain event being a pain.

Both subject-involvement and ontic category are plausibly epistemically parasitic upon our ‘core’ or ‘primary’ introspective access to the relevant token events themselves: ‘this event is occurring’. It may be argued that our introspective access to perceptual awareness itself is this way as well: it is indirect or secondary, as we only have direct or primary introspective access to the object/content of perceptual awareness. So, when it comes to our introspective access to perceptual awareness, we may be pushed towards a more intricate overall account introspection, though this is arguably not problematic. In any case, let us now turn to the issue of whether or not we can attend to perceptual awareness.

2.2: Can we attend to our own instantiations of perceptual awareness?

In examining whether or not we can attend to our own instantiations of perceptual awareness, we need to first consider certain obvious cases such as blurry visual experiences that may be thought to motivate the view that we can attend to ways of being visually aware of a content or object and (some of) its properties. If my visual awareness is blurry visual awareness, then that is a way my awareness is, as opposed to some property of an external object. Thus, if we can attend to a way of being visually aware of something, then it follows that we are, indeed, able to attend to visual awareness. What to make of this sort of case?

Well, we need to consider the suggestion that in the case of blurry visual awareness, we can only attend to the object of awareness. Take a myopic individual such as myself. With my glasses on, distant objects are in focus, but adjust the
prescription a bit in the wrong direction, and those same distant objects begin to look a bit blurry. Suppose I go to an optometrist for an eye exam. When I am looking through a phoropter at an eye exam screen, lens changes from my current prescription can make things look blurry. But things can also look blurry to me if the screen suddenly glitches and the sharpness of the display is altered. Suppose the screen glitches out in this way at the exact same time as my lens is being changed by the optometrist. Would I be in a position to tell, on the basis of introspection alone, whether the display suddenly became blurry or my visual awareness became blurry? I think that the answer to this question is clearly ‘no’, so it seems questionable that the blurry mode of visual awareness is even introspectively accessible, never mind a way of being perceptually aware of something that we can attend to. Since blurry visual awareness is, at best, just one mode of visual awareness, it certainly doesn’t follow that we can’t attend to perceptual awareness simply from our inability to attend to one particular way of being visually aware of something. So, let’s consider some other putative cases.

Chalmers (2013) briefly describes a case involving a purported shift in attention from the content of visual awareness to visual awareness itself. He writes,

[I]n a normal perceptual mode, look around a still room. Nothing seems to change. Now switch to an introspective mode and look around again. One is now quite clearly aware of a change: for example, when objects disappear from sight. But nothing in the world seems to change. (If one objects that one’s head or eyes seem to move, switch the case to a covert change in attention.) The change one is aware of is a change in one’s experience: a phenomenal change.18

I think that this sort of case meets with a response that is analogous to the one I gave for blurry visual awareness: the sense that ‘nothing in the world seems to change’ when one purportedly switches to an ‘introspective mode’ of attention here is

18 Chalmers, 2013, p. 347.
arguably just a reliable inference. Consider a case in which worldly conditions are abnormal such that as one visually scans a certain strange room, the moment before the subject ceases to be visually aware of an object \( o \) in that room as they move on to attend to another object, \( o \) disappears. Can introspection alone tell us whether we are in the strange room, or are noting shifts in perceptual awareness as we scan a normal room ‘in an introspective mode’? I don’t think so, but the initial example from Chalmers (2013) is challenging enough without throwing ‘strange room’ into the mix. I conclude that Chalmers’ (2013) case of being able to attend to perceptual awareness is, at best, inconclusive.

2.3: Harmanian Transparency

It is fairly plausible that a property like perceptual awareness is conditionally introspectively accessible, but if it is, then we face the issue we’ve just been examining of whether or not we can attend to it. For the record, in “the refutation of idealism”, G.E. Moore (1903) comes quite close to endorsing the view that we cannot attend to perceptual awareness. More recently, Gilbert Harman (1990) provides a well-known ‘transparency’ argument against intrinsic, non-relational, properties of phenomenally conscious perceptual events, but he seems to think that we can attend to perceptual awareness/representing a content. In his paper “the transparency of experience”, Michael Martin (2002) suggests that, when it comes to phenomenally conscious perception, we can only attend to \textit{externalia}, but he ultimately backslides on this point. It seems to me that Chalmers (2013) is correct in his assessment that most advocates of a transparency thesis concerning phenomenally conscious perceptual events only defend ‘Harmanian transparency’.

‘Harmanian transparency’, is simply Harman’s (1990) view that we can’t attend to any intrinsic, non-relational, properties of phenomenally conscious
perceptual events. The only ‘qualities’ we can attend to are those of *externalia/those* represented as being instantiated by some worldly particular. In my view, Harmanian transparency is inconclusive. If Harman (1990) is just arguing that we can’t attend to any intrinsic, non-relational, properties of phenomenal events, then, at best, this is just to say that we don’t have any ‘phenomenological’ or introspective support for a traditional ‘qualia-freak’/‘mentalist’ metaphysics of qualities such as colors. Importantly, *not having* introspective support for such a view *does not entail having* introspective support for a relevant alternative, such as the view that properties such as colors are properties of externalia (in which case either direct realism or representationalism would presumably be introspectively supported).

Thus, if introspection/phenomenology is *epistemically neutral* in this area, that is, if we lack introspective/phenomenological support for any of the relevant alternatives over others, then Harman’s (1990) ‘transparency’ thesis/argument loses its bite: qualia-freak views are no worse off than the others in this important respect.

Now, it’s safe to say that Harman (1990) rejects the relevant neutrality thesis in favor of the view that there is phenomenological/introspective evidence that colors are properties instantiated by *externalia*. Here is the famous passage from Harman (1990), *in context*:

Some sense datum theorists will object that Eloise is indeed aware of the relevant mental paint when she is aware of an arrangement of color, because these sense datum theorists assert that the color she is aware of is inner and mental and not a property of external objects. But, this sense datum claim is counter to ordinary visual experience. When Eloise sees a tree before her, the colors she experiences are all experienced as features of the tree and its surroundings. None of them are experienced as intrinsic features of her experience. Nor does she experience any features of anything as intrinsic features of her experience. And that is true of you too. There is nothing special about Eloise’s visual experience. When you see a tree you do not experience any features as intrinsic features of your experience. Look at a tree and try to turn your attention to intrinsic features of your visual experience. I predict that you will find that the only features there to turn your attention to will be features of the presented tree, including relational features of the tree ‘from here.’ The sense datum theorists’ view about our immediate experience of color is definitely not the naïve view; it does not represent the viewpoint of ordinary perception. The sense datum
theory is not the result of phenomenological study; it is rather the result of an argument, namely, the argument from illusion. But that argument is either invalid or question-begging, as we have seen.\(^\text{19}\)

Harman (1990) may be right that sense-datum and other mind-dependent views of the relevant properties are ‘not the result of phenomenological study’, but the question is whether or not the relevant ‘phenomenology’ suggests a view that is \textit{incompatible} with a ‘sense-datum’ or adverbial view. A number of issues require disentangling here, such as whether or not the phenomenology supports a relational or a representational approach, but let’s set this and other related issues to one side and just try to consider the issue of whether ‘the phenomenology’ of color experience is such that it favors the view that color properties are properties of \textit{externalia}.

The issue of whether we have introspective support for color properties being properties of the object of visual awareness is arguably sensitive to our choice of examples. If we think of a case in which colorful flashes are initiated in close proximity to a conscious viewer in a dark room for brief durations, is it so clear that these colorful flashes are such that the viewer has introspective support for these flashes occurring outside of themselves? Or is it when these flashes are further away from me in the dark room and are more prolonged in duration that it suddenly ‘seems clear’ that these colorful flashes are occurring outside of me?

If we can ultimately make sense of the notion of colors and other ‘secondary qualities’ being instantiated ‘in here’, then it seems that cases like this one, and certain after-image cases in which colorful after-images are only present for a very short duration as though they are ‘popping in and out of existence’, lend at least

some *prima facie* support to the notion that there can be introspective/phenomenological support for colors being instantiated by events *inside the perceiving creature/subject*, as opposed to outside of the perceiver. Naturally, it is far more tempting to suppose that color properties could be instantiated by events occurring in the perceiving subject when we are dealing with examples of events instantiating color properties as opposed to *objects* instantiating color properties.

So, if choice of example really does shift our ‘phenomenological intuitions’ or ideas about introspective support here, then Harman’s (1990) well-known transparency argument is undermined. The matter is far from clear, but I am inclined to think that Harman’s (1990) argument is influenced by his choice of example, where he assumes a relatively stable environment in which color properties persist for some time and the bearers of color properties are *objects*, not *events* occurring in the environment, like colorful explosions.

**2.4: Some venerable hard cases: still no inner attention to perceptual awareness itself?**

Earlier on, Harman (1990) addresses the issue of whether or not perceptual awareness itself is such that we can attend to or have introspective access to it. He writes,

Eloise is aware of the tree as a tree that she is now seeing. So, we can suppose she is aware of some features of her current visual experience. In particular, she is aware that her visual experience has the feature of being an experience of seeing a tree. That is to be aware of an intentional feature of her experience; she is aware that her experience has a certain content.\(^{20}\)

---

Unfortunately, this passage provides no guidance on the issue of whether or not we can attend to being visually aware of some object/content. Most will agree that sensory modes of perceptual awareness/ways of being perceptually aware of something are (conditionally) introspectively accessible, but the issue of attention raises the same sort of issue we encountered with blur (and the other case as well): are we attending to some aspect of the object of awareness/content, or are we attending to the relevant way being perceptually aware? Harman (1990) seems to want to leave it open that we can attend to ways of being perceptually aware of things, but he provides no positive support for such a thesis.

Harman (1990) briefly discusses a few venerable cases that have occasionally been taken as providing support for ways of being perceptually aware of something that we can attend to—cases involving what we might call ‘viewer-centered’ properties of perceptual awareness, such as the way a chair looks to a subject from a certain viewing angle. These properties, such as the ‘elliptical shape’ of a circular coin when viewed from a certain angle, have sometimes been treated as ways of being perceptually aware of the world, as opposed to ways we represent, or are aware of the world, as being. If we reject such an interpretation of this and other related cases, then we must find another way of dealing with the problem that the coin, when viewed from a certain angle, looks both elliptical and circular. In other words, the problem is that it seems we can be visually aware of a coin, c, as being both elliptical and circular (at the same time).

Again, we can avoid this problem of ‘inconsistent looks’ by holding that a coin’s ‘looking elliptical’ in the relevant contexts turns out to be just a natural way for us to express a subtle mode of visual awareness that we can attend to—what we might call the ‘modes of perceptual awareness’ view (MPA). Setting MPA to one
side, one might try to deny that we are, in such cases, visually aware of the coin as being both elliptical and circular, but then it is difficult to explain how we can apparently willfully shift from a phenomenally unconscious perceptual awareness of a relevant object (as being one of the two ways), to a phenomenally conscious one. The phenomenon in question seems better explained as an act of attention to one of the relevant aspects of ‘the object’ of visual awareness: the coin’s being elliptical or its being circular, but we can’t attend to both at the same time. In other words, we cannot attend to the inconsistent nature of the (intentional) object of the relevant visual awareness—instead, we can only ‘flip between’ both aspects when it comes to attention. This is arguably much like our experience in duck-rabbit cases. Let’s call this the ‘gestalt-shift’ view (GS) of ‘elliptical coin’ style cases.

The idea that we can have phenomenally conscious perceptual awareness that attributes inconsistent properties to an object, and be, in a sense, ‘phenomenologically blind’ to this fact, is obviously contentious, but it may also help to explain why we can only attend to one of the relevant aspects of this sort of visual awareness, and not both at the same time. Moreover, it honors the plausible suggestion that both of these aspects are aspects (of the intentional object) of our phenomenally conscious visual awareness at a given time. GS is also committal, however, in that it supposes a representationalist treatment of visual awareness: we cannot attend to the full inconsistent content of the relevant phenomenally conscious mental event, but we can attend to one or the other relevant aspects of its content.

It may be, however, that there are certain (presumably) contingent facts about our attention to our own phenomenally conscious events such that we can’t, at least in elliptical coin style cases, attend to both of the relevant properties of ‘the object’ of perceptual awareness at once, even if these properties aren’t inconsistent. Perhaps
the most promising suggestion in this regard is that visual awareness of the coin’s ‘being elliptical’, is visual awareness of (the instantiation of) a relation between some part of the perceiver/observer and the distal object of awareness, whereas the visual awareness of the coin’s being circular is simply awareness of an intrinsic property of the distal object of awareness. Something like this ‘viewer-dependent’ vs. ‘viewer-independent’ contrast is a good candidate for properties we are visually aware of objects having in this sort of case that can, indeed, be had by a distal object (unlike being both elliptical and circular).

So, if we can only attend to one these properties of the object of visual awareness, to the exclusion of the other, then we have a solution to elliptical coin style cases that doesn’t involve the claim that visual awareness involves representing an inconsistent content. Indeed, this view does not require one to hold that visual awareness involves representing a content at all: a direct realist could maintain that in this sort of case one is aware of the relevant extrinsic property in addition to the relevant intrinsic property of the distal object. Thus, this view definitely has certain dialectical advantages over GS, as it is apparently neutral between representationalist and direct realist views of perceptual awareness (the view that the ‘ellipticality’ of the coin is a mode of visual awareness [way of being visually aware] is also a direct realist friendly option).

This, ‘consistently attributable properties to the object’ view (CAPO), leads to some bold claims concerning the scope of visual awareness, regardless of whether or not one adopts a representationalist or direct realist unpacking of ‘visual awareness’. Admittedly, the extent to which we are visually aware of relationships between ourselves qua objects occupying physical space, and distal objects in our immediate environment is rather unclear. But to handle elliptical coin style cases, we must
recruit a relationship between part(s) of the viewer’s visual system and a distal object; is it plausible that we can be visually aware of visually represent such a relation? I shall not attempt to answer this question until later on, but it should be noted that this theoretical option has been enthusiastically pursued by a number of theorists, especially those endorsing a representationalist view of perceptual awareness. Here is Harman’s (1990) treatment of some cases from the relevant family:

Another point is that Eloise’s visual experience does not just present a tree. It presents a tree as viewed from a certain place. Various features that the tree is presented as having are presented as relations between the viewer and the tree, for example, features the tree has from here. The tree is presented as ‘in front of’ and ‘hiding’ certain other trees. It is presented as fuller on ‘the right’. It is presented as the same size ‘from here’ as a closer smaller tree, which is not to say that it really looks the same in size, only that it is presented as subtending roughly the same angle from here as the smaller tree. To be presented as the same in size from here is not to be presented as the same in size, period. I do not mean to suggest that the way the tree is visually presented as being from here is something is easily expressed in words. In particular, I do not mean to suggest that the tree can be thus presented as subtending a certain visual angle only to someone who understands worlds like ‘subtend’ and ‘angle’... I mean only that this feature of a tree from here is an objective feature of the tree in relation to here, a feature to which perceivers are sensitive and which their visual experience can somehow represent things as having from here.21

In this passage, Harman (1990) gestures toward the view just discussed, namely, that we are visually aware of relations between part(s) of our visual systems, and objects in the surrounding environment. Unfortunately, this view is not well-expressed by the idea that we represent an object as subtending a certain visual angle from here, which is what Harman (1990) occasionally implies in the passage quoted above (note especially the final sentence given the context of Harman’s subtending a certain visual angle analysis). We aren’t visually aware of the relevant relationship between our visual system and the distal object ‘from here’ any more than we are visually aware of the object’s ‘viewpoint-independent’ properties ‘from here’. The

‘size of an object that I am perceptually aware of from here’ is most charitably interpreted as a preliminary way of gesturing towards the proposal that such ‘viewer-centered’ properties of perceptual awareness are to be analyzed in terms of visual awareness of a relationship that obtains between a distal object and a part of the perceiver’s visual system—there is no intrinsic feature of the object that is actually represented in this case, contrary to what ‘viewer-centered size’ may unhelpfully suggest (the expression is still useful enough to earn its keep though, I say).

