



# Phenomenal Concepts, Transparency, and the Hard Problems of Consciousness

## Permanent link

<http://nrs.harvard.edu/urn-3:HUL.InstRepos:40046510>

## Terms of Use

This article was downloaded from Harvard University's DASH repository, and is made available under the terms and conditions applicable to Other Posted Material, as set forth at <http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA>

## Share Your Story

The Harvard community has made this article openly available.  
Please share how this access benefits you. [Submit a story](#).

[Accessibility](#)

*Phenomenal Concepts, Transparency, and the Hard Problems of Consciousness*

A dissertation presented by

by

Shantia Rahimian

to

The Department of Philosophy

in partial fulfillments of the requirements

for the degree of

Doctor of Philosophy

in the subject of

Philosophy

Harvard University

Cambridge, Massachusetts

May 2017

© 2017 Shantia Rahimian

All rights reserved.

Phenomenal Concepts, Transparency, and the Hard Problems of Consciousness

**Abstract**

The hard problem of consciousness is the problem of explaining why experiences—perceptions, sensations, emotions, and moods—are brain states. My dissertation is motivated by the thought that our current understanding of the hard problem is flawed, and one of the aims of my dissertation is to address this flaw. The other aim is to lay a foundation for a monistic view of the mind-brain relation that is distinct from physicalism.

Most philosophers assume that there is only one hard problem since the hard problem is the problem of explaining why there is something it is like to be in a brain state. In “Puzzles of Pain”, I argue that some experiences, e.g., pains, are such that, when we have them, we feel a certain way, while other experiences, e.g., color experiences, are such that, when we have them, something seems a certain way. This suggests that there are two hard problems, one associated with each type of experience. These problems are the subjects of my other papers.

In “Knowing What Pain Is”, I argue that the hard problem for pain is reconciling our awareness and knowledge of pain with pain being a brain state. I argue that a reconciliation can’t be achieved since what we learn about pain via our awareness is that it is *sui generis*. In “Does Transparency Solve the Mind-Body Problem?”, I argue that the hard problem for color experience is explaining why a brain state does a certain thing, viz., make us aware of color. By examining how neuroscientists explain why brain states perform functions, I argue that the

hard problem for color experience can't be solved since a physical mechanism showing how a brain state makes us aware of color can't be specified.

Implicit in these negative conclusions is a positive view that combines the monism of physicalism with the phenomenality of dualism. The roots for such a view can be found in Russell and Eddington, and it holds promise as a worthy alternative to the usual options of physicalism and dualism.

## Table of Contents

### Chapter 1: Aims, Motivations, and Conclusions

1. The Aims of the Papers . . . 2
2. Motivating the Papers . . . 3
3. “Puzzles of Pain” . . . 4
4. “Knowing What Pain Is” . . . 5
5. “Does Transparency Solve the Mind-Body Problem?” . . . 6
6. Understanding the Mind-Brain Relation . . . 8

### Chapter 2: Puzzles of Pain

1. Problems with the Experiential View of Pain . . . 16
2. The Argument for the Experiential View of Pain . . . 17
  - 2.1. Pain is a Sensation (or Feeling) of Some Kind . . . 18
  - 2.2. Sensations (or Feelings) are Experiences . . . 20
3. Transparency, Pain, and the Experience of Color . . . 23
4. The Transparent Framework . . . 26
5. Resolving the Problems . . . 28
6. Two Hard Problems . . . 37

### Chapter 3: Knowing What Pain Is

1. The Hard Problem of Consciousness and the Phenomenal Concept Strategy . . . 40
2. Preliminaries . . . 43
  - 2.1. Modes of Presentation and ‘Water is H<sub>2</sub>O’ . . . 43
  - 2.2. ‘Pain is B’ . . . 46
    - 2.2.1. Pain, the Phenomenal Concept ‘Pain’, and the Constitutional Theory . . . 47
  - 2.3. Back to ‘Pain is B’ . . . 52
3. The Hard Problem and the Phenomenal Concept Strategy . . . 53
4. The Problem with the PCS . . . 58
  - 4.1. The Test Case: Eden, Perfect Red, and Perfect Pain . . . 59
  - 4.2. Having Pain on Earth . . . 67
  - 4.3. Reducing Red and Reducing Pain . . . 70
5. Conclusion . . . 77

### Chapter 4: Does Transparency Solve the Mind-Body Problem?

1. Transparency-Intentionalism . . . 81
2. The Transparency of Experiences of Blue . . . 85
3. The Identity Conditions for Experiences of Blue . . . 92
4. The Transparency-Intentionalist’s Dilemma . . . 105
  - 4.1. Satisfying (Easy) . . . 105
  - 4.2. Satisfying (Property Identity) . . . 107
  - 4.3. Solving Neural Functional Problems . . . 111
  - 4.4. No Physical Mechanism . . . 118
5. Conclusion . . . 120

## **Acknowledgements**

First, I would like to thank my dissertation advisor, Professor Edward J. (Ned) Hall, for his continuous support and guidance throughout my dissertation. Ned is brilliant, kind, and thoughtful, and I couldn't have asked for anything more in a teacher and advisor. This dissertation could not have been written without his help, and one of things I will miss most about graduate school is talking philosophy with Ned.

Second, I would like to thank my other committee members—Professors Susanna C. Siegel and Mark Richard—for their invaluable help and guidance. Susanna and Mark were uncommonly generous with their time and feedback, and I benefited greatly—both in writing this dissertation and throughout graduate school—from their expertise and knowledge.

Finally, and most important, I would like to thank my father, Shahram Rahimian, my mother, Zoey Hashemi, my stepfather, Sohail Davoudian, my wife, Silena Paik, and the rest of my family. My parents have given me unconditional love and support throughout my life; and words cannot express how patient, loving, and supportive my wife has been. I am very lucky to have all of these people in my life, and their encouragement and backing have made this dissertation possible.

## **Dedication**

To my grandfather, Mahmoud Hashemipour, the person I look up to most



## Chapter 1

## Motivations, Aims, and Conclusions

### 1. The Aims of the Papers

The **hard problem of consciousness** is the problem of explaining and/or justifying the claim that experiences—perceptions, sensations, emotions, and moods—are brain states (or brain states are experiences). The three papers making up my dissertation—“Puzzles of Pain, “Knowing What Pain Is”, and “Does Transparency Solve the Mind-Body Problem?”—concern the hard problem. My dissertation is motivated by the view that our current understanding and characterization of the hard problem is flawed. One of the aims of my dissertation is to correct this flaw by accurately characterizing the hard problem. Specifically, I argue that, contrary to widespread assumption, there are *two* hard problems: an **opaque problem** associated with **opaque experience**, and a **transparent problem** associated with **transparent experience**.

A second aim of my dissertation is to argue that, once the opaque and transparent problems are distinguished and correctly characterized, we can see that **neither problem can be solved**. The opaque problem resists solution because our awareness<sup>1</sup> of feelings yields knowledge of them that is inconsistent with their being brain states. The transparent problem resists solution because we cannot specify a physical mechanism that shows how a brain state makes a subject aware of a *way* a thing seems.