Thus, according to this view, visual awareness can involve a representation of ‘size’ that is just a proxy for a representation of a relation between part of the observer’s visual system, and the distal object, in addition to our ordinary representations of ‘objective’ size qua intrinsic feature of a distal object (N.B. ‘intrinsic’ properties can also be relational properties of an object, so long as the relation is between different parts of that very same object). Harman (1990), at one point in the passage quoted above, tacitly suggests that the relation between the perceiver’s visual system and the distal object, is one that is not easily expressed in words. That’s a good hedge because, as Christopher Hill (2009) forcefully argues, Harman’s (1990) ‘subtending a certain visual angle’ analysis, “...is at best a gesture in the direction of an answer.”22 If Hill (2009) is correct (as I think he is), then an important problem here is that the more sophisticated the relation(s) needed to handle such cases is, the less plausible it becomes that we may be visually aware of such relations.

To summarize, I think that the major options in response to elliptical coin style cases are MPA, GS, and CAPO. Even if CAPO turns out to be an untenable option, we would still need to decide between GS and MPA. An important problem

---

with adjudicating between these two options, however, is that GS assumes that perceptual awareness involves representing a content. While such a view was the orthodoxy through the 1990’s and early 2000’s, anti-representationalist views, such as direct realism, have since been steadily gaining in popularity. For this reason, among others, we can’t resolve the matter here. But our discussion has (hopefully) helped sharpen the terms of the debate, as GS has been a neglected option in this area. Before closing the book on this case study of difficulties that can arise in identifying and characterizing properties of phenomenally conscious events, we will touch on a further controversial issue that arguably favors a form of MPA over its rivals.

2.5: Phenomenal events and the inner sanctum

Phenomenally conscious events occur in our heads, or, at the very least, inside our skins. This ‘inner’ character of phenomenally conscious events is problematic, as these events also seem to their (respective) subjects to occur in their (respective) heads—though this ‘inner location’ exhibits a good deal of vagueness/indeterminacy when it comes to introspection/phenomenological inspection. When we try to attend to the location of phenomenally conscious mental events, we seem to end up attending to a vague location within our bodies, not some external location. This is supported by examples such as silent speech and thinking, as well as mental imagery and imagination. In the case of phenomenally conscious perceptual episodes, however, it may be argued that there doesn’t seem to be any inner location of such events, and that, if anything, all we seem to be able to turn our attention to in these cases are events in the external world, including events involving our bodies.
At this point, however, it may be countered that perceptual awareness is always perceptual awareness from here, that is, from a phenomenologically accessible ‘inner location’ where I encounter this body, or some part(s) of it, interacting with other externalia, as well as simply those distal externalia themselves. Thus, Harman’s (1990) ‘slip-up’ in suggesting occurrences of visual awareness such as subtending a certain visual angle from here, may be viewed as giving (unwitting) expression to a long-standing ‘Cartesian theatre’ type view/picture. If we are visually aware of relations between (parts of) our bodies and distal objects, then there arguably does, indeed, seem to be a phenomenologically available sense in which a relevant subject is perceptually aware of such relations from here, that is, from a kind of inner sanctum behind the eyes, ears, nose and so on of this body.

Defenders of the view that we cannot attend to perceptual awareness plausibly need to reject the above suggestion concerning the ‘inner location’ from which we are aware of the objects of perceptual awareness, as it strongly suggests that we are, in fact, able to attend to a way in which we are perceptually aware of things. Despite the notorious metaphysical issues associated with a kind of location inside our bodies from which we are perceptually aware of things, such a view does seem to enjoy introspective/phenomenological support.

Indeed, the idea that the functioning brain (or some part(s) of it) is the seat of perceptual awareness, is a ‘Cartesian materialist’ account of this ‘Cartesian theatre’ phenomenon. The ‘theatre’ seems palpable when we consider examples such as silent speech and imagination, but it also seems to recede out of the picture when we try to attend to our phenomenally conscious perceptual events. Nevertheless, as suggested earlier, a kind of ‘inner location’ seems to survive even in these cases, and
it is arguable that we have a faint sense of this ‘inner space’ from which ‘acts’ of perceptual awareness ‘emanate’ or unfold. Here is how Mark Johnston (2010) expresses these ideas:

The modes of presentation of the items in my perceptual field are *perspectival*, that is, they present items to a particular viewing position, or more generally to a particular point from which someone might sense the surrounding environment. The implied position at which those modes of presentation seem to converge is the position of my head and body. To that same implied position, a bodily field, as it were a three-dimensional volume of bodily sensations, also presents. And that implied position is also one from which certain acts, presented as willed, emanate. Furthermore, it is the position where mental acts seem to be available for higher-order awareness. And when mental “images” and sounds are generated by imaginations in a space detached from their respective fields—as when I imagine the Hindenburg bursting into flames or rehearse a tune “in my head”—those imaginings appear at the center, in my mind’s eye or ear as it were.23

This is clearly contentious phenomenological territory, but I must admit that these remarks resonate with my own ‘phenomenological inspection/investigation’ of the relevant cases. I find it difficult to convince myself that such an interpretation of the relevant cases rests on some dubious metaphysical theory like what Benj Hellie (2013) calls the ‘soul-pellet’ view. Instead, ‘soul-pellet’ views seem a natural (though perhaps ‘scientifically discredited’) way of working out the consequences of taking the relevant introspective evidence at face value: introspective support for an inner sanctum *suggests an inner subject* of phenomenal events.

A major dialectical/meta-epistemic issue here is that presumably a fairly large number of competent ‘phenomenologists’ will *not agree* that there is introspective/phenomenological support for an *inner sanctum* or ‘virtual container’ (one of Johnston’s descriptive expressions) within which phenomenally conscious phenomena take place or unfold. But it may be argued that there has been much conflation of whether an *inner location* of phenomenally conscious events is

---
23 Johnston, 2010, p.139.
introspectively supported, with whether an inner subject of phenomenally conscious events, is introspectively supported. Assuming Hume is right that an inner subject of phenomenally conscious perceptual events is not introspectively supported, an inner space within which such a subject might reside or be closely connected, may, nonetheless, enjoy introspective/phenomenological support. Johnston (2010) evocatively refers to this inner space as the arena of presence and action. Others would no doubt derogatorily refer to it as the ‘Cartesian theatre’, and it is somewhat surprising that Johnston (2010) doesn’t address the similarities between his notion and the relevant Dennettian punching-bag notion, if for no other reason than to head-off certain obvious criticisms and concerns.  

In any case, the issue deserves more attention than it has received thus far in the literature, and it would be nice to collect more ‘data’ here in the form of philosophers’ judgements concerning whether there is ‘phenomenological/introspective’ support for ‘innerness’. Even if we are convinced that ‘innerness’ must be ‘illusory’, explaining this illusion of sorts would be an important task for future work. Having flagged certain epistemic/meta-epistemic issues with regard to introspection and phenomenal properties, let us now return to the issue of Harmanian transparency in this context.

3.1: Defeasible introspective support and Harmanian transparency

For all I have claimed thus far, it is quite possible to have introspective support for the existence of certain properties, and for it to be the case that nothing, as a matter of fact, instantiates them. The major alternative is that there is only

---

24 To be fair to Johnston, this may be because, ultimately, he takes the arena of presence and action to be a ‘merely intentional object’ that does not actually exist at the actual world. Needless to say, this proposal raises a number of puzzling questions of its own.
introspective support for properties that *can be instantiated* at the actual world.

There may occasionally be some *room for interpretation* when it comes to a putative property whose instantiation we have introspective support for. We may have introspective support for the existence of a certain property of some phenomenally conscious events, but the nature of this property is very much ‘up for grabs’.

Some might think this suggestion applies to ‘viewpoint-dependent’ properties of perceptual awareness, as well as the ‘innerness’ of phenomenally conscious events discussed in 1.6 of this chapter. Broadly speaking, the more rarefied a purported property of a phenomenally conscious event is, the more difficult it is to suppose that one ‘phenomenological description/interpretation’ of it is to be preferred/privileged over another. In what follows, we will consider the possibility that the relevant Harmanian claim concerning colors, while having defeasible introspective support, is mistaken.

According to Harman (1990) and many others, we have introspective/phenomenological support for colors being instantiated by *externalia*. Harman (1990) himself seems (sensibly enough) to think that the sort of introspective/phenomenological support we have for this claim is *defeasible*, such that colors and other secondary qualities may, in fact, turn out to be properties of phenomenal (or just *mental*) events. Now, it is important to distinguish here between introspective support for colors being instantiated in general, or *at all*, and introspective support for a *specific instantiation* claim like Harman’s (1990). I will argue that introspective support for this specific instantiation claim needs to have ‘epistemic priority’ over the relevant general instantiation claim, and that if there is *independent* introspective support for the relevant general instantiation claim, then Harmanian transparency is in trouble.
3.2: An argument against Harmanian transparency

The basic issue for defeasible introspective support for a specific instantiation claim is the possibility of independent introspective support for a relevant general instantiation claim, according to which there is a relevant kind of epistemic neutrality when it comes to which ontic categories or types of entity can instantiate the property at the actual world. Harmanian transparency is a good example of a thesis according to which we have introspective support for a certain broad category of entities that can instantiate colors which is incompatible, or at least in tension, with a different category of entities instantiating them at the actual world.

For this thesis to be correct, however, the introspective support for the relevant specific instantiation claim needs to have a kind of epistemic priority over the introspective support for the related general instantiation claim, otherwise the two introspective supports are in tension. A clear way for the relevant epistemic priority relation to obtain here is for the general instantiation claim to be merely a sound inference from one’s introspective support for the specific instantiation proposition.

Let us assume, for reductio, that we have introspective support for the notion that colors are instantiated by entities occupying external, physical, space. Since independent introspective support for the notion that colors are instantiated somewhere or other (whether in external physical space, or some sort of mental ‘quasi-space’) at the actual world, is incompatible with there being introspective support for colors being instantiated at a location in external, physical space, it follows that there can’t be such independent introspective support for the relevant general instantiation proposition. But suppose that one’s introspective support for the
relevant specific instantiation proposition is defeated, then it follows according to our assumptions thus far that the introspective support we had for the relevant general instantiation proposition is also defeated. Intuitively, however, even after we learn that there is no instantiation of a color property at a certain external location in physical space, we still have introspective support for that color property being instantiated *somewhere or other*. So, we get the contradiction in this sort of case that we both do and don’t have introspective support for the relevant general instantiation claim. Thus, Harmanian transparency, the claim that we have (defeasible) introspective support for the notion that colors are instantiated by entities occupying external space, is false.

The fundamental problem here, I think, is that we are plausibly in some sort of ‘cognitive contact’ with properties such as colors in hallucinatory cases such that even if we don’t self-evidently have introspective support for these properties being instantiated after the relevant introspective support for the specific instantiation claim is defeated, we can, nevertheless, *learn about* these properties through perceptual experience and introspective access to that experience. Indeed, this is why some have been attracted to the proposal that in cases of hallucinatory perceptual experiences/episodes of perceptual awareness, one is perceptually aware of and introspectively privy to, *universals themselves/uninstantiated properties.*

As Benj Hellie once casually pointed out to me in conversation, however, it is rather difficult to make sense of this suggestion, despite its obvious place in ‘logical space’. Moreover, once it is conceded that in hallucinatory cases we are in ‘cognitive contact’ with properties in some way or other, then it is tempting to suppose that this ‘cognitive contact’ is best worked out in terms of introspective support for ‘property-instances’/instantiations of universals, and that we can’t be
perceptually aware of, and introspectively access, universals independently of their instances. So, if there is a kind of independent first-personal epistemic support for the notion that we are at least in cognitive contact with properties such as colors in cases of hallucination, then this alone may put considerable indirect pressure on a thesis like Harmanian transparency.

Though I won’t attempt to settle the issue here, the subtle interplay between metaphysics and epistemology in this area clearly warrants more attention in future work. Overall, I’m inclined to think that Harmanian transparency, understood as the claim that we have introspective support for secondary qualities such as colors being instantiated by *externalia*, cannot be supposed to enjoy a kind of default status.

Only the weaker version of Harmanian transparency, namely, the view that we don’t have introspective support for secondary qualities such as colors being instantiated by mental events (alone), seems uncontroversial. The default view in this area, I think, should be that we *don’t have introspective support for any specific instantiation claim here*. In any case, we will now move on to address another major topic in consciousness studies, pain. Pain has often been thought to pose certain unique theoretical challenges in this area, and so it is a topic we would do well to concentrate on. In what follows, we will explore the issue in some depth, and I will ultimately argue for a novel theory of pain.

### 4.1: A problem about pain

Pain and other bodily sensations pose certain well-known theoretical challenges. One of the unique challenges posed by pain, in contrast to other bodily sensations such as nausea, arises from the wide range of bodily sensations we normally classify as pains. Here is Norton Nelkin (1994) on this point: “...consider the significant differences among phenomena we do call pain phenomena (those
experienced when cut, when suffering a toothache, or when having a headache). It is hard to understand what feeling they all share in common such that they are all pains.” Throbbing pains, sharp pains, dull pains, and shooting pains are all pains, but as Nelkin (1994) points out, such sensations can be very different from one another on a sensory level. So, what do these different sensations all have in common such that they are pains? Is there really some sensory quality or characteristic that all these different sorts of bodily episodes have in common?

One might answer that the relevant sensory property all these different episodes have in common is just ‘painfulness’ construed as a sensory, not an affective property of the relevant episodes. But this is to effectively give up on providing an informative answer to the question of what these diverse bodily episodes all share in common. We know that these sensorily diverse bodily episodes are all pain episodes, so if there is some sensory aspect that these episodes all have in common, then we must do better than say that the relevant sensory aspect is ‘pain quality’ or ‘painfulness’—understood as a sensory characteristic of the relevant episodes. If we can’t do any better than this, then we must content ourselves with a certain sort of primitivist view of pain. At this point, however, we should be considering more informative options.

4.2: The ‘painfulness’ theory of pain

One venerable theory in this area is that what unites all of these sensorily diverse bodily sensations under the category of pain sensation is not a sensory property of the relevant bodily episodes at all, but instead, a certain sort and/or level of negative affect or ‘bodily displeasure’ associated with them. It is this negative affect/unpleasantness/displeasure which is the property that unites these sensorily

---

diverse bodily episodes as pain episodes. We might call this the ‘painfulness’ theory of pain sensations. Now, the painfulness theory faces a number of important objections, particularly (putative) counterexamples where a bodily, or perhaps even an emotional occurrence, (allegedly) has the same negative affect/unpleasant character of pain, but isn’t a pain.

If there can be emotional events that have the same negative affect/unpleasantness associated with them as (bodily) pains, then clearly the painfulness theory of pain is incorrect. It may be urged, however, that the unpleasantness/painfulness property associated with certain bodily events is a different kind of unpleasantness/painfulness than the sort of displeasure associated with certain sorts of generally unpleasant mental events, such as a realization that one wasted the last ten years of one’s life, or a sudden feeling of inadequacy. If such unpleasant emotional and/or cognitive experiences are such that the displeasure associated with them differs in kind from ‘bodily displeasures’, then we can move on to consider putative bodily events other than pain that are arguably just as ‘painful’ as pain.

Certain bodily events such as one’s heart beating very quickly and irregularly, being too hungry to eat, and feelings of ‘movement’ had in an intense stupor, are a few obvious candidates for bodily events that may be just as painful as pains, but aren’t. It is, again, open to a defender of the painfulness theory to reply to these cases by suggesting that the sort of unpleasantness associated with these bodily events differs in kind from bodily events such as one’s knee snapping like a twig.

At this point, however, the theory seems to lose a fair bit of plausibility: intuitively, there is no difference in kind between the unpleasantness/displeasure of sensing a full bladder and a sharp pain from cutting one’s finger—there is only a
difference in degree of unpleasantness/displeasure. Uncomfortable bodily events such as pangs of hunger aren’t classed as pains since, qua uncomfortable, they differ in degree of unpleasantness from what we call ‘pains’. But not all of them do, and a bad enough hunger pang is, arguably, a pain in one’s stomach, precisely because it has reached a certain threshold of unpleasantness or displeasure for the subject that exceeds ‘very uncomfortable’. Nausea is made up of events that are often very uncomfortable, and some of these events may cross the threshold of discomfort to pain. That’s how unpleasant some of these events can be.