A third aim of my dissertation is apparent only when all three papers are taken together. On one hand, you might think that the papers show that mind-brain identities are false, i.e.,

---

<sup>1</sup> The notion of awareness—in particular, the notion of a special kind of awareness important to consciousness—plays an important role in all three papers. In this introduction, I use the terms ‘aware’ and ‘awareness’ in an intuitive and informal manner. I give a precise specification of the sense of ‘aware’ relevant to my arguments in the papers.

that **dualism** is true. On the other hand, it could be argued that what the papers *really* show is that a certain *interpretation* of mind-brain identities—**physicalism**—is false. Implicit in the papers is a *monistic yet non-physicalistic* interpretation of mind-brain identities. This *phenomenal* form of monism—**phenomenal monism**—could preserve the good features of physicalism and dualism while shedding their bad features. My dissertation, then, could be seen as laying the beginning foundations for a view of the mind-brain relation that is different from—and potentially superior to—the traditional options of physicalism and dualism.

## 2. Motivating the Papers

The mind-body problem is a perennial philosophical problem. In general terms, the mind-body problem is the problem of understanding how mind and body relate together. Prior to the 20<sup>th</sup> century, dualism—the view that mind and body are distinct—was so deeply ingrained in the minds of most philosophers that the primary form the mind-body problem took was: the problem of understanding how mind and body relate to each other *causally*.<sup>2</sup> In the mid-20<sup>th</sup> century and continuing to the current age, the rise of science and physicalistic metaphysics<sup>3</sup> has caused the mind-body problem to take a different form: the problem of understanding how mind and body—more specifically, mind and brain—are one and the same. A related version of the problem focuses on properties (or states): the problem of understanding how mental properties (states) and brain properties (states) are one and the same.

---

<sup>2</sup> See, e.g., the correspondence between Princess Elisabeth and Descartes (Shapiro 2014)

<sup>3</sup> See Papineau (2002, chapter 1).

This current form of the mind-body problem is thought to be especially difficult for *phenomenally conscious* mental states or **experiences**—things like perceptions, (bodily) sensations, emotions, and moods. In fact, the mind-body problem for experiences has come to be called the **hard problem of consciousness**.<sup>4</sup> Specifically, the hard problem is the problem of explaining and/or justifying the claim that experiences are brain states (or brain states are experiences). And this problem is thought to be hard (i.e., insoluble) because (i) there is something it is like to have an experience, and (ii) nothing could explain why—or justify the claim that—there is something it is like to be in a brain state.

Clearly, before we deem the current form of the mind-body problem hard, it is important that we correctly characterize the problem. After all, some phenomena—e.g., the relation of mind and body—could strike us as puzzling, but we could mischaracterize the *actual* problem (or puzzle) the phenomena presents us with. That is where my dissertation starts. In my view, our current understanding and characterization of the hard problem suffers from an important flaw: it assumes that experiences are uniform (in a specific way) with respect to what they make us aware of. I think that this isn't the case, and the consequence of this heterogeneity in experience is that we have, not one, but *two* hard problems of consciousness.

### 3. "Puzzles of Pain"

Pain has seemed a puzzling phenomenon to some philosophers because its mentality—specifically, the fact that it seems to be an experience—seems at odds with other of its features. More particularly, the experiential view of pain—i.e., the view that pain is an experience—is thought to have four problems: the semantic problem, the content problem, the

---

<sup>4</sup> Chalmers (1995).

awareness problem, and the location problem. In “Puzzles of Pain”, I argue that all of these supposed problems with the experiential view can be solved by recognizing that experiences come in two types. Some experiences, e.g., pain, are such that, when we have them, we *feel* a certain way, where this just means that we are aware of the *feeling* that is the experience we have. Other experiences, e.g., experiences of color, are such that, when we have them, *something seems* a certain way, where this just means that we are aware (because we have the experience) of a *way* a thing seems.

We can call the first class of experiences ‘**opaque experiences**’ and the second class ‘**transparent experiences**’. The fact that we have opaque and transparent experiences—coupled to the fact that the linguistic framework apt to talk about transparent experiences (the transparent framework) dominates the linguistic framework apt to talk about opaque experiences (the opaque framework)—resolves the four supposed problems of the experiential view. Just as important, the above fact about the heterogeneity of experience implies that there is no single hard problem of consciousness. The problem that our awareness of opaque experiences presents us with (the **opaque problem**) is different than the problem that our awareness of ways things seem presents us with (the **transparent problem**). These problems are the topics of my next two papers.

#### 4. “Knowing What Pain Is”

The split between opaque and transparent experiences is reflected in the fact that two different and irreconcilable approaches to the hard problem have been developed by physicalists. Physicalists who think that all experiences are opaque—i.e., all experiences are like pain—defend the view that, e.g., pain is a brain state by using the **phenomenal concept**

**strategy** (the **PCS**). According to the PCS, the hard problem is a result of the fact that the phenomenal description of pain is a non-contingent description of pain. This fact about the phenomenal description is due to the **constitutional theory of phenomenal concepts**. But the fact that pain is a constituent of the concept ‘pain’—and we, therefore, are aware of pain when we have it—does not imply—or even suggest—that pain is not a brain state.

In “Knowing What Pain Is,” I argue against this claim. I argue that our awareness of pain entails that our epistemic situation with respect to pain is indistinguishable from the epistemic situation of a subject S on Eden<sup>5</sup> having **perfect pain**. Perfect pain is a primitive, *sui generis* item having a distinct qualitative nature, and S can know—via their awareness of perfect pain—that perfect pain meets this description. They can know, then, that perfect pain is not a brain state. Because I am aware of pain when I have it, my epistemic situation is indistinguishable from that of S’s. Thus, I too can know that pain is a primitive, *sui generis* item having a distinct qualitative nature. I can know that pain is not a brain state.

##### 5. “Does Transparency Solve the Mind-Body Problem?”

The second approach to the hard problem comes from physicalists who think that all experiences are transparent, i.e., all experiences are like the experience of color. These physicalists (I call them ‘**transparency-intentionalists**’) claim that there is no hard problem, i.e., there is no insoluble mind-body problem associated with explaining why a brain state is, e.g., an experience of blue. Transparency-intentionalists make this claim because they think (i) the

---

<sup>5</sup> Eden (which has nothing to do with the biblical Eden) is a possible world described by Chalmers (2006) where everything is as it seems and how things seem matches how things seem on Earth. I discuss Eden in more detail in the paper.

mind-body problem is the problem of naturalizing intentionality<sup>6</sup>, and (ii) the problem of naturalizing intentionality is easy (i.e., it is difficult but solvable).

In “Does Transparency Solve the Mind-Body Problem?”, I argue that transparency-intentionalists face a dilemma in satisfying (i) and (ii). Because the identity conditions of, e.g., experiences of blue are captured by the functional description ‘makes S aware of perfect blue’<sup>7</sup>, and because the meaning of ‘represents’ and the semantic value of ‘blue’ in ‘represents blue’ are ambiguous, this implies: the mind-body problem is the problem of naturalizing intentionality iff ‘represents blue’ means ‘makes S aware of perfect blue’.

But if ‘represents blue’ means ‘makes S aware of perfect blue’, then the problem of naturalizing intentionality (= the mind-body problem) is hard. This is because, in order to solve the mind-body problem, we must explain why a brain state B makes S aware of perfect blue. This type of explanation is a neural functional explanation, i.e., an explanation that explains why a brain state does a certain thing or performs a certain function (the thing or function described by the functional description). Given how neuroscientists solve neural functional problems, it follows that, in order to solve the mind-body problem, we must specify a physical mechanism that shows how B makes S aware of perfect blue. Since this can’t be done, the problem of naturalizing intentionality (= the mind-body problem) is hard.

However, if transparency-intentionalists specify ‘represents’ and ‘blue’ in ‘represents blue’ so as to make the problem of naturalizing intentionality easy, then ‘represents blue’ will

---

<sup>6</sup> The problem of naturalizing intentionality is the problem of explaining why a brain state represents something.

<sup>7</sup> Perfect blue is a primitive, *sui generis* property having a distinct qualitative nature. It is not instantiated in this world since it is not included in a scientific characterization of objects. Whether or not it is the semantic value of ‘blue’ is a complicated issue I discuss in the paper. Also, I save a full characterization of the functional description—one where the sense of ‘aware’ used in the description is fully specified—for the paper as well.

not mean 'makes S aware of perfect blue'. This implies that the mind-body problem is not the problem of naturalizing intentionality. Transparency-intentionalism then is false, and the mind-body problem for experiences of color is hard.

## 6. Understanding the Mind-Brain Relation

In summary: "Puzzles of Pain" argues that there are opaque and transparent problems; "Knowing What Pain Is" specifies the opaque problem and argues that it cannot be solved; and "Does Transparency Solve the Mind-Body Problem?" specifies the transparent problem and argues that it cannot be solved. Though these conclusions have a negative tone, there are four positive points that come out of this dissertation.

First, "Puzzles of Pain" argues for a positive view about experience that has merit independent of issues related to the mind-body problem. In my view, an unwarranted commitment to the **uniformity thesis** has had a negative and undue influence on discussions of experiences. Most philosophers assume that, with respect to properties like *being transparent* and *being opaque*, experiences are not heterogeneous. This unwarranted assumption is responsible for a number of confusions outside of those having to do with the experiential view of pain.

For example, **(1)** the phenomenological and intuitive differences between sensations and colors; **(2)** the fact that sometimes it makes sense to say that mental states make us feel a certain way and sometimes it makes sense to say that mental states make things (external things) seem a certain way; and **(3)** the relation between knowing what *it is like to have a mental state* and knowing what *the referent of a mental state is like* all come into focus only if we accept that experiences differ with respect to transparency. Consider:



(1) is explained by the fact that sensations are opaque experiences that we are aware of while colors are what transparent experiences make us aware of. (2) is explained by the following facts: when we have an opaque experience, we are aware of the feeling that is the experience (that is why we feel a certain way); and when we have a transparent experience, the transparent experience makes us aware of a way a thing seems (that is why something seems a certain way). Finally, (3) is explained by the following facts: in the case of opaque experiences, knowing what it is like to have an opaque experience = knowing what the referent of the opaque experience (= the opaque experience itself) *feels* like = knowing, i.e., being aware of, the feeling that is the opaque experience; and in the case of transparent experiences, knowing what it is like to have a transparent experience = knowing what the referent of the transparent experience seems (e.g., looks) like = knowing, i.e., being aware of, a way the referent seems. (1), (2), and (3) are all difficult to explain under the uniformity thesis. Once that thesis is discarded though, they become explicable.

Second, in “Does Transparency Solve the Mind-Body Problem?”, a key problem afflicting discussions of the mind-body problem is highlighted. Often, the difficult nature of the mind-body problem (for experience) is obscured by the use of theoretical terms we loosely grasp. For example, equivocation on ‘represents’ makes the following line of thought seem reasonable: the mind-body problem is easy because **(i)** words represent things; **(ii)** brain states represent things; **(iii)** to say that (ii) is hard implies that (i) is hard; and **(iv)** (i) is not hard. This line of thought equivocates on ‘represents’ because what ‘represents’ means varies depending on context. What’s more, given the various ways that ‘represents’ is used, it is not plausible that it has a unified sense or meaning. This is important because it highlights the need to center our

discussions of the mind-body problem on what we have an independent, non-theoretical grip on with respect to consciousness. Only by identifying and theorizing from such a point can we correctly understand and characterize the mind-body problem.

Third, it's hard to deny that one of the reasons philosophy—in particular, *doing* philosophy—is valuable is that it helps us get clear on difficult problems. Often, this involves recognizing that our current understanding and characterization of a problem is flawed. In order to make progress on such a problem, we need to understand and characterize it better. To the extent that the three papers making up this dissertation do that for the hard problem of consciousness, then they, despite their negative conclusions, make progress on the problem.

Finally, sometimes, a positive view is implicit in negative conclusions. This is because, often, merely by understanding and characterizing a problem better, a path forward is made clear. How do my conclusions hold a positive view? Well, from “Puzzles of Pain”, we have the conclusion: experiences are items that make us aware of things, either themselves (if they are opaque) or ways (if they are transparent). From “Knowing What Pain Is”, we have the conclusion: pain is a primitive, *sui generis* item having a distinct qualitative nature. And from “Does Transparency Solve the Mind-Body Problem?”, we have the conclusion: experiences of blue make us aware of perfect blue.

Now, on one hand, it could be said that these three conclusions show that mind-brain identities are false. Pains are not brain states because pains are primitive, *sui generis* items while brain states are not; and brain states are not experiences of blue because we cannot

explain why brain states make us aware of perfect blue.<sup>8</sup> These claims imply that **dualism** is true: experiences and brain states are distinct.

There is, however, another way of taking the conclusions. If mind-brain identities require interpretation, then what the conclusions *really* show is that mind-brain identities are false *only under particular interpretations*. For example, mind-brain identities are false under **physicalist** interpretations, i.e., interpretations that favor the brain-side of the identities. What exactly does this mean?