So, we have a broad outline of a painfulness theory of pain: a (bodily) pain is a bodily event that is displeasurable/unpleasant to a certain degree, namely, that degree beyond ‘very uncomfortable’. This is an important theory of pain. What’s more, it seems to me that David Armstrong (1962) once suggested a ‘rear-guard’ manoeuvre that can deal with one of the most prominent recent threats to this theory: the phenomenon of pain asymbolia.

4.3: Pain asymbolia and the painfulness theory

Pain asymbolia is a rare medical condition caused by damage to a certain area of the brain (the posterior insula—see Berthier et al. 1988) wherein sufferers claim to experience pains that aren’t unpleasant and ‘don’t bother’ them. In Marcelo Berthier et al.’s (1988) study, all six of the study participants avoided withdrawal as well as any characteristic expressive reactions such as grimacing, when exposed to noxious or injurious stimuli. None found exposure to any of the stimuli unpleasant, but all claimed that such stimuli did, indeed, cause pain. Assuming, then, that such rare individuals aren’t to be classified alongside those insensitive to pain (as seems plausible, see Bain, 2014, for a helpful summary of the reasons why), pain
asymbolics seem to present a counterexample to the painfulness theory: (genuine) pains that aren’t painful.

As Colin Klein (2015) emphasizes, however, pain asymbolics have important additional deficits that need to be taken into account as well, and this may suggest a different sort of interpretation of this case. In Berthier et al.’s (1988) study, and the pioneering work by Erwin Stengel and Paul Schilder (1928), the patients seemed worse than pain insensitives in learning to avoid circumstances that generally lead to bodily injury. These researchers also noted that some of the patients seemed prone to self-harm, and did not react appropriately to threats and warnings of impending bodily harm/damage, including both verbal and auditory threats.

Pain asymbolics do, however, seem capable of experiencing emotional displeasure (see Schilder and Stengel, 1928; Hemphill and Stengel, 1940), so it is not the case that these individuals are simply incapable of negative affect altogether. Klein (2015) argues that these additional deficits suggest that pain asymbolics have lost the capacity to care about their own bodily integrity, and that this can explain the apparent absence of negative affect associated with bodily injury. It is not, according to Klein (2015), that pain asymbolic’s pains aren’t unpleasant or displeasing, it is just that this displeasure doesn’t motivate in the context of a lack of concern about one’s own bodily integrity.

The above suggestion, however, is difficult to understand: why should a lack of concern about one's own bodily integrity block the unpleasantness/displeasure of pain from motivating one to avoid it and try to get it to cease once it begins? Klein’s (2015) ‘answer’ to this question at one point is that, “Pains motivate because we care about our bodies. Were we to stop caring—something that is ordinarily impossible, for good biological reasons—then pains would not matter. Asymbolics are an
unusual realization of this possibility.” 26 If this is the lynch-pin of Klein’s (2015) account, then I think it is in serious trouble.

The reason why is that, as I already suggested, the unpleasantness/displeasure of pain has motivational force, and Klein (2015) only acknowledges the motivational force of (awareness of) bodily damage/potential bodily damage. Once one appreciates that the unpleasantness/displeasure of pain is a reason for action, then we require an explanation as to why the unpleasantness of asymbolics’ pains don’t motivate them to act accordingly. If the unpleasantness/displeasure of pain were just the unpleasantness or displeasure associated with awareness of one’s own bodily integrity being compromised, then this would explain why asymbolics aren’t motivated by the unpleasantness of pain.

But such an instrumentalist view of the disvalue of pain affect is highly implausible and has the consequence that pain asymbolics pains aren’t unpleasant after all: they can only be so for the sake of the unpleasantness of one’s own bodily integrity being compromised, and pain asymbolics, ex hypothesi, don’t care about their own bodily integrity. So (as Bain, 2014 also argues), Klein (2015) seems to leave the motivational force of the unpleasantness/displeasure of pains unaccounted for.

4.4: Pain asymbolics don’t have pains after all—a defense of the painfulness theory

According to the painfulness theory, a bodily event that lacks the relevant negative affect isn’t a pain at all; the relevant negative affect is constitutive of pain. So, if we reject Klein’s (2015) and other related accounts according to which pain

26 Klein, 2015, p.500.
asymbolia is a kind of ‘bodily de-personalization’ disorder, then what should we make of pain asymbolics’ insistence that they have pains but their pains lack negative affect? Here is an interesting passage from Armstrong (1962) that I think sheds considerable light on this issue:

...the possibility of separating impression and reaction may explain what is happening in those situations where people report ‘pain’, but also report that they are quite unworried by the pain. Perhaps we can construe such extraordinary reports as implying that they feel something take place in their body, a feeling which they recognize would ordinarily evoke the pain-reaction, but which is not doing so in this case. The reaction is abolished, but not the impression. In such an unusual situation, a linguistic decision has to be taken as to whether they can be said to be feeling ‘pain’ or not.27

I think Armstrong’s (1962) suggestion that the relevant individuals ‘...feel something take place in their body, a feeling which they recognize would ordinarily evoke the pain-reaction...', is on the right track. Pain asymbolics remember which sorts of bodily sensations were such that, prior to their brain damage, bodily sensations of these sorts were painful/unpleasant. It should not come as too much of a surprise, then, that these individuals continue to refer to these sensations as ‘pains’ even now after they are no longer painful: they are communicating that the sensory aspect(s) of these sensation types hasn’t changed, and ‘pain’ is still as good a term as any for expressing this sensory persistence fact to the experimenters. Moreover, it is not for them to say whether or not there is some relevant sensory property that is common to all these sensations which would unite them as ‘pains’.

So, I think that pain asymbolics have good practical reasons for calling these relevant sensations ‘pains’ in certain contexts. But even if I am wrong about this, it would still be at least understandable why asymbolics should continue to

call these sensations ‘pains’ even if, strictly speaking, they aren’t pains anymore following their brain damage. Moreover, the knowledge that asymbolics have concerning which types of bodily sensations used to be painful is problematic in terms of how these cases have been studied, as I will now explain.

4.5: Pain asymbolia research and risk of bias

Asymbolics know that certain types of bodily sensations used to be painful. This represents a potential source of bias in the relevant experiments as patients may use this knowledge to give answers that, on some level, they suppose the experimenters want or are looking for. For this reason, familiar bodily sensations and circumstances that allow for asymbolics to utilize their relevant previous knowledge should be minimized. Ideally, asymbolics should be kept in the dark as much as possible as to why exactly they are being studied, and exposed to stimuli that, in ordinary individuals, would undoubtedly cause pain, but appear relatively neutral or innocuous so as to prevent asymbolics from using their previous knowledge, in combination with task demands, to provide an answer that they might think experimenters are looking for, or ‘seems right’, given their background knowledge about the relationship between certain types of bodily sensations and negative affect.

Thus, avoidance of familiar pain-inducing stimuli well-known to cause the relevant negative affect in ordinary individuals is badly needed. To my mind, this suggests a better way to study such individuals: expose them to ‘unorthodox’ noxious stimuli that cause pain in ordinary individuals, and then determine whether or not asymbolics classify the relevant bodily sensation as a pain or not. This would ensure that asymbolics cannot rely on their previous knowledge of the relationship between certain bodily sensations and negative affect in providing
answer. My hypothesis is that upon encountering unfamiliar sensory stimuli that are classed as pains by ordinary individuals, asymbolics would not call such bodily sensations pains—they would be unable to tell whether the relevant sensation is a pain or some other type of bodily sensation associated with a different affective valence in ordinary individuals.

To avoid the problem of guessing here, we could intersperse the normally pain inducing stimuli with (also unfamiliar) stimuli that normally induce a mere feeling of discomfort. No doubt the study design details would need to be more intricate and refined than this, but, as a general strategy, this seems to be a principled approach that could rather easily be applied here.

My suggestion, then, is that asymbolics can use their background knowledge about the relationship between certain types of bodily sensations and a certain type or level of negative affect, in combination with task demands, to answer questions concerning the presence or absence of a pain in a way that is not merely a subjective report. If there are genuine cases of ‘congenital indifference to pain’, then I think this suggestion gains further traction.

Congenital indifference to pain is rare a condition in which children do not find exposure to any normally pain-inducing stimuli to be unpleasant, and they lack any neuropathy or other abnormalities that would plausibly interfere with the sensory processing of pain, in addition to passing the usual tests for sensing and discriminating between various stimuli agreed by ordinary individuals to be pain-inducing. Importantly, these individuals don’t normally call any sensations ‘pains’ at all, and the evidence in favor of them having ‘pains that don’t hurt’ simply comes from the evidence for an intact and ‘normal’ functioning
of the sensory pathways we take to be relevant or potentially so in ‘pain perception’.

Unfortunately, as Toscano et al. (2003) point out, most cases of so-called ‘congenital indifference to pain’ diagnosed before the 1960’s were probably just cases in which patients had some sort of sensory abnormality that went undiagnosed, such that they were actually cases of ‘congenital insensitivity to pain’ (there have been numerous advances in detecting various subtle sensory/nervous system abnormalities that didn’t exist back then). A true case of congenital indifference to pain is one such that the sensory character of the (normally) painful bodily sensations is in no way damaged or degraded—if there is any evidence of this, then such a case can only be considered one of ‘insensitivity’ and not ‘indifference’ (for a challenging case in this regard, see Cox et al. 2006, and especially the ‘clinical note’ in the methods section).

Nevertheless, as Toscano et al. (2003) acknowledge, there has been at least one recent case in which no sensory abnormalities were detected and a non-sensory hereditary abnormality was detected (see Landrieu et al. 1990). So, if there are genuine cases of ‘congenital indifference to pain’, then it is unclear whether these individuals actually have pains at all, as we don’t get any sense from them that there is some common sensory core to a relevant range of different types of sensations which unites them as ‘pains’ (a point I will reiterate shortly). Moreover, the distinction between ‘congenital insensitivity to pain’ and ‘congenital indifference to pain’ simply assumes that pain is a sensory phenomenon, not an affective one.

It should also be noted that while damage to our ability to sense our own bodily events may eliminate (bodily) pain from the picture, this does not suggest
that pain is not a (sensed) bodily event that is unpleasant for its subject to a certain degree. If anything, the relevant research in both areas supports a picture according to which negative affect is at least partially constitutive of pain: one would expect that if pain were a purely sensory phenomenon, then those individuals with congenital indifference to pain would learn to recognize the relevant diverse sensory kinds of bodily events as ‘pains’, or as being such as to belong to a common kind of bodily event. But this is not the case; instead, it is as though these individuals do not know what pain is at all. My tentative conclusion, then, is that neither pain asymbolia nor congenital indifference to pain pose any major problem for the painfulness theory of pain; if anything, ‘congenital indifference to pain’, assuming there are genuine cases of it, may offer some support for the painfulness theory.

4.6: Why a headache isn’t a kind of pain: a consequence of the painfulness theory

The painfulness theory of pain has an apparently counterintuitive consequence: that bodily event-types such as headaches, aren’t kinds/types of pain. Let me explain. On the painfulness theory of pain, pain, qua type of bodily event, refers painful (unpleasant to a certain degree), qua type of bodily event. Those bodily event (tokens) that are painful are tokens of the type painful bodily event, and thus, according to the painfulness theory, tokens of the type pain. Now, a headache is not a kind of painful bodily event, it is a certain kind of bodily event in its own right that is painful. At least, it is a certain kind of bodily event that is typically so; perhaps there are asymbolics who have headaches that are not painful—this is not incompatible with the painfulness theory of pain, since a headache qua type bodily event, is a certain ‘sensorily manifest’ type of bodily
event. Indeed, this was presupposed in the two preceding sub-sections: pain asymbolics remember the sensorily manifest bodily character of certain bodily event-types such that when they have an ache, or a throbbing bodily sensation/event they can recognize it as an instance of the relevant bodily event-type, without this token event even being unpleasant at all. So, there are no different kinds of bodily pain, only different kinds of bodily events whose tokens may be of the type painful bodily event, and thus tokens of the type pain.

Why are certain bodily events-types such that (most of) their tokens are painful and thus pains, while others are not? Well, I think this should be viewed as an empirical question that is still being worked on to this day. Undoubtedly, this is why there is generally no problem with thinking of aches, and other rather difficult, but not impossible, to describe bodily event-types, being considered pains. Strictly speaking, however, only bodily events of the type painful bodily event are pains, and pain asymbolia teaches us that many bodily event-types are such that while the vast majority of tokens of these types are also tokens of the type painful, they aren’t all this way, as a matter of contingent fact.

The sad cases of pain asymbolia and congenital indifference to pain also drive home how lucky most of us are that the relevant kinds of bodily events, including, of course, those involving actual bodily damage, are such that tokens of them in us are also tokens of the type painful. Though we touched on the motivating aspect of painfulness earlier, it should also be noted that it is quite plausible that painfulness is inherently or intrinsically motivating—it is a ‘stand-alone’ reason for action due to its plausibly intrinsic or non-instrumental disvalue. Painfulness sucks on its own merits, and this explains why pain does too, if we adopt the painfulness theory of pain. To further appreciate the merits of the
painfulness theory, however, we need to consider an important and closely related alternative view in the vicinity: what is sometimes called the ‘composite view’ of pain.

4.7: The composite view of pain

According to what I, and others (e.g. Klein, 2015) call the ‘composite view’ of pain, a bodily event or just an event (whether psychological or bodily) is an instance of the kind (whether mental or physical and so on) pain, if and only if that (token) event has a certain ‘sensory’ or ‘sensorily manifest’ property or, perhaps, properties, and a certain ‘non-psychological/non-emotional’ kind of unpleasantness/‘negative affective’/displeasure, to a certain degree, namely, the/that degree beyond ‘very uncomfortable’/beyond the upper threshold of ‘mere discomfort’ (I don’t want to rule out some epistemic vagueness concerning this threshold such that we can, and perhaps even frequently do, make mistakes about whether something was just uncomfortable or actually painful). Now, here is a simple version of the ‘composite view’ of pain (CV):

Simple CV—

A token event, e, is a token of the type pain, if and only if e is a token of the type, bodily disturbance, and is a token of the type painful (unpleasant/displeasing to a certain degree): in other words, a token bodily event e is a token of the type pain iff e is a token of the composite type: painful bodily disturbance.

Now, Simple CV is clearly a starting point for an important view in this area. Its main problem, in my view, is that ‘bodily disturbance’ is either too broad and vague, in which case it is not really an alternative to the painfulness theory of pain, or it is too narrow and specific to be plausible. Hill (2009), a defender of the ‘bodily disturbance theory’ of pain, shows an awareness of this problem, but he doesn’t do much to resolve it. He writes, “...I will argue for a theory of the nature
of pain that I will call the bodily disturbance theory. This view identifies pains with peripheral disturbances involving actual or potential damage."\(^{28}\) Earlier on the same page, Hill (2009) doesn’t use the term ‘peripheral’, and just says ‘bodily’: ‘pains are bodily disturbances that involve actual or potential damage’.

In an attached footnote, Hill (2009) immediately says the following about this understanding of ‘bodily disturbance’:

As David Bennett has pointed out to me, there is a problem with the idea that pains are bodily disturbances that involve actual or potential damage — it often happens that bodily disturbances are felt as pains even though they do not pose threats to the integrity of the parts of the body in which they occur. Migraine headaches are examples. This shows that if we are to identify pains with bodily disturbances, we will need to characterize the relevant disturbances in broader, more inclusive terms. It will be necessary, for example, to invoke abnormally high values of degrees of pressure and inflammation. I do not think, however, that it is widely off the mark to characterize the relevant disturbances as ones that involve actual or potential damage. This characterization does justice to the normal case. Because of this, and because of its simplicity, I will continue to rely on it in the sequel.\(^{29}\)

I think that Hill (2009) underestimates the scope of the problem here: bodily disturbances don’t just intuitively encompass other bodily event-types that are normally thought of as pains, such as migraines, but also bodily events such as uncomfortable but not painful ‘twinges’ and all manner of bodily sensations whose tokens are (for the most part anyway) unpleasant, but not painful. Why isn’t a sudden pang of hunger or a sudden need to visit the restroom a ‘bodily disturbance’? Well, presumably, ‘bodily disturbance’ is being used here in such a way as to refer to bodily events that are just in some way potentially or actually deleterious to the body—roughly as Hill (2009) says. But vomiting is an example of a bodily event that is potentially deleterious or harmful to the body, yet we don’t want to count it, and others like it, as pains.