Well, suppose that we have a mind-brain identity of the form ‘M = B’, where M is a mental state (e.g., an experience) and B is a brain state. We know at least two things about this identity. First, it implies that the concept ‘M’ and the concept ‘B’ have a common referent F. Second, it implies that ‘M’ and ‘B’ conceive of or **describe**<sup>9</sup> F in different and inferentially isolated ways (since the identity is *a posteriori*). Given this, we can ask the following question: which of the two descriptions of F—the description associated with ‘M’ or the description associated with ‘B’—get things right with respect to F? In asking this question, I am drawing on a familiar idea made popular by Wilfrid Sellars in “Philosophy and the Scientific Image of Man”

---

<sup>8</sup> The reason why the order of the mental states and brain states are reversed in these two statements is that pains are opaque while experiences of blue are transparent. Because we are aware of pain when we have it, our concept of pain is rigid. That is, we know pain, not by functional description, but by being aware of it. This entails that, if pains and brain states are one, then pain must be a brain state B, necessarily. Though pain does make us aware of itself, the necessity of the identity implies that we don’t have to explain why B makes us aware of itself. The problem, instead, is that our awareness of pain yields knowledge that it is primitive and *sui generis*, and it is this which conflicts with pain being B. By contrast, we know experiences of blue, not by being aware of them, but by functional description. So our concept of an experience of blue is not rigid, and this entails that, if experiences of blue and brain states are one, then brain states must be experiences of blue, contingently. This does require explaining why a brain state makes us aware of perfect blue (i.e., why a brain state satisfies the functional description).

<sup>9</sup> My use of the terms ‘describe’ and ‘description’ are loose. A description, as I use the term, can include things other than mentalese sentences, e.g., it can include phenomenal images or appearances.

and “Empiricism and the Philosophy of Mind.”<sup>10</sup> According to Sellars, there are two images of the world—a manifest image and a scientific image—and, of these two images, it is the scientific image that characterizes the world as it really is (it gets things right with respect to the world).<sup>11</sup> What does this mean? Well, when we examine the *actual* similarity relations among biological species, we find that our manifest concepts of these species get these relations wrong. According to our manifest concepts, a humpback whale is more similar to a whale-shark than it is to a bat, but this we know is wrong. The humpback whale is more similar to the bat than it is to the whale-shark, and it is only our scientific concepts of these species that get these relations right. This same pattern is evident in numerous other cases, e.g., our manifest concept of gold, but not our scientific concept of gold, classifies fool’s gold as real gold.

Extending this idea to mind-brain identities, we would say that ‘B’, not ‘M’, gets things right with respect to F. Roughly, this means that the scientific description of F associated with ‘B’ necessarily and intrinsically describes F (i.e., it is a non-contingent description of F’s intrinsic properties and nature). But conceiving of F in this way is problematic given that our three conclusions imply that (i) F is either an item that makes us aware of itself or a way, and (ii) if F makes us aware of itself, then it is primitive and *sui generis*. Note, however, that this problem dissolves if we abandon the idea that, in the case of a mind-brain identity, it is the brain-side (the scientific side) that characterizes the common referent as it really is. Maybe, for a mind-

---

<sup>10</sup> See also Shoemaker (1994).

<sup>11</sup> Sellars defends the “primacy of the scientific image” (1963a, p. 32), and he claims: “I am quite prepared to say that . . . in the dimension of describing and explaining the world, science is the measure of all things, of what is that it is, and of what is not that it is not” (1963b, p. 173).

brain identity, it is (to modify an expression used by Leopold Stubenberg (1997)) the “north-pole” of the identity—the mind-side—that should be favored.

This is not an unprecedented view of the mind-brain relation. Bertrand Russell once said “we know nothing about the intrinsic quality of physical events except when these are mental events that we directly experience” (1956, p. 164), and “[a]s regards the world in general, both physical and mental, everything that we know of its intrinsic character is derived from the mental side” (1927, p. 402). And Arthur Eddington (1928) claimed:

[O]ur knowledge of the nature of the objects treated in physics consists solely of readings of pointers and other indicators . . . The schedule [of pointer readings] is . . . attached to some unknown background . . . But in one case—namely, for the pointer readings of my own brain—I have an insight which is not limited to the evidence of the pointer readings. That insight shows that they are attached to a background of consciousness . . . If we must embed our schedule of indicator readings in some kind of background, at least let us accept the only hint we have received as to the significance of the background—namely, that it has a nature capable of manifesting itself as mental activity. (258-260)

One way of interpreting these comments by Russell and Eddington is to see them as endorsing the thought that, for mind-brain identities, it is the phenomenal description of F—the description of F associated with ‘M’—that necessarily and intrinsically describes F. The cost of embracing this view is that it is no longer the case that, for all things, the scientific image is prime. In the case of consciousness at least, it is the phenomenal (or manifest) image that is the “measure of things” (see note 11). The benefit, however, of the view may be that it could offer us a way of looking at the mind-brain relation that has many of the strengths of physicalism and dualism and few of their weaknesses.

Let’s say that **phenomenal monism** is the view that ‘M’ and ‘B’ refer to F and the phenomenal description of F gets things right with respect to F (it characterizes F as it really is).

Since dualism holds that 'M' and 'B' don't co-refer, phenomenal monism is clearly different than dualism. It, therefore, doesn't (necessarily) acquire dualism's faults. For example, it need not posit (as dualism must posit) brute psycho-physical laws that account for mental state-brain state correlations. Since physicalism is the view that, though 'M' and 'B' refer to F, the scientific description of F gets things right with respect to F, phenomenal monism is also clearly different than physicalism. It, therefore, doesn't (necessarily) acquire physicalism's faults. For example, it need not deny that pain is a primitive, *sui generis* item having a distinct qualitative nature.

Phenomenal monism, then, seems to occupy a position in-between physicalism and dualism. And, to some degree, it retains some of what is good about physicalism (e.g., its monism) and dualism (e.g., its favoring of phenomenal descriptions) while, at the same time, shedding some of what is bad about these views (roughly, intractable explanatory and justificatory problems and brute psycho-physical laws). Since phenomenal monism not only concords with but is suggested by the three conclusions of my dissertation, one can say that, rather than supporting dualism, my dissertation lays the beginning foundations for a view of the mind-brain relation that is distinct from—and perhaps superior to—the traditional options of physicalism and dualism.

## Chapter 2

## Puzzles of Pain

### 1. Problems with the Experiential View of Pain

To some philosophers, pain is a puzzling phenomenon. On one hand, pain seems to be an **experience**.<sup>12</sup> More specifically, the **experiential view of pain** says that pain is an experience because (i) pain is a sensation (or feeling) of some kind, and (ii) sensations (or feelings) are experiences. But the experiential view of pain seems to have four problems.

First, expressions like ‘experiencing an experience’, ‘sensing a sensation’, or ‘feeling a feeling’ hardly make sense. We don’t have experiences of our experiences, sensations of our sensations, or feelings of our feelings. We just have experiences, sensations, and feelings. But we do say: ‘I am having an experience of pain’, ‘I am experiencing pain’, ‘I am having a sensation of pain’, ‘I am sensing pain’, ‘I am having a feeling of pain’, and ‘I am feeling pain’. This suggests that pain isn’t an experience, sensation, or feeling. Call this the **semantic problem**.

Second, when I have pain, my pain tells me something: it tells me that something is wrong with my body. This suggests that pain is not an *experience* but instead a *content* of an experience. Call this the **content problem**.

Third, when I have an experience of F, the object of awareness (what I am aware of) can be distinguished from the act of awareness (what makes me aware of the object of awareness). What I am aware of is F; what makes me aware of F is having the experience of F; and the experience of F and F are distinct. But when I have pain, what I am aware of is pain. This

---

<sup>12</sup> I discuss later in more detail what experiences are. For now, we can take experiences to be special kinds of mental states (roughly, mental states such that there is something it is like to be in them).



collapses the object of awareness (pain) and the act of awareness (*having pain*). Pain, therefore, can't be an experience. Call this the **awareness problem**.

Finally, because experiences are located in the head (if anywhere at all) and pains are located in distal body parts (e.g., toes), unless I'm willing to admit that toes have experiences (i.e., have mental states), I better conclude that pains aren't experiences. Call this the **location problem**.

In one way or another, I think that all of these supposed problems with the experiential view of pain make the same mistake: they fail to recognize that experiences differ with respect to transparency (or the property of being transparent—I explain later what this means). For reasons that are difficult to know, philosophers have been reluctant to accept the idea that experiences are heterogeneous in regards to what they make you aware of even though, with respect to other properties mental states have, heterogeneity seems to be the rule rather than the exception. In this paper, I will argue that experiences differ with respect to transparency, and once this is recognized, the putative puzzles of pain dissolve. The structure of the paper is as follows. In §2, I argue for the experiential view of pain. In §3, I argue that experiences come in two distinct types: transparent experiences and opaque experiences. And in §4 and §5, I show how, by recognizing this heterogeneity in experiences, the four problems mentioned above can be solved.

## 2. The Argument for the Experiential View of Pain

Michael Tye (2005a) says “pain is a feeling or an experience of a certain sort” (100), and Murat Aydede (2013) claims: “That pain is a subjective experience seems to be a truism.” Many philosophers think that pain is an experience, and, as is evident from the first quote, this

thought is rooted in the experiential view of pain (henceforth, the **experiential view**): (i) pain is a sensation (or feeling) of some kind; and (ii) sensations (or feelings) are experiences (i.e., to have a sensation is to have an experience). In this section, I will argue for both claims of the experiential view.

### 2.1 Pain is a Sensation (or Feeling) of Some Kind

David Lewis (1980) once said: “Pain is a feeling. Surely that is uncontroversial” (222). Why is this uncontroversial? The reason, I think, is this. If, right now, you had a pain (or were in pain), there would be a sensation (or feeling) of some kind you’d be *aware* of. By ‘aware’, I don’t mean ‘*de dicto* aware’; that is, I don’t mean that you are aware of the sensation only by being aware of some fact about the sensation. You are aware of the sensation *itself*: you are *de re* aware of the sensation. Also, by ‘aware’, I don’t mean ‘indirectly aware’, i.e., aware of the sensation by being aware of something else. You are *directly aware* of the sensation. So, if you had a pain right now, there would be a sensation of some kind you’d be directly *de re* aware of, and in what follows whenever I say ‘aware’ I mean ‘directly *de re* aware’ (unless I specify otherwise).

Now consider the general conditions under which you’d say: ‘I have a pain’ (or ‘I am in pain’). If you were *not* aware of an instance of that kind of sensation (the kind mentioned above), would you say ‘I have a pain’? This seems highly doubtful. Moreover, if you *were* aware of an instance of that kind of sensation, it is highly plausible that you’d say: ‘I have a pain’. This suggests that (abstracting to an arbitrary, normally functioning<sup>13</sup> subject of experience S): S

---

<sup>13</sup> The significance of assuming that S is normally functioning is that S will judge ‘I have a pain’ only if S is aware of an instance of that kind of sensation.

judges 'I have a pain' (or 'I am in pain') iff S is aware of an instance of that kind of sensation (or feeling).

Now consider: do (instances of) sensations (or feelings) occur unbound? That is, are there free-floating sensations? That possibility is hard to make sense of. Sensations seem to be things that are *had* by things. What kinds of things have sensations? Do tables have sensations? How about rocks? Neither tables nor rocks have sensations because neither tables nor rocks have *minds*. Evidently, the *only* things that have sensations are things with minds. Things like S.

So S has sensations, and S is aware of sensations. Is it possible for S to have a sensation but not be aware of it? For example, could you right now be having a sensation (or feeling) even though you are not aware of any sensation (or feeling)? That does not seem right. It also doesn't seem right to say that you could be aware of a sensation (or feeling) but not have it. In fact, the following seems true: to have a sensation is to be aware of it, and to be aware of a sensation is to have it. Thus, if S has a sensation, then S is aware of that sensation, and if S is aware of a sensation, then S has that sensation. Combining this with the earlier biconditional, we have: S judges 'I have a pain' (or 'I am in pain') iff S has an instance of that kind of sensation (or feeling).

Unless you think that S could be having an instance of that kind of sensation but be wrong in judging 'I have a pain', the above biconditional suggests the following: to have pain is to have that kind of sensation, and to have that kind of sensation is to have pain (i.e., to have a pain = to have an instance of that kind of sensation). Pain, then, is that kind of sensation (or feeling), and a pain is an instance of that kind of sensation (or feeling). To the extent that the

preceding argument is uncontroversial, it is uncontroversial that pain is a sensation (or feeling) of some kind. This yields the first part of the experiential view.

One consequence of pain being that kind of sensation is that pain is not localized, peripheral body damage (or ‘body damage’, for short). If your hand is anesthetized and your finger is cut, you have a cut (i.e., an instance of body damage). Since your hand is anesthetized, you are not aware of an instance of that kind of sensation; so you do not have an instance of that kind of sensation. Therefore, you do not have a pain (that, of course, is the point of the anesthesia). This implies that ‘S has an instance of body damage’ and ‘S has a pain’ have different truth conditions. And this, in turn, implies that, though body damage can cause pain, body damage *is not* pain.

## 2.2 Sensations (or Feelings) are Experiences

Pain is commonly thought to be an experience because sensations (or feelings) are commonly thought to be experiences. Why is that the case? It helps here to contrast sensations with colors.

On one hand, sensations and colors seem to be similar in that they both have *qualitative natures* that are difficult to describe using “non-sensation” and “non-color” vocabulary (respectively). For instance, it is difficult to describe the qualitative natures sensations (colors) seem to have—even what sensations (colors) are—to someone who has never had a sensation (seen a color) before. This suggests that sensations and colors are both *sui generis* things having distinct qualitative natures.

But there are also important differences between sensations and colors. Colors seem to inhere in **external objects**, i.e., objects distinct from and external to the mind. Because of this,

colors seem to be **intrinsic and objective properties** of external objects. By ‘objective’, I mean that colors seem to be properties that can be instantiated by objects regardless of whether subjects are aware (in any sense of ‘aware’, e.g., *de dicto* aware, indirectly aware) of them. For example, if tomorrow some calamity befalls the universe so that every being with a mind dies, it’s not hard to imagine things still being colored (e.g., green). In fact, this is probably the naïve or common sense view.