\(^{28}\) Hill, 2009, p.177.

\(^{29}\) Hill, 2009, p.177.
Hill (2009) could insist that vomiting isn’t potentially damaging to the body in some restricted sense of ‘damage’ according to which not just anything that does or might compromise one’s bodily integrity constitutes bodily damage or potential damage, but then it is not clear what is being talked about anymore. ‘Damage’ is usually just synonymous with ‘harm’ or ‘impairment’ and so on. And there are plenty of bodily events involving actual or potential bodily harm or impairment that, intuitively, aren’t pains such as bending one of my legs into a certain awkward position that is uncomfortable, but not painful.

Moreover, bodily harms and impairments can be quite subtle and so we only want to count certain more ‘serious’ harms and impairments here as the relevant ‘bodily disturbances’. But there is a certain arbitrariness problem that arises from just saying, as Hill (2009) does, that ‘abnormally high values of degrees of pressure and inflammation’ should also probably be counted as ‘bodily disturbances’ in the relevant sense. Plausibly, we are only wanting to count these ‘abnormally high values of degrees of pressure and inflammation’ as pains because of their association with the relevant negative affect—high values of other potentially deleterious or deleterious bodily activity presumably won’t be counted as pains, such as the overproduction of certain neurotransmitters, high blood pressure, and other sorts of ‘internal’ and potentially or actually deleterious bodily activity.

So, as I said earlier, I think that bodily disturbance qua type of bodily event, is either too broad and vague to be relevant in partially characterizing/constituting pain, or it is too narrow and specific to plausibly be a component pain. Having said all that, I don’t take this to be the last word on the matter, and there may yet be other ways of developing the bodily disturbance theory, or a bodily disturbance component theory of pain that have a better chance at success. But I am inclined to
think that, as it stands, the bodily disturbance theory, and bodily disturbance versions of CV, face an uphill battle. Bodily disturbance versions of CV such as Simple CV, also seem to get ‘swamped’ by painfulness/ a certain kind of negative affect.

4.8: Simple CV and the painfulness swamping problem.

According to Simple CV pain qua type, is just the type painful bodily disturbance. We have seen in the previous section that there are problems with how to make sense of this notion of ‘bodily disturbance’, but let’s set them aside here. Now, a further problem for Simple CV is that the notion of a bodily disturbance doesn’t seem to in any way constrain painfulness: we can’t point to any plausible cases where painfulness is present, and bodily disturbance is absent, such that we intuitively want to conclude that the case is not one of pain.

Indeed, any case in which painfulness is present is one we seem to want to call a case of pain. But if that is correct, then the bodily disturbance component of Simple CV is swamped by its painfulness component. Nikola Grahek (2007), however, has argued that there is at least one real-world counterexample to the claim that painfulness (of the relevant sort, not emotional or psychological painfulness) is such that any event that instantiates it is a pain.

4.9: Painfulness without pain?

Grahek (2007) draws our attention to a case studied by Ploner, Freund, and Schnitzler (1999) in which a patient with a certain kind of brain damage is described as exhibiting painfulness or pain affect, without pain. Here is the relevant passage from Ploner et al. (1999) for our purposes (quoted by Grahek, 2007, as well):

Pain thresholds were 200mJ for right hand and both feet. Evoked pain sensations were characterized as “pinprick-like” and were well localized within 2-3 cm. For left hand, up to an intensity of 600 mJ, no pain sensation could be elicited. However, at intensities of 350 mJ and more, the patient spontaneously described a “clearly unpleasant” intensity dependent feeling emerging from an ill-localized and extended
area “somewhere between fingertips and shoulder,” that he wanted to avoid. The fully cooperate and eloquent patient was completely unable to further describe quality, localization, and intensity of the perceived stimulus. Suggestions from a given word list containing “warm,” “hot,” “cold,” “touch,” “burning,” pinprick-like,” “slight pain,” “moderate pain,” and “intense pain” were denied nor did the patient report any kind of paraesthesia.30

Now, as Klein (2015) rightly points out, ‘clearly unpleasant’ here doesn’t add up to painfulness, so this is the first clear obstacle to considering this to be an actual case in which painfulness is present without any clear bodily event or sensorily manifest bodily event occurring such that it could be a pain event. If we waive this issue, then it may be (I will ultimately claim, incorrectly,) argued that we have a case of painfulness without pain. As Grahek (2007) describes it,

...we have a clear and vivid picture of what pain experience comes to when it is deprived of its sensory-discriminative component while its affective and behavioral components are kept intact. The subject who has lost the sensory-discriminative capacities of pain is no longer able to precisely localize noxious stimuli applied to his body. The most that he can do is refer to a diffuse area “somewhere between fingertips and shoulder.”31

I basically agree with Grahek’s (2007) description of this case, but I also don’t think it supports painfulness without pain: even if the patient can’t, in any helpful way, characterize the sensory character of the relevant bodily event occurring—they do seem to think that there is a bodily event of some sort occurring, albeit one that can only be vaguely pinned-down as occurring somewhere on the arm (where ‘arm’ includes hand).

Now, as to why the patient is reticent to call this bodily event anything at all, given that the painfulness theory entails that they should call such a ‘vaguely localizable’ painful bodily event, a pain, we might speculate that in such an aberrant case in which the relevant part of one’s own bodily is basically devoid of any sort of

‘sensory character’, it is natural that one would be conservative in attesting to the nature of the bodily event just had, even if, strictly speaking, it should be considered a pain since if one feels something or other going on somewhere on one’s arm, and it is painful, then that event/ ‘bodily what-not’ is a pain. Ultimately, then, I think that this sort of empirical case for painfulness without pain fails. Thus, the painfulness swamping problem for a composite view of pain such as Simple CV cannot be easily met, and this also speaks in favor of the painfulness theory of pain, given the initial plausibility of a composite view of pain.

4.10: Isn’t it just a verbal issue? A deflationist objection to the painfulness theory

Someone sympathetic to a view along the lines of Simple CV (including representationalist developments of it) or perhaps even just a neutral party here may think that my interest in the painfulness theory of pain somewhat obtuse: there are bodily events that are bodily disturbances in a broad sense of the term, and bodily events that are painful—we can call all painful bodily events ‘pains’ but presumably we could do this for bodily disturbances as well; either way, the phenomena of interest include bodily disturbances of various sorts and painful episodes. This attitude is fair enough, except that pain itself, qua type of event, has always been of intrinsic interest as well. And, according to a view like Simple CV, there is such an event-type, in addition to just bodily disturbances and painful events. The painfulness theory has the consequence that there is no such additional event-type: there are painful events, and bodily disturbances, with some overlap here.

So, there is a substantive philosophical issue at stake here, contrary to what the initial ‘deflationist’ idea implied. Moreover, much of the interest in the notion of ‘bodily disturbances’ has not been for its own sake, but for the sake of the idea that
this broad sort of bodily event might be able to unite the otherwise ‘sensorily
disparate’ bodily event-types we often call ‘pains’ (strictly speaking, inaccurately,
according to the painfulness theory). As argued in 4.7 of this chapter, however, this
hope seems rather forlorn. So, the dispute between the different theories here is not
merely a verbal one, but, instead, is highly substantive.

4.11: Pains as sources or causes of a certain kind displeasure to a certain
degree

So far, I have assumed that there is a kind of painfulness, ‘bodily’
painfulness, that differs from the sort/kind of painfulness associated with certain
emotional and/or psychological events. This, however, is open to dispute, and an
important alternative here is the suggestion that the same sort of negative affect is
present in both cases, and it has a kind of intentional ‘evaluative’ structure such that
the difference simply lies in a difference in intentional objects: in the one case I
‘evaluate’, and thus mentally represent, a certain emotional or psychological event as
‘bad’ or ‘unpleasant’ to a certain degree, in the other I do the same except I represent
a bodily event or whatnot as being ‘bad’ or ‘unpleasant’ to that same relevant degree.
I will now argue, however, that we can, to a certain extent, by-pass this issue and
other related issues in this context by adopting a non-standard (and perhaps novel)
version of the painfulness theory.

Thus far, I have used the terms ‘bad’, ‘unpleasant’, ‘displeasure’, and
‘painful’ more or less interchangeably, but the time has come to be more precise. In
my view, painfulness is to be identified with a certain kind or sort of feeling, namely,
a certain kind of displeasure. The view I am most attracted to here is that pain qua
type is just the type: source or cause of (a certain kind of) displeasure. This is a bit
of a departure from our earlier discussion, but it is not in tension with it.
One of its virtues is arguably that it doesn’t identify painfulness with awareness or representation of a certain bodily event or whatnot as ‘bad’ or ‘unpleasant’, but nonetheless accounts for the importance of such bodily awareness/representations: when we represent or are aware of a bodily event as ‘unpleasant’ we are aware of or representing it as a source or cause of displeasure, in other words, we are representing it as a pain.

This honors the intuition that we can be interoceptively aware of or represent pains, not just bodily events that, as it were, happen to be pains, thanks to some coalescence of pain affect and some sensory property or properties of that event. It also honors the intuition that pain is, at least typically, a bodily event, not a psychological phenomenon like representing or being aware of a (non-pain) bodily event. So, when we are interoceptively aware of or (appropriately) represent certain bodily events or what-not as being unpleasant to a certain degree, then we are aware of or representing them as pains, because pains are just causes or sources of a certain kind of displeasure to a certain degree. It is crucial to this theory that ‘displeasure’ is understood as the opposite of ‘pleasure’ and as a feeling that doesn’t (at least on the surface) have an intentional structure.

This, in my view, has always been the sort of half-hidden truth about pain that has tacitly motivated qualia-freak accounts of it. When we experience pleasure or displeasure, what’s true is that we generally represent or are aware of something or other as its cause or source. This is true of pain: when painfulness is instantiated in us, we seem to always take something or other to be its cause or source. But we can leave it open that this can, as a matter of fact, fail to occur. Painfulness, on this account, is such that whenever it is instantiated, pain is instantiated, because something must have caused this feeling to occur. Whatever causes it to occur is a
pain, and sometimes, of course, we can have multiple sources/causes of the relevant displeasure occurring. We happen to be so constructed as to be quite good at detecting various sources/causes of the relevant feeling, and this feeling is intrinsically motivating. Pains on this account, are generally motivating in so far as they are sources or causes of this intrinsically motivating feeling.

4.12: Objections to the pains as causes of displeasure view

An important challenge to the pains as causes of displeasure version of the painfulness theory of pain is the suggestion that whatever negative affect we associate with pains, it is a ‘manifest quality’ located wherever pains are located, contrary to the ‘pains as causes’ view, according to which only the cause of the relevant negative affect, is located where pains are located: the ‘unpleasantness’ of pain being just the source of a certain kind of displeasure or, perhaps, displeasurable feeling, to a certain degree (namely, as stated earlier, that beyond ‘very uncomfortable’ in the relevant sense of ‘discomfort’ that refers to the non-emotional, non-psychological sort).

My reply to this objection, however, is that it is actually quite unclear whether there is phenomenological/introspective evidence for such negative affect ‘quality’ being instantiated at bodily locations—if there is, then it’s not aptly referred to as ‘displeasure’, I think, because displeasure, qua feeling, isn’t ever phenomenologically supported as being located at certain parts of our bodies, such as at the tip of one’s right index finger. In my view, those who find this sort of objection tempting may be tacitly assuming a picture of pain according to which there is something about the relevant sensations which is inherently/intrinsically displeasing to us, a kind of ‘intrinsic nastiness’, these sensations have such that they are sources of displeasure. But I have already argued that there is no obvious
introspective/first-personal support for a ‘sensory pain quality’ that unites the relevant disparate sensory/sensorily manifest types of bodily events such as headaches and ‘shooting pains/sensations’.

A related objection has it that if pains are just sources/causes of a certain kind of displeasure to a certain degree, then we would expect that this would be introspectively/phenomenologically accessible to us, but clearly, it is not. I’m inclined to agree that there is not some sort of direct phenomenological support for this proposal, but there may yet be, and I think there is, indirect introspective/phenomenological support for it, garnered through phenomenological reflection on the place of displeasure itself in the theory.

Moreover, I disagree that we should expect there to direct introspective/phenomenological evidence for any theory of pain, though I acknowledge, of course, that there could be such support, and acknowledge that the ‘pains as causes’ view is probably not a good candidate for a view that enjoys such support. I’m inclined to think that there really isn’t direct phenomenological/introspective support for any relevant view over others, but this is obviously a rather large topic that deserves its own examination. Unfortunately, I won’t be undertaking it here.

A final objection I’ll consider here suggests that an alternative development of the painfulness theory is to be preferred: “it would be more ecumenical to work with ‘badness’ here, since we have a better understanding of it than we do of some rather amorphous kind of ‘feeling’, and we could just say that a pain is an event that is bad for the subject to a certain degree”. My response is first to point out that pain can’t just be understood as the type: bodily event that is bad for the subject to a certain degree, as many bodily events can be ‘bad’ to the same degree that pain
events are ‘bad’, but they are certainly not pains, unless we mean something special here by ‘bad’, such as ‘bad in a certain way’, but this is not a promising development.

Now, it could be held that pains are bodily events that we represent as being bad to a certain degree, but this is unpromising unless, again, we mean something ‘special’ here by ‘bad’, since, plausibly, representing a bodily event as bad to certain degree is different from representing it as ‘painful’ in the relevant sense. For this reason, those who wish to think of the negative affect of pain simply in terms of badness and not in terms of displeasure (which, on a kind of ‘old-fashioned’ but intuitively plausible view, is a feeling that intrinsically/inherently bad), rely on a separate sensory or sensorily manifest property of the relevant bodily events to arrive at a reasonably plausible version of a composite view of pain.

On this view, my suggestion that ‘badness’ isn’t enough painfulness is rebutted by the integration of a component such as being a bodily disturbance, or a certain broad sort of sensory event, such that the conjunction of being a bodily disturbance and being represented as bad for one, is taken to deliver both painfulness and pain. It is not clear to me whether badness + some sort of bodily/sensory property/kind could ever plausibly amount to the relevant negative affect of ‘painfulness’ which, intuitively, is bad, or, at best, a kind of badness itself, not badness + some non-evaluative property. But even if we find it reasonably plausible that it could, we have already seen that there are major obstacles to success in locating the bodily/sensory half of the equation here, so to speak, that could launch such a view.

So, I think there is a good reason why philosophers have preferred to use the term ‘unpleasant’ when describing pains, not just ‘bad’. Obviously, pains are ‘bad’
but they are bad because they are unpleasant. And ‘unpleasant’ here, I have argued, should be interpreted as: a source/cause of displeasure. Pains are a source/cause of displeasure, this much, we should all agree. The controversial view I have tried to motivate and defend here is that a pain just is a source/cause of (a certain kind of) displeasure to a certain degree.

With that, I propose to leave the topic of pain—hopefully the ground clearing accomplished in this section, as well as the view developed here, will make future theorizing in this area less ‘painful’. In the following final section of this chapter, I introduce and make some preliminary comments on what is, perhaps, the most popular topic in consciousness studies: reductionist approaches to phenomenal properties such as colors as we understand/know them through introspection. We take this topic up in more depth and generality in the upcoming final chapter of this thesis.