Note that the objectivity of colors extends to your own body. Even if you or no one else is aware (in any sense) of the color of your finger, it’s not hard to imagine your finger still being colored. Even if your finger is severed, placed in an opaque box, and buried, you still wouldn’t expect it to suddenly lose its color. The fact that colors seem to be intrinsic and objective properties of external objects pushes them out of the mind. Colors, we can say, seem **external** in nature (i.e., external to the mind).

Given the discussion in §2.1, sensations seem different. Whereas it seems possible for something to be colored even though no one (including the thing that is colored) is aware (in any sense) of the color or the thing colored, it does not seem possible for something to have a sensation and no one be aware of it. This is because the thing having the sensation must have a mind, and it must also be (directly *de re*) aware of the sensation they are having. Also, if tomorrow every being with a mind dies, it doesn’t seem possible for things—dead bodies, rocks, trees, etc.—to have sensations. And a severed finger in a buried box doesn’t somehow contain a sensation that no one is aware of.<sup>14</sup>

---

<sup>14</sup> The phenomenological and intuitive differences between colors and sensations are discussed by many philosophers. See, e.g., Chalmers: “[T]he phenomenology of pain is quite different from the phenomenology of color, where we have no trouble conceiving of an object being perfectly colored even though no one ever

These considerations push sensations into the mind. But why are sensations *experiences*? To answer this question, we have to think about what experiences are. Experiences are supposed to be mental states of a distinct sort. It goes without saying that we have many different kinds of mental states. Some of our mental states are sub-personal in that we could never become *de dicto* aware of them merely through introspective reflection. The early mental processes involved in understanding language are examples of these kinds of states. Some of our mental states are unconscious in that, though it is possible to become *de dicto* aware of them merely through introspective reflection, it is very hard, for various psychological reasons, to do so. Some of our deepest desires and beliefs may be like this. Other of our mental states are unconscious—in another sense—in that we are merely not currently *de dicto* aware of them through introspective reflection. Beliefs that you are *not* currently thinking about are examples of such states. Finally, some of our mental states are conscious in that we are currently *de dicto* aware of them through introspective reflection. Beliefs that you *are* currently thinking about are examples of such states.

Some of our mental states, however, seem to have another feature. When we have these mental states, there is something it *feels* like to us to have these mental states.<sup>15</sup> When I say ‘there is something it *feels* like to us’ I mean the following: when we have these mental

---

experiences its color” (113-114); and Block (2005): “[W]e have to recognize that colors are objective in a way that Subjective Qualities [i.e., sensations] are not” (140).

<sup>15</sup> See Nagel (1974). Also, see Stoljar (2016) for an analysis of the semantics of ‘what it is like’ sentences. According to Stoljar, ‘what it is like’ sentences express propositions about events (e.g., S having a mental state) affecting subjects (e.g., S) in affective ways (i.e., subjects *feel* a certain way as a result of the event).

states, there is a *feeling* (or sensation) we are (directly *de re*) aware of. If a mental state has this feature, then we say (or we can say) that it is an **experience**.<sup>16</sup>

Now, earlier, I said that sensations (or feelings) are items that things with minds have. I also claimed that S has a sensation iff S is (directly *de re*) aware of the sensation. Finally, I said that sensations are pushed into the mind (i.e., they are mental in nature). These considerations suggest that sensations are mental states. Moreover, they suggest that sensations (or feelings) are experiences because, when subjects have them, there is something it *feels* like to these subjects to have these sensations, i.e., these subjects are aware of a sensation (or feeling). In particular, they are aware of the sensation (or feeling) they have.

Note that colors do not satisfy the above *sufficient condition* for being an experience. This is because there is nothing it *feels* like to have a color (to be colored). If an object is colored, that does not imply that the object feels a certain way, i.e., is aware of a feeling. Thus, despite colors and sensations both seeming to be *sui generis* things having distinct qualitative natures, only sensations satisfy the above sufficient condition for being an experience.

As a result, because pain is a sensation and sensations are experiences, pain is an experience. This completes the argument for the experiential view.

### 3. Transparency, Pain, and the Experience of Color

In §2.2, I claimed that we have many different kinds of mental states. We have subpersonal, unconscious (in two senses of ‘unconscious’), and conscious mental states. In addition, we have mental states such that, when we have them, there is something it *feels* like

---

<sup>16</sup> So if I want to know whether or not my computer has experiences, one way to determine this is as follows: determine whether my computer has an internal state such that, when the computer is in that state, there is something it *feels* like to the computer to be in that state, i.e., there is a feeling the computer is aware of.

to us to have these mental states (there is a feeling we are aware of). But, in addition to these mental states, we seem to have another kind of mental state. Some of our mental states are such that, when we have them, it's not that we *feel* a certain way; instead, something *seems* a certain way. And when I say 'something *seems* a certain way' I mean that we are (directly *de re*) aware of a *way* that thing seems.<sup>17</sup>

The reason why it seems that we have such mental states is that, though having a color does not entail that the thing having the color feels a certain way, we do seem to be aware of colors in much the same way we are aware of sensations.<sup>18</sup> Though in the case of sensations we are aware of them by having them, the same can't be said for colors. It must, therefore, be the case that we are aware of colors by having mental states—distinct from the colors themselves—that make us aware of the colors (that make it seem that something is colored). These mental states, then, are such that: when we have them, something seems to be colored, i.e., we are aware of a color.

If we call mental states like these **experiences** too, then we (evidently) have (at least) two types of experiences. One type of experience is such that, when we have it, we *feel* a certain way, i.e., we are aware of the *feeling* that is the experience we have. The other type of experience is such that, when we have it, something seems a certain *way*, i.e., we are aware of a *way* that thing seems. We can call the former experiences '**opaque experiences**', and the

---

<sup>17</sup> See Stoljar (2016) where he emends his original analysis of 'what it is like' sentences in terms of a person feeling a certain way to: "it is stereotypical to use a 'what it is like'-sentence in contexts in which what is at issue is either how a person feels or how things seem to the person" (1181-1182).

<sup>18</sup> At least, if colors are taken to be the *sui generis* qualitative items they seem to be. If they are not, e.g., if they are surface spectral reflectances, then the claim made in the text is not true. In what follows, I'll assume that colors are *sui generis* qualitative items.



latter experiences ‘**transparent experiences**’.<sup>19</sup> Sensations are opaque experiences because, when we have a sensation, we *feel* a certain way, i.e., we are aware of the sensation (or *feeling*) that we have. Experiences of color, on the other hand, are transparent experiences because, when we have an experience of color, something seems to have a certain color, i.e., we are aware of the color the thing seems to have.<sup>20</sup>

Is it implausible that experiences are heterogeneous with respect to transparency? As was made clear in §2.2, we have many different kinds of mental states. So heterogeneity seems to be the rule rather than the exception for mental states. Curiously, philosophers have been reluctant to extend the general heterogeneity of mental states to the realm of experiences. Many think that *experiences* can’t be heterogeneous because they endorse what might be called the **uniformity thesis for experience**: experiences cannot differ with respect to properties like *being transparent* or *being opaque*.

Theses similar to the uniformity thesis, though rarely argued for, are often taken for granted in the literature. Amy Kind (2014), for instance, in arguing against representationalism, says “the plausibility of the view hinges in large part on its applicability across all phenomenal states” (115). Daniel Stoljar (2004), in considering the possibility of endorsing the existence of qualia for some experiences but not others, says: “[S]urely the qualia realist does not want to be maneuvered into the position of saying that color experiences lack qualia. It would be an

---

<sup>19</sup> This terminology is motivated by the thought that, e.g., contact lenses are transparent because, when I wear them, I am not (directly *de re*) aware of them. I am *de dicto* aware of them, i.e., aware *that* I am wearing them. Experiences of color are like contact lenses in that, when I have an experience of color, I am not (directly *de re*) aware of the experience. I am *de dicto* aware of the experience, i.e., aware *that* I am having the experience. See Dretske (1995, chapter 2).

<sup>20</sup> Note that, under this view, part of what makes something an experience is that it makes the subject who has it aware of something (either itself, e.g., a sensation, or something external, e.g., a color).

odd sort of position indeed which postulates qualia but then adds that qualia are only instantiated in cases in which you face the sun with closed eyes, or else are in states of sexual climax” (368). And Clare Batty (2010) takes the thesis that “certain philosophical issues about perception should be settled in the same way for each of the sensory modalities” as a starting point for her discussion of the representational content of olfactory experiences, claiming that this thesis—which she calls the “unification thesis”—is “implicit in much of the contemporary perceptual literature” (513).

If the uniformity thesis were true, then theorizing about experiences would be easier. And perhaps something like the uniformity thesis should be, other things being equal, an operating assumption. But when it comes to mental states, in general, and experiences, in particular, other things don’t seem to be equal. Mental states seem to differ in dramatic and important ways, and, in virtue of these differences, they fall into different categories. Why should we expect the heterogeneity present throughout all of mentality to suddenly stop when the mental state in question becomes an experience (i.e., a phenomenal mental state)? Add to this the fact that sensations and experiences of color, despite both being experiences, seem different with respect to what we are aware of when we have these experiences, there is little reason to think that the uniformity thesis is actually true.

#### 4. The Transparent Framework

I’ve argued that sensations (e.g., pains) are experiences. I’ve also argued that experiences come in two distinct types: opaque experiences (e.g., sensations) and transparent experiences (e.g., experiences of color). In §1, I claimed that the four problems associated with the experiential view—i.e., the semantic, content, awareness, and location problems—can be

resolved by recognizing that experiences differ with respect to transparency. Before I explain why that is the case, I need to argue for one more point.

Clearly, we have experiences; and, clearly, for reasons having to do with reproductive fitness, we've developed a language—a linguistic framework—to talk about our experiences. If our experiences were uniform with respect to transparency, then we would need only one linguistic framework to talk accurately about all our experiences. Unfortunately, our experiences haven't cooperated with us: we have transparent and opaque experiences.

Because our experiences differ in this way, if we want to talk accurately about all our experiences, we need two linguistic frameworks: we need a **transparent framework** apt to talk about **transparent experiences**, and we need an **opaque framework** apt to talk about **opaque experiences**. Of note, transparent (opaque) frameworks are not apt to talk about opaque (transparent) experiences.

Though we have both frameworks (i.e., both frameworks have evolved), the transparent framework has, unsurprisingly, become the dominant way we talk about experiences. I say 'unsurprisingly' because, when it comes to increasing our odds for survival, it is crucial that we communicate information about *external* items—e.g., predators, prey, and food. In order for us to, say, survive an encounter with a lion, we need to communicate information about the lion and other elements in the external world, and we need to coordinate our behavior based on this information. In light of this, and because transparent experiences are externally directed, the transparent framework is far more important to our survival than the opaque framework. Moreover, given the critical role the transparent framework plays in promoting survival, it is likely that it evolved first.

For (at least) these two reasons, one would expect the transparent framework to be the dominant framework, and, in fact, it is. It is the default and primary way we talk about experiences, in particular, the default and primary way we attribute experiences to ourselves. As a result, though we should and we can talk about opaque experiences using the opaque framework, like an old habit that is hard to break, we have a hard time resisting talking about opaque experiences using the transparent framework. This causes infelicities when we attribute opaque experiences (e.g., pains) to ourselves (using the transparent framework). In ordinary life, these infelicities are harmless. But in philosophical life, they can cause us to think that pain is puzzling.

## 5. Resolving the Problems

In this section, I will show how the fact that (i) we have transparent and opaque experiences, and (ii) we have a dominant transparent framework apt only for transparent experiences (and a sub-dominant opaque framework apt only for opaque experiences) can resolve the four problems associated with the experiential view.

**(1) The Semantic Problem.** According to the semantic problem, expressions like ‘experiencing an experience’, ‘sensing a sensation’, or ‘feeling a feeling’ make little sense. We don’t have experiences of our experiences, sensations of our sensations, or feelings of our feelings. We just have experiences, sensations, and feelings. But we do say things like: ‘I am having an experience of pain’, ‘I am experiencing pain’, ‘I am having a sensation of pain’, ‘I am sensing pain’, ‘I am having a feeling of pain’, and ‘I am feeling pain’. Doesn’t this show that pain isn’t an experience, sensation, or feeling?

The problem here (as Kripke (1980, p. 152) observed) is that the experience of pain and pain are one and the same; the sensation of pain and pain are one and the same; and the feeling of pain and pain are one and the same. In other words, to have an experience of, a sensation of, or a feeling of pain is just to have pain. So all of the expressions involving pain are equivalent to: 'I have pain' (or 'I am in pain').<sup>21</sup> Moreover, they are infelicitous ways of attributing pain to myself and they can't be taken at face value. Here is why.

'I have pain' (or 'I am in pain') are felicitous ways of talking about pain because they straightforwardly attribute to myself an experience I have (or a phenomenal mental state I am in). Why then do we talk in the infelicitous ways? It is because of the dominance of the transparent framework. Transparent experiences have the following feature: the experience (experience of F) and what you are aware of (F) when you have the experience are distinct. And when we attribute transparent experiences to ourselves, we say things like, 'I have an experience of F' or 'I see F'. So when we attribute transparent experiences to ourselves, the item (F) we are aware of fills the argument place in 'I have an experience of \_\_\_' or 'I see \_\_\_'. Because the transparent framework is dominant (i.e., it is the default and primary way we talk about experiences), this leads to the following tendency: to attribute an experience to myself, I say, 'I have an experience of \_\_\_' or 'I  $\phi$  \_\_\_', where ' $\phi$ ' is an experiential verb (e.g., 'see', 'hear', 'taste', 'feel', etc.) and the argument place (the blank) is filled in by what I am aware of. In the case of pain, this yields: to attribute pain to myself, I (infelicitously) say, 'I have an experience of pain' and 'I feel pain'.