5.1: Physicalism and reductionist approaches to phenomenal character

Much of the interest in phenomenal character and introspectively accessible properties of phenomenally conscious events has been in whether or not the relevant properties are metaphysically and explanatorily irreducible. Unfortunately, this has come at the cost of considerably less attention being paid to unpacking the basic phenomena in need of accounting for here. For example, as we saw in the previous section of this chapter, it would be a mistake to simply ask ‘is pain irreducible?’ without alighting on the issue of pain affect or ‘painfulness’, and the possibility of some sensory or bodily property that events of this type might have in common (I argued that there is no such sensory ‘pain quality’—thus it would need to be thought of as primitive, a suggestion I also reject in favor of the ‘painfulness theory’ of pain).
The above sort of problem has been rectified to some extent in recent years as it became clear to many that intentionality is crucial to an initial understanding of, *inter alia*, ‘sensory qualities’, such as ‘sensory redness’ that were previously held, as a kind of default view, to be candidates for metaphysical and perhaps even explanatory reduction to physical/brain-based or functional-role properties. Intentionalist theories ‘kicked sensory qualities back out into the world’ and/or held that such properties could be metaphysically and explanatorily understood in terms of perceptual awareness of/perceptually *representing* certain well-behaved physical properties such as, in the case of color, ‘surface-reflectance profiles’, and so on.

Such issues came to attract a fair amount of attention, and this was a good thing, since, *inter alia*, it reinforced the point that we need to do at least some rigorous theorizing about ‘problematic’ phenomenal properties at the *psychological level itself*, prior to considering whether such properties might also be explanatorily and/or metaphysically accounted for by properties at some other level, such the structural/functional. Mental notions ‘as we find them’ in introspection, aren’t obviously reducible to the functional, so it is important to avoid conflating the psychological (or, at the very least, the *phenomenal* psychological) level of understanding with the functional.

This, in spite of a good deal of encouragement in the past for a view according to which *more or less everything other than* some narrowly circumscribed range of properties of phenomenally conscious events, and phenomenal consciousness itself, is *conceptually functional* or causal-role relational. Those with reductive aspirations, however, rarely, if ever, attempt a comprehensive treatment of the different ‘families’ of phenomenal properties/event-types. Among other families, *phenomenal emotions*, have often been left out of the picture, or been given a
superficial and (in my view anyway) rather implausible representationalist treatment (but see Hill, 2009 for a better effort than usual in this regard).

Fundamentally, I think the issue of reducibility when it comes to phenomenal properties and phenomenal character overshadowed the issue of how we should think about such phenomena on their own terms. Perhaps the best theory of (phenomenal) pain in its own terms will ultimately favor or even entail its irreducibility, but if we are only concerned with the matter of whether or not pain is irreducible, then much conceptual and metaphysical preliminary work, such as the work done in the previous section, gets short shrift. And this is to the detriment of arriving at a solid overall understanding of the phenomenon/phenomena that we can sensibly inquire into regarding reducibility.

The same could probably be said for ‘viewer-dependent properties’ of perceptual awareness discussed at various points in section 2 of this chapter, as well as ‘innerness’ from 2.5 of this chapter, though the former has received a fair amount of serious engagement in recent times. In any case, I think it can be problematic to engage the issue of reducibility when we haven’t done much theorizing at the psychological level itself. No doubt some will suppose that this isn’t much of a problem when the issue is, specifically, reducibility to the physical or functional, but, at least in some cases, this seems rather unclear. We will address the issue in the next chapter on the topic of the reducibility of phenomenal consciousness and phenomenal properties.
1.1: Phenomenal consciousness, first-person warrant, and the explanatory gap

If phenomenal consciousness either is or entails first-person warrant, then by virtue of its being this way/playing the relevant *epistemic role*, I argue that phenomenal consciousness entails an ‘explanatory gap’ between it and ‘the physical’ *qua* structural/functional. This point was introduced and briefly discussed in chapters 1 & 2, and we return to it here in the context of whether or not phenomenal consciousness is *reducible* to the physical. I have already suggested in previous chapters as well that the issue of whether or not phenomenal consciousness itself *qua* psychological event-type or ‘mere property’ is reducible to the physical, should be carefully distinguished from the issue of whether or not ‘phenomenal properties’ entail an explanatory gap(s).

I will argue in this chapter that plausibly neither phenomenal consciousness, nor certain phenomenal properties, are *conceptually reducible* to the physical, but for different reasons. In the case of phenomenal consciousness, the explanatory gap is due to the problem of explaining any plausible development of first-person warrant in structural/functional terms. In the case of certain phenomenal properties, such as ‘sensory colors’/colors as we understand them through introspection, the explanatory gap is between structural/functional properties and phenomenal properties that we take (through introspection) to be non-relational properties or *qualities*. There is also
a potential explanatory gap between certain relational phenomenal properties and structural or functional properties.

Before we come to examine phenomenal properties in the context of epistemic/conceptual reducibility, we will examine phenomenal consciousness itself. We don’t require anything beyond our conditional introspective access understanding of phenomenal consciousness, for it to be the case, as I argued earlier, that there is ‘no obvious reductive explanation’ of it. In this chapter, I will be arguing for a stronger anti-reductive explanation view here that may require some additional assumptions for the argument to go through. The key, however, remains first-person warrant itself—it is this normative epistemic property of phenomenal consciousness, or property entailed by phenomenal consciousness, that underwrites my view that phenomenal consciousness itself is not conceptually reducible to the physical qua structural/functional.

1.2: The distinctive source of the explanatory gap problem for phenomenal consciousness

The explanatory gap problem for phenomenal consciousness has generally been located at the level of a putative non-epistemic property of (some) mental events. While this property is often assumed to be, in some sense, introspectively accessible, I briefly argued in chapter 2 that this is far from clear. My assumption in what follows is that phenomenal consciousness is not primarily introspectively accessible, and that the conditional introspective access characterization of phenomenal consciousness reflects the concept in the vicinity we often do and, more importantly, should operate with when it comes to phenomenal consciousness (we must, again, be careful to distinguish between ‘phenomenal character’ and ‘phenomenal consciousness’).
If we had some other concept of phenomenal consciousness acquired solely through ‘first-personal reflection’ and/or ‘inner-attention’, then this would very likely be the sense of ‘phenomenal consciousness’ that interests us. But, unlike the case of phenomenal character/phenomenal properties, it is not compelling that we do possess such a concept when it comes to phenomenal consciousness itself. If we did possess such a concept, this would plausibly be just as obvious to us as it is in the case of many of our concepts of phenomenal properties like our concept of the painfulness/negative affect of pain, but it’s not obvious, so we (plausibly) don’t.

So, we will assume here that the right starting point for the relevant phenomenon to be explained is an epistemic one: our understanding of phenomenal consciousness lies in terms of a certain distinctive epistemic property/role: this property either is or entails, first-person warrant. Now, if all there was to the problem of phenomenal consciousness was this epistemic property/role, then we might think that the problem of phenomenal consciousness is just an instance of the problem of explaining normative properties in non-normative and, specifically, structural/functional, terms. As I briefly argued in chapters 1 and 2, though at least enough to ensure that the relevant understanding of phenomenal consciousness meets what I called the ‘no currently obvious reductive analysis’ constraint, this doesn’t show that phenomenal conscious (or, better, the relevant notion thereof) raises a unique explanatory gap.

It is important to note, then, that the conditional introspective access understanding of phenomenal consciousness/the understanding of phenomenal consciousness as either entailing or constituting first-person warrant, is, on the surface anyway, physicalist/reductionist-friendly. Those with the relevant reductive aspirations may even be liable to suppose that a fully reductive explanation of
phenomenal consciousness should ultimately be possible here (assuming the relevant epistemic understanding of it), since even high-profile anti-physicalists such as David Chalmers (1996; 2012), have thought that a reductive explanation of normative phenomena in structural/functional terms has a good chance at success.

I’m inclined to disagree with Chalmers (1996) on this, but the point remains that even if such a reductive analysis can’t succeed, phenomenal consciousness will no longer be thought of as posing a unique challenge to reductive explanation in this context. And this point remains physicalist/reductionist-friendly, since it is one less epistemic gap to worry about when it comes to arguing for a physicalist metaphysics or a related reductive monist metaphysics.

I will argue in this chapter, however, that on the major options for working out the relationship between phenomenal consciousness and first-person warrant, we end up with a more robust epistemic role for phenomenal consciousness. Moreover, I will argue that it is one which generates or gives rise to a unique (or plausibly so) explanatory gap, as it suggests that only a certain primitive property is able to play the relevant epistemic role at the actual world.

2.1: Self-support and awareness accounts of first-person warrant

Broadly speaking, there are two major ways of developing first-person warrant: first-person warrant may be developed as a uniquely self-supporting form of epistemic support, or it may be developed as a form of awareness of either phenomenally conscious events or simply mental events that provides (non-inferential) epistemic support for the occurrence of the relevant events. As long as first-person warrant is understood as being a form of non-inferential epistemic support/ ‘evidence’ these are the two main options. And there is both intuitive and
theoretical support for first-person warrant being a kind of non-inferential epistemic justification/warrant/support/.

If first-person warrant is taken to be a kind of inferential epistemic support, then we face the difficult task of articulating a plausible inferential structure here: seeing that the gas gauge reads low provides me with inferential support for the proposition along the lines of, ‘the car I’m driving is low on gas’; what might such indirect epistemic support for a proposition such as, ‘this pain is occurring in my right big toe’, consist in? It can’t be some inference from my own bodily behavior or the like, since it would then no longer be first-person warrant at all, but a publicly accessible third-person inferential epistemic warrant. Moreover, it can’t be some inference from another mental event of mine since (sheer implausibility aside in this particular example) this just passes the buck: we then need to address our epistemic warrant for this other mental event.

In general, while some phenomenal events may be such that we have a kind of uniquely first-personal inferential epistemic support for them, it is highly plausible that there are at least some phenomenal events for which we have non-inferential epistemic support that epistemically grounds the relevant inferential first-person epistemic support. If that is so, then we can allow that while there are other sorts of first-person warrant that can be developed along very different lines from that which we will consider here, the sort of first-person warrant we are concerned with is non-inferential first-person warrant. And, in this context, it is plausible that we do just have the two major options mentioned above for developing this sort of first-person warrant. At any rate, this is what I will endeavor to show in what follows.
2.2: Self-support and awareness accounts of first-person warrant are the only games in town

In the previous section, we clarified that the relevant form of first-person warrant we are concerned with here is *non-inferential* or ‘direct’. Now, if we think about non-inferential or direct epistemic support for a proposition along the lines of ‘a certain mental event $e$ is occurring’, it seems clear that the options are quite constrained. Either we are in some sort of ‘cognitive contact’ with $e$ such that this form of cognitive contact provides direct/non-inferential epistemic support for the relevant proposition, or $e$ is such that it provides ‘evidence of itself’/is epistemically self-supporting. ‘Awareness’ and ‘cognitive contact’ are similar notions in this context, such that a suitably broad notion of ‘awareness’ plausibly captures this first option: the relevant form of awareness here might be ‘purely cognitive’ or ‘quasi-sensory’, ‘non-propositional’ or ‘propositional’, and so on.

So, if these two options basically exhaust the field, then I understated the case earlier in claiming that the awareness and epistemic self-support options are the major ones in this area. We needn’t commit to so strong a view here, but I must confess that I can’t think of any additional options that don’t fall under one of the two. I am thus led to the conclusion that, though I can’t rule out some additional option beyond these two, the awareness option is broad enough that it, alongside the self-support option, are the ‘only games in town.’

2.3: Clarifying the self-support option

According to the self-support option, first-person warrant is a special sort of epistemic support: it is a non-trivial form of epistemic support that a fact provides itself. ‘Non-trivial’, since in a trivial sense, every fact supports/favors itself—so this is clearly not the intended epistemic view here. If we reject the view that there is
such a form of non-trivial epistemic self-support, then we reject the relevant option for developing first-person warrant, and are plausibly left with just the awareness option. I am open to the idea that we can’t make good sense of the relevant notion of ‘non-trivial’ epistemic self-support here, and that this option may, ultimately, just collapse into the awareness alternative.

Perhaps a tacit understanding of having first-person warrant in terms of attending to it provides a source of illicit epistemic support here: its epistemic enabling role may be covertly shifted in our thinking to an epistemic grounding role and back again, leaving an unaccounted for ‘epistemic reside’ that we confusedly explain in terms of ‘non-trivial’ epistemic self-support. So, when we consider the issue of whether phenomenal events provide some sort of distinctive non-trivial self-support, we need to avoid this potential hazard. If we let attention do double-duty here as both epistemic support and epistemic enabler, then we can easily be fooled into supposing that phenomenal events are epistemically special in providing non-trivial epistemic self-support—that’s the sort of error just described in which we illicitly lean on attention as epistemic support, only to subsequently retract and restrict its role to an epistemic enabler, suggesting that phenomenal events themselves are providing the relevant epistemic support.

2.4: Arguments for first-person warrant as non-trivial self-support

Once we recognize this possible slippage with respect to ‘inner’ attention, it seems problematic to simply lean on the ‘intuitive plausibility’ of the idea that phenomenal events are such that they are ‘evidence of themselves’. We need to make a case for the relevant non-trivial epistemic self-support here, but this task seems far from easy. One way of arguing for the relevant epistemic view is by elimination: assuming that there plausibly is non-inferential first-person warrant,
then if some apparently general sort of argument against the awareness option succeeds, then that only leaves the self-support option standing.

In terms of positive arguments, there is a kind of regress or epistemic-buck passing argument to the effect that some relevant mental event in the vicinity must be epistemically self-supporting or self-standing in order for us to adequately explain the first-person access we have to phenomenal consciousness itself, but I pointed out earlier on that it is not particularly plausible that we do have first-person epistemic access to phenomenal consciousness itself. Moreover, it is unclear why such alleged first-person epistemic access to phenomenal consciousness itself couldn’t just be some form of awareness that we lack (non-inferential) first-person access to, thereby averting the potential for an awareness regress here.

A more promising alternative (positive) strategy here is to argue that the epistemic strength of first-person warrant is such that it is best accounted for by the view that it is non-trivial epistemic self-support, not some form of awareness of a relevant (mental) event. While the details of such an argument could probably be developed in a number of different ways, I am inclined to think that, at the very least, there is no decisive reason that goes against a kind of ‘primitive higher-order awareness’/ ‘higher-order acquaintance’ account of first-person warrant here in favor of the non-trivial self-support alternative.

Ultimately, it seems to me that those who have been attracted to the non-trivial epistemic self-support view, such as Chalmers (1996), argue more or less from elimination here. Chalmers (1996) takes it to be plausible that phenomenal consciousness itself is/constitutes (non-inferential) first-person warrant, and that higher-order treatments of phenomenal consciousness fail since, inter alia, we have non-inferential first-person access to phenomenal consciousness itself, and we don’t
uncover higher-order awareness through such epistemic access, as we would expect. Again, however, this sort of strategy assumes that phenomenal consciousness itself is introspectively accessible.

In any case, let us suppose that we can, indeed, make sense of the notion that phenomenally conscious events are such that they are non-trivially epistemically self-supporting, and that this option should be favored over the awareness alternative. If this is correct, then I argue we plausibly have a (unique) explanatory gap here.

2.5: First-person warrant as non-trivial self-support and the explanatory gap

Plausibly, there is no conceptual entailment from structural:functional truths to normative epistemic truths such as ‘(the instantiation of) phenomenal consciousness provides non-trivial epistemic self-support for certain propositions’ (assuming arguendo, that this is such a ‘normative epistemic’ truth). This is because in order for such an entailment to obtain, there should, at least in principle, be some structural or functional analysis in the offing here.

But there is not much of an inkling as to how even a normative epistemic notion such as evidence could be roughly analyzed in structural:functional terms without eviscerating it by taking evidence to be ‘truth-conduciveness’, or something along such lines. ‘Truth-conduciveness’ seems like a notion that might, in principle, admit of a functional analysis; we can broadly envisage how such an analysis might go that doesn’t immediately change the subject and result in an ‘eliminativist’ view, as a pure truth-conduciveness analysis of evidence plausibly does.
But if this is basically all we are saying about first-person warrant, then, as previewed earlier, phenomenal consciousness/a property entailed by (instantiations of) phenomenal consciousness (such as ‘acquaintance’) does not pose a unique explanatory gap problem here. If first-person warrant is/represents a unique form of epistemic support, however, then it may be that this gives rise to a unique explanatory gap between the structural/functional and the normative epistemic.

The matter is somewhat fraught, however, as it is difficult to come to grips with the relevant epistemic notion on its own explanatory level: what can we say in ‘epistemese’, as it were, that would help unpack this notion of epistemic self-support that isn’t simply the kind of ‘self-support’ that the fact that Nadine is walking across Gordon Street provides for the fact that Nadine is walking across Gordon street, mutatis mutandis, for other banal non-phenomenal facts? Not much, I’m inclined to think.

Ironically enough, if this sort of epistemic support is legitimately primitive at its own explanatory level, then it does seem quite plausible that we would have a unique explanatory gap on our hands. Still, this idea may be liable to arouse the suspicion that the relevant epistemic notion is a kind of theorists’ ‘hack-job’. But if we do accept the development of first-person warrant along the lines of non-trivial self-support, then I think it is plausible that we secure a unique explanatory gap in this context. Let us now consider how the alternative way of developing first-person warrant fares in this regard.

2.6: The awareness account of first-person warrant and the explanatory gap

On the awareness option for developing first-person warrant, our non-inferential epistemic support for the relevant present-tense psychological
propositions is/derives from some form of direct awareness relation we stand in to those mental events that figure in the relevant propositions/contents. So, if a mental event \( e \) is such that its subject, \( S \), is directly aware of it, then the idea here is that this constitutes first-person warrant for \( e \) (a warrant that, presumably at least as a matter of contingent fact, can only be \( had \) by \( e \)'s subject).

Now, there are obviously a number of ways of developing the relevant form of ‘direct awareness’ that we allegedly have to at least some of our own mental states/events here. But this awareness very plausibly needs to be ‘special’ here as phenomenally unconscious ‘first-order’ (perceptual) awareness doesn’t provide normative epistemic support for relevant worldly propositions; one would expect there to be some relevant difference between direct awareness of our own mental events and direct awareness of the world, other than simply their subject matters. If no such difference is suggested by a relevant notion of direct awareness, then this leaves us to wonder: why does this awareness play the relevant epistemic role? That is the nub of my argument for a unique ‘explanatory gap’ in the case of the awareness option for developing first-person warrant.

As suggested earlier, I’m inclined to think that any understanding of the relevant ‘higher-order’ or ‘meta-mental’ direct awareness seems to give rise to this explanatory problem unless we are working with some primitive notion of ‘awareness’ here that is taken to differ from first-order awareness such that only the former provides (non-inferential) normative epistemic support for propositions. Strictly speaking, however, all we need here is the claim that it seems any functionalist analysis of meta-mental (direct) awareness will be such that we can (coherently) ask: why should the relevant functional-role awareness property provide
(non-inferential) normative epistemic support, ‘evidence’, broadly construed, for the relevant present-tense mental propositions?

Overall, then, it seems quite plausible we have a legitimate unique explanatory gap on our hands in this case as well. If there is some obvious notion of higher-order awareness here that plausibly explains the relevant epistemic support, then I must confess that I am not familiar with it. On higher-order approaches to phenomenal consciousness, the main division has always been between ‘inner-sense’ (higher-order ‘perception’) and ‘cognitive’ (higher-order thought) approaches to the relevant awareness. On neither approach is it at all obvious that such awareness has the relevant epistemic credentials (normative epistemic and non-inferentially supporting)—especially when developed in such ways as to be ‘reductionist-friendly’.

So, I think only a bit of reflection is really needed here to tell that on any mainstream non-primitive approach to higher-order awareness, the relevant explanatory gap arises: why should this ‘awareness’ of a mental event provide normative non-inferential epistemic support for the occurrence of that mental event? ‘Inner awareness,’ after all, might be non-epistemic and merely a psychological enabling condition for introspective accessibility (in the sense of ‘introspective accessibility/access’ outlined in chapter 2).

We discussed inner attention in such a capacity in chapters 2 & 3. ‘Inner attention’, more specifically, was suggested as constituting having first-person warrant, and it is clearly intended to be a form of non-epistemic meta-mental/higher-order awareness. So, it should be plausible for at least a couple of reasons by this point that developing first-person warrant in terms of (direct) meta-mental awareness, with no obvious exceptions, leaves the relationship between that
metaphysical property (whether genuinely relational or not) and the relevant epistemic property *mysterious enough* that we are left to wonder: why does this metaphysical property confer the relevant epistemic support here?

**2.7: The problem of different kinds of explanatory gaps**

I have claimed that there may be different kinds or sorts of explanatory gaps: explanatory gaps between the structural/functional and the normative epistemic are not all explanatory gaps of the same sort, is what I committed to in the *previous section*. But this suggests that we have a certain understanding of the source of the relevant explanatory gaps: we know enough about the nature of the different phenomena to be explained in structural/functional terms to know the reason why they can’t be so explained differs in the one case from the other such that there are different kinds of explanatory gap relations, as opposed to different instances of the same sort of explanatory gap relation.

I think we need to walk a fine line here and embrace a view according to which we can have good reason for thinking that the *underlying sources* of certain explanatory gaps differ in kind from one another, without knowing much of anything about the sources themselves. If we know that the *explananda* belong to different metaphysical categories or families, then that is at least some support for the explanatory gaps being ‘different in kind’ or, more precisely, deriving from different kinds of sources. That, to be clear, is the intended interpretation behind my use of the phrase ‘different kinds of explanatory gaps’.

Now, I explained earlier on why I think it is important to understand phenomenal consciousness, or the property that constitutes/entails first-person warrant, as a property such that the explanatory gap it gives rise to has a unique source. Unless we assume that phenomenal consciousness is the only property that
gives rise to an explanatory gap between it and the structural/functional, then we require a unique source of the explanatory gap here to distinguish the relevant problem of phenomenal consciousness from others. Since, in my view, phenomenal consciousness is not plausibly alone in giving rise to an explanatory gap, we need to have some support for the view that phenomenal consciousness is associated with an explanatory gap that has a unique underlying source. I have argued that this is so when it comes to phenomenal consciousness itself, regardless of whether or not first-person warrant is developed along higher-order awareness or self-support lines.

2.8: But is this really an ‘explanatory gap’?

The issues surrounding ‘a priori’ or ‘conceptual’ entailment here are somewhat subtle and confusing, but we could also just be a bit more relaxed in our understanding of the ‘explanatory gap’ itself, since a recognizably important theoretical problem has plausibly been shown to emerge here, and it is plausibly a more specific version of one that has an established, if not always properly acknowledged, importance: reconciling a reductionist account of phenomenal consciousness with its epistemic credentials.

This issue usually arises in the context of ‘infallibilist or nearly so’ views of the epistemic status of the sort of epistemic support that phenomenal consciousness is taken to provide: how could any ‘physical’ or ‘physical-friendly’ property or properties be such as to confer this sort of epistemic support? In a causal/structural world, how do we ever get this sort of strong epistemic traction? I think it is fair to say that this has been a problem for traditional physicalist views according to which metaphysical reduction is always underwritten by a kind of explanatory reduction.

But the (main) difference between this traditional challenge and my own is that I have operated from the weaker premise that the epistemic credentials of
phenomenal consciousness are such that they are just *normative epistemic*, phenomenal consciousness is ‘evidential’ not just ‘reliable’ or ‘truth-conducive’ when it comes to its epistemic credentials—and that this normative epistemic support is *non-inferential*. If we follow this path, then I have tried to show here that on either major way of developing a normative epistemic account of phenomenal consciousness, namely, the awareness or the non-trivial self-support account of first-person warrant, we get a less committal and thereby improved version of this venerable challenge to physicalism. It is also one I think is well-considered as a unique ‘explanatory gap’—in the sense deriving from Chalmers (2010b;1996).

In any case, I will now move on to discuss what is certainly a more ‘traditional’ explanatory gap issue in the vicinity: the explanatory gap that arises with respect to certain *phenomenal properties*, that is, properties of phenomenally conscious events, such as those we discussed in the *previous chapter*.  

3.1: Phenomenal properties and the explanatory gap

Phenomenal properties are properties such as the painfulness or the negative affect of a (phenomenally conscious) pain, (phenomenally conscious) perceptual awareness and the properties of its ‘objects’. Even if most of these properties turn out to be properties of externalia, it is still reasonable to think of such ‘secondary qualities’ as colors, tastes, sounds, and so on as ‘phenomenal properties’. This, thanks to our introspective access to them by way of our introspective access to conscious perceptual awareness, even if the relation of perceptual awareness itself is not *primarily* introspectively accessible, a possibility we considered at some length in the *previous chapter, sections 2.2 & 2.4*. In other words, even if our primary or direct introspective access to conscious perceptual awareness is restricted to what we
are aware of, this is still a form of introspective access to perceptual awareness itself, just not direct or primary introspective access to perceptual awareness itself.

Now, it is usually assumed that despite certain (in some cases) obvious differences between phenomenal properties, they all give rise to a common explanatory gap phenomenon. In other words, it is assumed that, despite certain important differences between phenomenal properties, they all represent a single unified source-type of an explanatory gap: there isn’t a unique explanatory gap associated with colors that differs from pain affect and so on. This seems a rather dubious assumption, but we can avoid making it without threatening to void the issue of independent interest.

Even if there are different explanatory gaps associated with different phenomenal properties, so long as these explanatory gaps are plausibly to be distinguished from those that arguably obtain in domains, such as the normative and the evaluative, then the integrity of the explanatory gap problem for phenomenal properties is preserved. That said, I think it’s clear that traditionally many properties of phenomenally conscious events have been thought to be qualities we come to learn about through introspection. And when it comes to such ‘qualities’, structural/functional properties provide no immediate explanatory purchase here and seem to be clearly such that we can coherently suppose any structural or functional property be present while a relevant quality is absent.

The problem here is worse than just a lack of explanatory purchase on apparent/putative qualities in general from structural/functional properties, because, as I said, the existence of such qualities is often thought to have introspective backing, such that they cannot be written off in the way we might think to do when it comes to qualities in other domains such as goodness in the sense of ‘goodness’ that
isn’t just ‘goodness for...’. But it is, of course, quite controversial whether we have introspective access to the relevant phenomenal properties qua qualities.

Moreover, representationalist/intentionalist theories and others have made the important contribution of suggesting ‘psychological level’ explanations of such ‘phenomenal qualities’ that may be ‘reductionist-friendly’, such that once we understand phenomenal qualities in other psychological terms, we can then analyse these ‘other psychological terms’ in structural/functional terms. For example, it has often been suggested that ‘sensory colors’ such as ‘sensory redness’ are just a matter of a mental event e’s representing a certain content. Now, if at least some such ‘phenomenal qualities’ can be understood in terms of representing a certain content, and representing a content is broadly susceptible to an analysis in functional terms, then this suggests a way in which the explanatory gap for these phenomenal properties might be closed.

A proponent of the explanatory gap for the relevant phenomenal properties might respond to this sort of challenge in a number of ways, but two reactions seem particularly germane. First, one could argue that any functionalist construal of the relevant representation relation here will be such that it is coherent that the functional-representation of a certain content is present, and the relevant phenomenal property is absent. Alternatively, one could deny the initial analysis of ‘phenomenal qualities’ in terms of representing a content and suggest that irrespective of whether or not representing a content here is understood as being susceptible to a functionalist analysis, that it at least seems coherent that representing a content be present while a relevant phenomenal quality is absent. I will argue that we should favor the second option here.
3.2: Why analytic representationalism can’t succeed

Representationalist/intentionalist theories of phenomenal qualities were developed in the context of a metaphysics of phenomenal qualities according to which they are, or least seem to be on the basis of introspection, qualities of *mentalia*. This is why Harman’s (1990) ‘transparency argument’ was so startling and controversial in its time: it had largely passed unnoticed that a controversial metaphysics of secondary qualities had been at work in the background of much of the early discussion of whether or not phenomenal events could be reductively explained in terms of neural events.

In the *previous chapter*, I suggested that Harman’s (1990) transparency argument is liable to stronger and weaker interpretations. On the weak interpretation of Harman’s (1990) transparency argument, there is no direct introspective support for the suggestion that phenomenal qualities are qualities of *mentalia* as opposed to properties of *externalia* or properties that are represented in perceptual experiences as being properties of *externalia*. I find a thesis along these lines to be quite plausible. Now, Harman (1990) is plausibly also interpreted as arguing for the view that we actually have *direct introspective support for a contrary metaphysics of such qualities*, namely, that introspection/ ‘phenomenology’ supports the view that these qualities are actually qualities externalia, or, at least, are qualities represented as being such in perceptual experience.

Though I am less sanguine about this stronger transparency thesis (as I explain in 2.3 of chapter 3), the point I wish to draw attention to here is that this stronger transparency thesis does not support an analysis of phenomenal qualities in terms of representing a certain content. If anything, it supports the notion that these qualities are qualities that are *represented* as, or simply are, qualities of externalia.
So, I think it is a mistake, on a number of counts, to suppose that qualities such as colors as we know them through introspection are such that they are plausibly analyzed in terms of representing a content.

One gets the impression of a category mistake here: confusing the property of representing a content with a feature of that content itself. And it does not gain any support through direct introspection; again, if anything, all that is supported here is that these qualities are qualities of externalia, or we are perceptually aware of them (perhaps incorrectly) as being such. Thus, even if phenomenal perceptual awareness is analyzable in terms of representing a content, it does not follow that phenomenal qualities such as ‘redness’ can be successfully analyzed in terms of representing a certain content. Redness and other qualities are, at best, what is represented by representations, not representations themselves of certain contents.

The above error arose from the notion that, pre-theoretically, there are some psychological/mental qualities that require explaining here. But, if anything, there is just ‘pre-theoretical’ support for a contrary view: namely, that these qualities we come to learn about through introspection are actually just qualities of externalia. To be clear, I’m inclined to think that ‘pre-theoretically’ there is no support for either view: that such qualities are qualities of the mental, or that they are qualities of the non-mental. Introspection, as I suggested in chapter 3, is plausibly neutral on the matter, and while we might think that there is some ‘common-sense’ support for the view that colors and other secondary qualities are properties located in the ‘external’ world, this, to my mind anyway, is far from clear (though, of course, I grant that we generally operate as though this is true).

In any case, the point, again, is just that it seems mistaken for a number of reasons to attempt to analyze color qualities and relevant ‘secondary qualities’ in
terms of *representations of* representing a certain content. So, I think it’s safe to say that phenomenal qualities such as redness pose an explanatory gap, since wherever we end up locating them—in the mind or in the world, *prima facie* anyway, they are not structural or functional properties that can readily be explained by other structural/functional properties (more on this later).

I should also point out that on representationalist views according to which colors *qua* qualities do not figure in the content that is represented in such perceptual experiences, there also seems to be a clear explanatory gap. Waiving the issue of representing a content vs. the content represented, representing a content such as a certain dispositional property being instantiated, or a physical property being instantiated, opens up an explanatory gap at the level of content: once we take colors *qua* qualities out of the content of a representation, it seems quite coherent that representing a content in which the relevant physical or dispositional property figures will be such that it can be present, while ‘colors’, as ordinarily understood, are absent.

It is partly for this sort of reason, I think, that there has been renewed interest in understanding secondary qualities such as colors as representations of certain ‘primitive qualities’. But once we take this line, it is more plausible to suppose that these ‘primitive qualities’ themselves, not representing them in some appropriately ‘perceptual’ way, generates the relevant explanatory gap here. Either way, once this sort of concession is made so as to include colors as we understand/know them through introspection in the content of perceptual representations, then we plausibly get an explanatory gap here.

My considered view remains, however, that it is a mistake to try to analyze the relevant phenomena in terms of representing a content at all, even if that content
includes colors as we understand/know them through introspection. That is not to say that the best overall theory in this area can’t be that colors and other secondary qualities are representations of certain primitive qualities: my suggestion earlier that this may be a kind of ‘category mistake’ is merely to point out that, at the very least, this is not a theory that has a good claim to being a ‘conceptual analysis’ here. In other words, we don’t need the suggestion that such a theory is ruled out a priori (as would be the case if it really is a category error)—only that there is considerable pressure on views according to which colors and secondary qualities are conceptually analyzable in this way.

Ultimately, then, this is how I think a proponent of the explanatory gap for phenomenal qualities should react to the challenge from what we might call ‘analytic representationalism’. I noted earlier, however, that this is one of two ways in which we can resist this suggestion. Though I am not persuaded by the alternative, it is, nevertheless, instructive to consider it here.

3.3: Chalmers on analytic representationalism

Chalmers (2010a) addresses the possibility of analyzing phenomenal qualities in something other than structural-functional terms which, subsequently, might be further analyzed in such terms. He writes, “One might hold that there is some intermediate notion X, such that truths about X hold in virtue of structural/dynamical descriptions and such that truths about consciousness hold in virtue of X.”32 He rejects this general strategy for closing the explanatory gap on the grounds that either X itself is structurally/functionally analyzable, in which case the relevant phenomenal truth(s) cannot be analyzed in terms of X, or X is not structurally/functionally analyzable, in which case the explanatory gap is, at best,

32 Chalmers, 2010a, p.122.
just booted down to the level of \( X \) itself. He then goes on to suggest that, when it comes to the ‘notion of functional representation’, we can’t successfully analyze phenomenal consciousness in such terms. In other words, if we are working with a notion of representing a content that is taken to be functionally analyzable, then Chalmers’s (2010a) claim here is that this notion can’t succeed in analyzing phenomenal qualities.

The problem with this argumentative strategy, however, is that it is not obvious we can’t analyze phenomenal qualities in terms of representing a content, unless we take the line discussed in the previous section, according to which this dubiously locates the relevant phenomenal qualities at the level of representation, as opposed to the level of the content represented. Suppose, for example, that it is held that colors as we understand/know them through introspection, are representations of certain primitive qualities/contents in which such qualities figure.

Here I don’t find it particularly compelling that this sort of representation might (for all we know a priori) obtain without a relevant phenomenal quality obtaining, unless we alight on the representation vs. represented undermining consideration from the previous section. So, I don’t find Chalmers’s (2010a) general strategy here, on its own anyway, to be adequate: the problem is not functional representation per se, as an analysis of phenomenal qualities, the problem is with representation in general, as an analysis of phenomenal qualities.

Once we clearly separate out the problem of phenomenal consciousness itself from the problem(s) of phenomenal properties, then I think it becomes apparent that it won’t do to simply suggest that a functional understanding of representing a content can’t successfully analyze phenomenal qualities. While it’s true that a functional-representation of a content, in and of itself, needn’t be phenomenally
conscious, this is orthogonal to the issue of whether a functionalist understanding of representing a content can successfully analyze certain phenomenal properties.

With these points in place, let us now move on to consider the issue of phenomenal qualities from a slightly different angle: does our introspective access to these qualities provide a kind of direct argument against a broad spectrum of conceptual reductionist views according to which these qualities can be fully grasped or understood in third-person or publicly accessible terms?

### 3.4: Phenomenal qualities and the Nagelian argument against analytic functionalism

In Thomas Nagel’s (1974) celebrated paper, “What is it like to be a bat?” he distinguishes between facts that are only accessible to certain knowers/rational-epistemic agents, and those which, in principle, are accessible to any knower/rational-epistemic agent. Now, even if there are no such facts on a coarse-grained notion of ‘fact’, Nagel (1974) seems be getting at something important here: our introspective access to certain phenomenal qualities seems to be such that those without the relevant epistemic capacity are not in a position to know certain fine-grained facts, since they lack a certain way of thinking about colors that we and others who happen to be constituted like us in the relevant ways have or possess. And if this is so, then analytic functionalism and other views according to which we only have one substantive way of thinking about such properties that can, in principle, be had by any rational epistemic agent, are incorrect.

Now, as far as I can tell, analytic functionalism and other views according to which there is one fully adequate way of thinking about phenomenal properties that can, in principle, be had by any (fully) rational epistemic agent, only allow that if there are any other ways of thinking such properties, these ways are brute, along the
lines of a ‘bare demonstrative’ that contains absolutely no information about its referent—a kind of ‘blind pointer’. That is to say, if there is a way of thinking about such properties in addition to the (purported) one which is, in principle, accessible to any and all rational epistemic agents here, then it is only accessible to some rational epistemic agents, and this way of thinking must be ‘brute’ in the manner just described. If not, then this other way of thinking about the relevant properties threatens the status of the (purported) third-personally accessible and ‘canonical’ way of thinking these properties as conceptually reducing them or analyzing them, such that we can’t meaningfully (coherently) ask: why should the relevant third-personally accessible property always be such that when it is present, the relevant ‘phenomenal quality’ is also present as well.

It may be thought that even allowing for some other way of thinking about phenomenal properties that is akin to a ‘blind-pointer’, is incompatible with the conceptual reduction of the relevant phenomenal truths to the relevant ‘third-personally accessible’ truths. This may be correct, but it is obviously of no help to analytic functionalism or other conceptual reductionist views here. If anything, we have arguably just arrived at a useful starting point for thinking about metaphysical reductionist views that grant the existence of an explanatory gap between phenomenal qualities and the relevant reducing properties.

A defender of conceptual reductionist views of phenomenal qualities might respond to the above ‘Nagelian argument’ by denying that introspective knowledge of phenomenal qualities involves an introspective ‘mode of presentation’ or way of thinking about such qualities, and that it is just ‘knowledge’ that results from a certain distinctive kind of psychological process that tends to produce true beliefs concerning the relevant mental events. If, instead, it is allowed that we have a certain
distinctive source of non-inferential epistemic support for the relevant propositions, then this raises problem that such knowledge plausibly goes by way of epistemic access to either the property itself, or some aspect of that property which we don’t have epistemic access to except through introspection. If it is suggested at this point that this unique epistemic access is just unique epistemic access to a fine-grained proposition, then, once again, we have given up on conceptual reduction here.

So, I think the options for the conceptual reductionist here are limited: either a very bare or thin introspectively acquired notion of such qualities figures in our introspective access to them, or our ‘introspective’ access itself to these properties is very bare or thin, and runs along the lines of a purely evaluative truth-conduciveness account, that does not implicate anything like first-person warrant. As I briefly argued in chapter 2, however, it is something of a plausible default assumption in this area that there is a kind of distinctive normative epistemic warrant for the relevant propositions.

Concerning the alternative option: if it is assumed we have a certain introspectively acquired understanding of phenomenal qualities/properties such as colors, then the idea that this understanding is just ‘blindly referential’ or not revealing of any substantive information concerning its subject-matter, seems quite dubious. But introspective access to colors seems to furnish us with a substantive though difficult to express grip on them, and conceptual reductionist approaches to such secondary qualities struggle to accommodate this idea. Indeed, our ‘Nagelian argument’, strongly suggests that we can’t fully understand colors and other secondary qualities by means of some notion that any (fully) rational epistemic agent can, in principle, acquire. Our introspective access to phenomenal qualities seems substantive (why else would we be so resistant to third-person publicly accessible
analyses here\textsuperscript{33}, and this ‘datum’ places considerable pressure on conceptual reductionist views in this area.

We might think about it this way: conceptual reductionist views problematically entail that someone without introspective access to color experiences and colors could come to fully understand and know what colors are. But, as classic examples such as Frank Jackson’s (1982) ‘Mary’ suggest, it is very tempting to suppose that, at the very least, such individuals will be missing a substantive way of thinking about such qualities. In general, while the ‘knowledge argument’, or Jackson’s (1982) ‘Mary’ may not show that physicalism is false, it does plausibly reveal that we at least take our introspective access to colors to be substantive: substantive enough anyway that we wish to credit Mary with a kind of significant epistemic progress once she experiences colors for the first time and reflects on them.

Some (e.g., Dennett, 1991) notoriously deny that any such epistemic change takes place in Mary once she leaves her monochromatic environment, but this widely considered an implausible stance here. Moreover, part of the appeal of Dennett’s (1991) discussion of the Mary case is his attack on the highly idealized nature of the case itself as in one which a human individual like us knows all the narrowly and broadly physical facts in physical terms. But unless we assume there is more to the physical than what Chalmers and others have referred to as the ‘structural-dynamical’—what I have been calling the ‘structural/functional’, then it is unclear whether this idealization is problematic. In any case, we have seen that we needn’t rely on Mary-type cases for the Nagelian argument to have force, though such cases may help drive the point home.

\textsuperscript{33} Cf. Chalmers, 1996, p.15.
3.5: Is there introspective support for ‘phenomenal qualities’ *qua* qualities?

I have argued that we have introspective support for the existence (instantiation) of colors and other secondary qualities at the actual world. I have also argued that this introspective support is a distinctive sort of normative epistemic support. I don’t think this can be without us also having a kind of substantive epistemic access to these properties themselves such that we at least know that colors aren’t can’t plausibly be identified a priori with any *obvious* structural/functional property. Some think that our introspective support for the existence of colors is such that we also have introspective support for colors being *primitive qualities* or just (difficult to express) *qualities*. Others might think that, strictly speaking, we only have introspective support for the view that colors and other secondary qualities are *not* structural or functional properties.

But it is not particularly clear how we could have *direct* introspective support along these lines. Presumably whatever ‘introspective support’ there is for thinking that colors are not structural or functional properties derives from direct introspective support for either the proposition that colors are *qualities*, or for some difficult to express—perhaps even inexpressible, proposition that *entails* that phenomenal properties such as colors aren’t structural/functional properties. Since there seems to be some sort of ‘pre-theoretical’ support for the view that colors, for instance, are *qualities*, it is difficult to avoid the suggestion that the relevant source of support here is something along the lines of introspection.

In any case, the ‘intuitive’ backing behind the view that colors are qualities, whatever it is exactly, continues to exert a strong influence against both conceptual
and ‘metaphysical only’ reductionist views here. That said, the issue is complex and requires a more detailed discussion than I can offer here.

At the very least, it is hopefully clear by this point that our introspective access to phenomenal properties such as colors provides a steep challenge to conceptual reductionist views in the vicinity, and it is highly plausible that, contrary to these views, phenomenal properties such as colors open up an explanatory gap that cannot be closed except by approaches proposing the existence of certain exotica along the lines of ‘proto-phenomenal qualities’ or other unfamiliar qualities. Qualities seem fit to conceptually entail other qualities, even if the explanans in this sort of case is unfamiliar and postulated. Structural or functional properties, on the other hand, don’t seem fit to conceptually entail qualities, hence the need for ‘conceptual analysis’ of notions that seemingly refer to qualities in structural/functional terms, or terms that are ultimately analyzable in structural/functional terms.

This is familiar territory from the relevant consciousness studies/philosophy of mind literature. My main contribution has been to emphasize the importance of introspection in this context. Depending on how one works out the details of introspective access to phenomenal properties, irrespective of the well-known question of its epistemic force/strength, we get pushed or pulled in a certain direction when it comes to the issue of whether or not phenomenal properties such as colors generate an explanatory gap. And, as we saw in the previous section, we needn’t rely on well-known thought experiments and cases such as Jackson’s (1982) ‘Mary’ and Chalmers’ (1996) ‘zombies’ to motivate or reveal the existence of a unique explanatory gap when it comes to certain much discussed phenomenal properties. In the next section, we will turn to discuss the important issue of whether certain
phenomenal properties, or phenomenal properties in general, are *phenomenally proprietary* in the context of reducibility.

**4.1: Phenomenally proprietary properties and reduction**

Phenomenal properties such as colors which we seem pre-theoretically inclined to think of as *qualities* can pack an additional punch against reductionist views if they are also considered the preserve of the phenomenally conscious. This suggestion is liable to sound outrageous until it is recalled that it is by no means obvious whether colors as we know them through introspection are properties of *externalia*. That said, there does not seem to be any obvious problem with the view that phenomenally *unconscious* visual awareness involves visual awareness of colors as we know them through introspection. It may be tempting to suppose that any property we have introspective access to is such that it can’t be instantiated without at least implicating a phenomenally conscious mental event. Our direct epistemic access to colors always goes by way of phenomenal consciousness and introspection, but this in no way shows that colors aren’t instantiated without our having the relevant epistemic access to them.

While the idea that we can be unconsciously visually aware of colors as we understand/know them through introspection may seem a bit strange, it is certainly a live option. Perhaps one of the main reasons why this possibility has been unattractive to some is that it suggests colors themselves aren’t familiar structural/functional properties: colors as we introspectively understand/know them aren’t plausibly structural/functional properties, so if these ‘colors’ aren’t phenomenally proprietary, then it follows that colors don’t seem to be structural/functional properties—and this hurts those with reductive aspirations here. Worse still, if our introspective grasp of colors is a grasp of *qualities*, then if we can
be unconsciously visually aware of ‘colors’ in this sense, it follows that colors are qualities, not reductionist-friendly structural/functional properties (unless we assume an eliminativist view of color here).

Moreover, the problem is not just that the colors as qualities or apparent qualities view pose an explanatory gap here. It is also that if we suppose colors qua qualities/apparent qualities are properties externalia, then this seems to be tension with the prevailing scientific understanding of colors according to which there aren’t any such qualities of externalia. This was presumably the initial motivation for locating any such qualities ‘in our heads’ in the first place. Our scientific picture of distal objects would be greatly complicated, if not undermined, by such qualities being properties of externalia, unless we assume such qualities are entirely epiphenomenal, which, aside from arguably not dealing with the problem here, would make our visual awareness of such properties a primitive affair: visual awareness of such properties (property-instances) would then be an entirely non-causal matter.

At the very least, then, it is clear why theorists of all stripes—reductionist and non-reductionist alike, have wished to avoid the view that colors are qualities of externalia: it seems in tension with a naturalistic/scientifically informed view of colors. For this reason, it should come as little surprise that theorists have had reservations about holding that colors as we introspectively understand/know them are properties we can be unconsciously visually aware of: this seems to lead us dangerously close to the view that colors are qualities of externalia, or just properties of externalia that don’t sit well with a naturalistic but not necessarily reductionist approach in this area.
We could, of course, locate colors as we understand/know them through introspection in our heads, and maintain that such properties can be instantiated in our heads without being phenomenally conscious. This ‘old-fashioned’ metaphysical view is not a popular option these days, though it has been advocated and defended by David Rosenthal (2005). While I don’t find there to be anything obviously wrong with such a view, it must contend with major theoretical rivals in the area. Moreover, if I am right in thinking that at least a weak version of Harmanian transparency is quite plausible, then there is no introspective support for the view that colors and other ‘secondary qualities’ are instantiated ‘in our heads’.

Finally, it has been argued that bodily sensations such as pain pose a problem here: it is implausible that the negative affect or painfulness of pains is a property that might be instantiated in one’s head without being phenomenally conscious. While we could treat painfulness as an aberration of sorts here, this seems ad hoc in the context of defending the view that phenomenal properties such as color (as understood/known through introspection) are not phenomenally proprietary. So, let us now turn to examine this ‘hard case’ since it is also of independent interest to our discussion here.

4.2: Is painfulness phenomenally proprietary?

It has often been suggested that the very idea of phenomenally unconscious pain is oxymoronic or contradictory: there is a conceptual entailment from being a pain to being phenomenally conscious. Now, if there is such a conceptual entailment from being a pain to being phenomenally conscious, then I think that this is only (at least initially) plausible on the assumption that displeasure is (at least) partly definitional of pain such that it is a conceptual truth that there are no pains lacking
the property of being displeasurable. I argued for a view of pain along these lines in
the previous chapter.

Even if this claim is correct, however, it’s not entirely clear whether
displeasure is such that *phenomenally unconscious displeasure* is
incoherent/contradictory. It does seem strange to suggest that something can be
displeasurable to someone without their being aware of it, but, if anything, this may
be just a mildly intriguing empirical fact: as a matter of fact, something can’t be a
source or cause of the feeling of displeasure in someone unless they are aware of
said source or cause. In other words, I must, as a matter of fact, be aware of
something in order for it to cause the feeling of displeasure (to more or less of an
extent) in me.

This is surely correct for the most part, but might there not, in fact, be cases
in which I am not aware of anything at all, but the feeling of displeasure is tokened
in me? Couldn’t my brain be ‘zapped’ in such a way that while I am not aware of
anything in particular going on, the feeling of displeasure gets tokened in me? There
is also the problem of hallucinations here, though this potential empirical
counterexample is obviously difficult to evaluate. Overall, then, I’m inclined to think
that there is no entailment from pain *qua* type of event, to phenomenal
consciousness.

4.3: No good case for secondary qualities being phenomenally
proprietary?

If the above verdict can be upheld, then I think this is a substantive set-back
for to the view that certain phenomenal properties are the preserve of phenomenally
conscious events. Pain, after all, is among one of the most promising candidates for a
mental event-type whose tokens are always phenomenally conscious. And we have
seen that even as a matter of contingent fact, it is unclear whether pains are always phenomenally conscious—so they are definitely not the preserve of the phenomenally conscious, since being phenomenally proprietary, at the very least, requires that a property be such that, at the actual world, it can only be instantiated by a phenomenal event.

So, if pain is taken to be a ‘hard case’ from the category of ‘secondary qualities’ such as color, then this may suggest that it is a mistake to suppose that ‘secondary qualities’, in general, are phenomenally proprietary. It may be replied, however, that all this would show is that we can’t use cases like pain to argue for the view that secondary qualities are phenomenally proprietary. That seems correct, but then it is not particularly obvious where additional support for the view that secondary qualities are phenomenally proprietary comes from (assuming there is such support).

An important suggestion here is that additional support comes from the theoretical/explanatory payoffs of the theory itself. If we accept that the relevant phenomenal properties are phenomenally proprietary, then, inter alia, this explains why introspective access to the relevant properties seems uniquely informative: we are learning about properties that can’t be instantiated by the non-phenomenal world, and our introspective access to those properties which plausibly isn’t uniquely informative, such as our introspective access to ‘primary qualities’ such as size and shape, is, by contrast, to be explained by those properties not being phenomenally proprietary. In a way, this is perhaps what we would expect: if certain properties can be instantiated by the non-phenomenal world, then our introspective access to such properties isn’t uniquely informative about them, nor does it seem to be.
While this is an interesting explanation as to why our introspective access to certain phenomenal properties seems uniquely informative about them, it is by no means the only reasonable explanation in the vicinity: the explanatory gap associated with secondary qualities might also explain why our introspective access to these properties seems uniquely informative while our introspective access to primary qualities does not. Moreover, if secondary qualities are, indeed, *qualities*, then that would also explain why our introspective access to these properties seems uniquely informative.

Overall, the fact that there are other plausible explanations in the vicinity which don’t entail that secondary qualities are phenomenally proprietary, and the fact that there plausibly are other theoretical/explanatory considerations that favor the view that such properties are not phenomenally proprietary, suggests that there is no obvious additional support for the view that secondary qualities are the preserve of the phenomenally conscious. Instead, its support seems to accrue solely from the considerations canvassed in *section 3.6 of this chapter*. But here we saw that these considerations don’t militate against the view endorsed by Rosenthal (2005) according to which secondary qualities are mental or ‘inner’ properties that can be instantiated by phenomenally unconscious events.

If there is some independent reason for rejecting Rosenthal’s (2005) view that isn’t also a reason for rejecting the view that secondary qualities are phenomenally proprietary, then this would obviously help the case for the view that secondary qualities are the preserve of the phenomenally conscious. It is rather unclear, however, what sort of reason might be proffered in this regard. It also seems unlikely that we will be able to alight on some sort of decisive consideration here.
any case, I think it may be rather helpful to revisit the broader context within which this debate takes place.

4.4: No overlap, partial overlap, and full overlap views phenomenal properties

On what I will call a ‘no overlap’ view of phenomenal properties, the phenomenal properties of phenomenally conscious events are all only had by those events. On a ‘partial overlap’ view of phenomenal properties, at least some, but not all, phenomenal properties of phenomenally conscious events can be had by events that aren’t phenomenally conscious and don’t implicate phenomenal consciousness. On a ‘full overlap’ view of phenomenal properties, all phenomenal properties can be had by those that aren’t this way.

If only secondary qualities such as colors are such that they can’t be had by any phenomenally unconscious events and never, at least as a matter of fact, implicate them, then a partial overlap view is correct here. It may be thought that the most likely version of a partial overlap view to be correct here is that most phenomenal properties can, in fact, be had or implicate non-phenomenal events, with some exceptions such as, perhaps, secondary qualities. As I briefly explained in chapter 1, however, this sort of partial overlap thesis requires some account of why there are these ‘exceptional’ phenomenal properties that are phenomenally proprietary.

If we must choose between no overlap and full overlap views, full overlap views seem to have the upper hand. No overlap views hold that all properties we come to understand/know through our introspective access to perceptual awareness are such that they are phenomenally proprietary, but then what about ‘primary qualities’ we encounter through phenomenal perceptual awareness and thus have
introspective access to? Such structural/functional properties obviously can’t be phenomenally proprietary, but then how do we, in a non-\textit{ad hoc} way, find a way to disarm their status as phenomenal properties. We can’t simply say that secondary qualities such as colors are \textit{mental} properties here, since ‘mental’ is no longer a unified kind on the no-overlap view. Moreover, we wouldn’t want to try this sort of move anyway, because we want to leave open the possibility that secondary qualities are qualities of \textit{externalia}, so it won’t help to switch from ‘mental’ to ‘inner’ in this context.

In the \textbf{previous section}, we discussed a partial overlap view that is rather in the spirit of no overlap views, according to which only those phenomenal properties that introspection seems to be uniquely informative about are phenomenally proprietary. But we came up empty handed in our search for a persuasive direct argument for this claim. In any case, we are now in a position to say that the choice here is plausibly just between full overlap and partial overlap views; this is an important area in which more substantive theorizing is needed. Moreover, even if those properties to which we seem to have uniquely informative introspective access aren’t phenomenally proprietary, this idea that some phenomenal properties are to be distinguished from others in that we seem to have the relevant sort of introspective access to them, is a good one.

Recent work by Philip Goff (2011; 2017) has reinforced the importance of the above sort of claim about our introspective access to the relevant phenomenal properties in arguing against both conceptual and metaphysical-only reductionist approaches to these properties. As Goff (2011; 2017) plausibly argues, both conceptual and metaphysical-only reductionist approaches must ‘tame’ our introspective access to the relevant properties by denying that such epistemic access
is uniquely informative in some way. Instead, such epistemic access must be treated as providing no such unique information about these properties: it is either a kind of ‘blind knowledge’ as on purely truth-conducive views of epistemic support, or it is knowledge that employs ‘blind concepts’ that don’t reveal any unique information about their referents—that is, information about their referents which isn’t already captured by concepts that can, in principle, be had by any rational epistemic agent.

This is my own gloss on some of the details of Goff’s (2011; 2017) argument, but the basic idea is that when it comes to secondary qualities, reductionists of all stripes plausibly can’t allow that we have a kind of introspective access to these properties which is uniquely informative about them. But if introspective access involves a unique form of non-inferential normative epistemic support, and we seem to have introspective support for properties beyond just those that are just structural/functional, then it should come as no surprise that this sort of introspective thesis is a threat to reductionist theses in general, and must be denied by their exponents. Whether reductionists can plausibly deny such a thesis remains a vexed issue, and we shall not attempt to resolve it here. Again, I just want to emphasize the importance of introspection in this context—this is plausibly the major source of the problem here, and I think we would do well to keep this in mind for future work in this area.

5.1: The two main explanatory gaps in consciousness studies

In this chapter I have argued that there are two main explanatory gaps when it comes to phenomenal consciousness: the explanatory gap that derives from phenomenal consciousness itself and the explanatory gap that derives from certain phenomenal properties, plausibly, those we seem to have a unique insight into by way of our introspective access to them. In the case of the latter, this unique insight
is usually expressed in terms of these phenomenal properties being such that they aren’t, *prima facie* anyway, structural or functional properties: they are not, on the surface at least, *relational* properties of any sort at all.

Still, as I argued in the **previous section**, we don’t seem to need the claim that phenomenal properties such as secondary qualities don’t seem to be structural/functional properties to generate a (unique) explanatory gap here. Instead, our introspective access to these properties themselves generates a problem for *conceptual reductionist* approaches to secondary qualities, since conceptual reductionist approaches hold that, in principle, any (fully) rational epistemic agent can fully understand colors and other secondary qualities. But this seems incorrect: our introspective access to colors and other secondary qualities seem to furnish us with some unique information about them.

At the very least, as I argued earlier, it is tempting to credit Jackson’s (1982) Mary with some substantive epistemic progress when it comes to colors after she consciously sees them for the first time and reflects on them, epistemic progress that plausibly at least puts her in a position to know a certain *fine-grained* fact about colors that she didn’t know in her monochromatic environment. But if that’s so, then there is an explanatory gap with respect to secondary qualities, and conceptual reductionist approaches to these properties fail without our having to rely on a kind of ‘brute’ intuition to effect that for any structural/functional property we can currently think of, it is coherent that this structural/functional property may be present while a relevant secondary quality is absent.

So, I think we have made some philosophical progress in both clarifying solidifying the status of the explanatory gap such phenomenal properties pose. Additionally, we have reconceived of the explanatory gap in the case of phenomenal
consciousness itself in way that will hopefully attract renewed interest and attention. May future work in this area pick up where I am leaving off!
In this thesis, we explored a number of important topic areas in the consciousness studies literature. In chapter 1, I developed a framework for thinking about the major problems of consciousness, and its centerpiece was the understanding of phenomenal consciousness itself in terms of ‘conditional introspective accessibility’. This way of understanding phenomenal consciousness is, as far as I am aware, a unique contribution to the literature that avoids the problems associated with characterizations of phenomenal consciousness in terms of introspection alone, such as the fact that those without the relevant psychological enabling conditions needed for introspection may yet have mental events that are phenomenally conscious.

Our understanding of phenomenal consciousness from chapter 1 also fostered a principled distinction between phenomenal consciousness and phenomenal properties that has often been elided or conflated in the literature. Moreover, I argued that the ‘kind view’ of phenomenal consciousness, which would offer support for thinking about phenomenal consciousness itself in terms of phenomenal properties, once made explicit, is to be rejected for a number of reasons—not the least of which is that it immediately rules out certain theoretical approaches to phenomenal consciousness, such functionalist and higher-order approaches.

Chapter 1 also addressed an important quasi-empirical challenge to the conditional introspective access understanding of phenomenal consciousness from Ned Block (1995; 2007), offering a (as far as I am aware) novel line of response to this challenge that should be of independent interest. We ended chapter 1 with a ‘logical geography’ of the major issues in consciousness studies: phenomenal
consciousness itself, phenomenal character/phenomenal properties, and introspection. The rest of the thesis explored these issues, with an emphasis on the connection between introspection and certain phenomenal properties, and introspection and phenomenal consciousness itself.

In **chapter 2**, I developed an epistemological framework for introspection, thereby addressing what I argued in **chapter 1** is one of three major problems in the consciousness studies literature. The framework developed in **chapter 2** filled in a number of specific details absent from earlier accounts developed along similar lines, such as that found in Chalmers (1996; 2010c). The major assumption in **chapter 2** is that we have a kind of unique normative epistemic support for propositions along the lines of ‘*this* mental event is occurring right now’.

While I did attempt to motivate this assumption, some will no doubt be inclined to think that there is a plausible account of introspection according to which there is no such epistemic support. I disagree, but we can take what follows from this claim as *conditional*—if there is such unique normative epistemic support for the relevant propositions, then such and such follows. This would still be of considerable interest, I think, since the relevant epistemological assumption here still has a large number of supporters, and is at the core of major anti-reductionist works such as Chalmers’ (1996) *The Conscious Mind*.

In **chapter 3**, we explored the issue of phenomenal character/phenomenal properties from a number of different angles, with a focus on colors and pain. I argued that Harman’s (1990) well-known ‘transparency argument’, namely, the argument that we have introspective support for colors being (solely) instantiated by *externalia* should not enjoy a kind of default status in this area, as is often supposed nowadays. Instead, we should be careful to distinguish between ‘weak’ and ‘strong’
Harmanian transparency—I argued that the weak version is preferable and that, if anything, the view that introspection is neutral or silent on the issue of whether colors (as we know them through introspection) are mental or not should be the default in this area.

We also explored the related issue of whether or not we have direct or primary introspective access to conscious perceptual awareness. It has often been suggested that we only have direct introspective access to the object/content of perceptual awareness, but the issue is highly fraught, and there are a number of potential counterexamples to such a claim. We considered a number of such counterexamples and concluded that while none are obviously decisive, there is a need for future work to further engage certain ones that have yet to receive the attention they deserve, such as the (alleged) property I called ‘innerness’ in section 2.5.

After leaving the topic of perceptual awareness, we then turned our attention to pain—a classic ‘hard case’ from the family of phenomenal properties. We examined pain in some depth and ended up with a novel account of it that leaves open the issue of whether it is reducible or not. This part of chapter 3 is a self-contained discussion of a venerable topic that should be of interest to theorists of all stripes. We covered certain interesting empirical data that may have a bearing on the metaphysics of pain, and found that some of the relevant empirical research in this area, the research on ‘pain asymbolia’, suffers from a study design problem that can hopefully be addressed in future work.

We concluded chapter 3 with a brief discussion of the issue of the reducibility of phenomenal properties such as color that foreshadowed chapter 4 in which we turned to examine the issue of reduction in more depth—for both
phenomenal consciousness itself and phenomenal properties such as colors as we understand/know them through introspection.

**Chapter 4**, the final one of this thesis, examined the explanatory gap in the context of our unique understanding of phenomenal consciousness itself and in the more familiar setting of phenomenal properties such as colors. I argued that phenomenal consciousness itself and certain phenomenal properties such as colors each pose a unique explanatory gap. This differs from the traditional view according to which there is just one explanatory gap in the vicinity, but this arguably just reveals how the distinction between phenomenal consciousness itself and certain phenomenal properties such as pain and colors, has often been conflated.

Having made the case for these importantly distinct explanatory gaps (explanatory gaps deriving from distinct sources), we then proceeded to examine the potential for a somewhat different way of arguing against conceptual reductionist approaches to phenomenal properties such as color. I argued that there is a Nagel (1974) inspired argument from our introspective grasp of such phenomenal properties that plausibly refutes conceptual reductionist approaches while bypassing the potential for a deadlock of intuitions here—what Chalmers (1996) once called the ‘great divide’ in consciousness studies. This highlights the importance of considering introspection and the nature of our introspective access to certain properties in the context of the debate over physicalism, a point Philip Goff (2011; 2017) has also emphasized.

We also discussed the important issue of whether or not certain phenomenal properties are phenomenally proprietary, that is, the preserve of phenomenally conscious events. I defended the view that colors and pains are not phenomenally
proprietary, but the issue is vexed and other phenomenal properties such as
intentionality deserve attention in this regard as well—a task for future work.

In general, while the conclusions arrived at in chapter 4 are fairly modest,
the ways we arrived at them are of some interest since, among other things, they
broaden and deepen our understanding of what is at stake here. Ultimately, it is
unclear whether our development of the ‘explanatory gap’ in the case of phenomenal
consciousness itself helps or hurts the reductionist’s cause: it may be thought that
this new understanding of the problem is less daunting to physicalism, and no longer
suggests the sort of popular anti-reductive argument from conceivability to
metaphysical possibility in this case.

I think this is probably correct, especially if we are operating with the higher-
order awareness way of developing first-person warrant, but it is not clear to me
whether this ultimately makes the reductionist’s task all that much easier. In the case
of the higher-order awareness development of first-person warrant, we have a more
general explanatory challenge that doesn’t rely on conceivability intuitions; this may
make it more difficult for metaphysical-only reductionists to avoid the charge of
offering an ad hoc account of the phenomenon in question.

In any case we are still left with the more traditional explanatory gap
phenomenon in the case of phenomenal properties such as colors. If our novel
understanding of the ‘explanatory gap’ in the case of phenomenal consciousness
itself can be bridged, then the division of conceptual labor suggested in this thesis no
doubt helps to focus the issue squarely on the explanatory gap associated with
certain phenomenal properties. I’m inclined to think, then, that regardless of how we
ultimately interpret the novel explanatory challenge from phenomenal consciousness
itself developed in this thesis, it helps to clarify and advance the debate in this area.
References


Hellie, B. 2007. “‘There’s something it’s like’ and the structure of consciousness”, *Philosophical Review*, 116:441-463