---

<sup>21</sup> See Lewis (1980): "To have pain and to feel pain are one and the same" (222).

Statements like these can't be taken at face value. They are the product of using the transparent framework to attribute to myself an opaque experience. Under an opaque framework, to attribute an experience to myself, I say, 'I have E' or 'I am in E', where E is both what I am aware of and the experience I have (to have an experience of E is to have E, and to feel E is to have E). If I used the linguistic framework (the opaque framework) apt for the kind of experience pain is (an opaque experience), then I would only get the felicitous claims 'I have a pain' or 'I am in pain' when I attributed pain to myself. The semantic problem would then dissolve.<sup>22</sup>

**(2) The Content Problem.** According to the content problem, pains tell me something about the world: they tell me that something is wrong with my body. Alex Byrne (2001), for instance, says: "[W]hen one has a pain in the toe, *the world seems a certain way*, namely, that there is 'some hurt or disorder' in the toe" (228-229; italics in the original). The concern for the experiential view, then, is this. Because a pain in my toe tells me that something is wrong with (i.e., there is disorder or damage in) my toe, that pain must be the *way* that an experience represents the damage in the toe. Pain, then, is an *appearance property* of body damage, and this rules out pain being an experience since appearance properties are *ways* things are represented by experiences. They are not experiences themselves.

Again, in my view, the dominance of the transparent framework is to blame for the existence of the content problem. In the case of a transparent experience (an experience of F), what we are aware of (F) is not an experience. Thus, it *could be*—and often is *taken as*—a way

---

<sup>22</sup> Note that there is no tendency, in the case of color, to say, e.g., 'I have blue' or 'I am in blue' since the opaque framework is not dominant. The dominance of the transparent framework is why we apply it to experiences it is not apt for.

some item is represented by the experience. So when we talk about having an experience of F—or even seeing F—F is often regarded as an *appearance property* of another item G, where G could be related to the experience causally (i.e., it could cause the experience). This leads (because of the dominance of the transparent framework) to the following tendency: when I attribute an experience to myself, I should conclude that what I am aware of when I have the experience is an appearance property of the cause of the experience.

Applied to the case of pain, this yields: when I attribute pain to myself, I should conclude that the pain is an appearance property of, e.g., the cut (since the cut causes the pain). This is problematic since experiences (pains) can't be appearance properties of things (cuts); but, of course, the real problem here is that rules apt for transparent experiences embedded in the transparent framework are being applied to an opaque experience (pain). Under an opaque framework, it isn't the case that, when I attribute an experience (a pain) to myself, I should conclude that what I am aware of (the pain) is an appearance property of the cause of the experience (the cut) since what I am aware of (the pain) is the experience.<sup>23</sup>

This doesn't mean that the experience (the pain) doesn't have content. By virtue of being caused by the cut, the pain could have content about the cut in much the same way that fire alarms have content about fires. Such content (i.e., derived content) is had by virtue of

---

<sup>23</sup> This is made clear by the fact that, while we say 'experience of a table' or 'see a table', we don't say 'experience of a cut' (we infelicitously say 'experience of a pain') or 'feel a cut' (we infelicitously say 'feel a pain') (we might however say 'it feels *like* I have a cut'). What this shows is that, in the transparent case, we naïvely take what we are aware of to *be* the putative cause of the experience—i.e., no distinction is made between F and G—and this supports the view that F is an appearance property of G. We don't, however, have this tendency in the case of pains and cuts, which undercuts the idea that pains are appearance properties of cuts.

beliefs about causes,<sup>24</sup> and pain could have this kind of content about the cut without being an appearance property of the cut. The content problem, then, is a by-product of the dominance of the transparent framework, and it is not a real problem for the experiential view.

**(3) The Awareness Problem.** The awareness problem states that pain can't be an experience because (i) the object of awareness (F) and the act of awareness (*having an experience of F*) are always distinct, and (ii) pain is an object of awareness. Joseph Levine's (2006) concern with pain seems to be due to the awareness problem:

The difficulty in understanding how to treat pain is manifested beautifully by the apparent ambiguity of the word 'feeling', as it occurs in 'feeling pain'. David Lewis famously said that pain is essentially a feeling . . . [B]ut what is a feeling anyway? . . . On the one hand, we talk of 'feeling a pain', which makes the pain seem like the object of awareness, and the feeling the act of awareness . . . On the other hand, if one adopts Lewis's slogan, 'pain is a feeling', then it's the feeling that is the object of awareness, it is not the act of awareness itself . . . Treating [this] . . . as a matter of verbal ambiguity . . . certainly eases the naturalization task; but I think it hides what is in fact the very puzzling nature of pain, and feelings generally. They seem to involve an act-object structure, so that there is something one is feeling when one is having a feeling. On the other hand, what's being felt, the apparent object, seems to itself be a mode of conscious experience, not an extramental feature of an extra-mental object. If you just contemplate that throbbing in your finger when, say, you jam it in a door, you know what I mean. You are aware of the throbbing, it is the object in that sense, but throbbingness seems to be itself part of the very stuff of conscious experience. I think this is indeed puzzling, and I do not know how precisely to understand the phenomenon. (278-279)

Levine's puzzlement seems to be centered on the difficulty reconciling talk of 'pain being a feeling' with talk of 'feeling pain'. I've already discussed the problems with the locution 'feeling pain' but Levine's concern runs deeper than that. Since to *feel* something is to have a *feeling of*

---

<sup>24</sup> You might think that pains at least carry original content to the effect that "something bad is going on with my body". But even that isn't the case since the pain I have after a hard workout or after applying alcohol to a wound could indicate something good is going on (e.g., my muscles are getting bigger and stronger, and my wound is getting disinfected). Sometimes pain is a good sign.



some kind, and since *feeling* something is an act of awareness (i.e., it is something that makes you aware of the thing you are feeling), *having a feeling* is an act of awareness. So *having a feeling* is like *having a (transparent) experience of F* in that they are both acts of awareness.

Now in the case of *having an experience of F*, the object of awareness—F—is different from the act of awareness (*having an experience of F*). This leads to the following tendency (assuming the dominance of the transparent framework): when I attribute an experience to myself, I should conclude that the object of awareness is distinct from the act of awareness.

If pain is a feeling and I have pain, then *having pain* is an act of awareness. But when I have pain, I am aware of pain, so pain is the object of awareness. Given the above tendency, this implies that pain is distinct from *having pain*, which of course is puzzling. Clearly, though, the problem (again) is that a rule embedded in the transparent framework is being applied to an opaque experience (pain). Under an opaque framework, the act of awareness (*having pain*) is not distinct from the object of awareness (pain) since, when you have an opaque experience (pain), you are aware of the experience (pain).<sup>25</sup> The awareness problem thus dissolves once we recognize that experiences—and their corresponding linguistic frameworks—are heterogeneous with respect to transparency, and one of these frameworks—the transparent framework—is dominant.

**(4) The Location Problem.** Finally, the location problem states that, because experiences are located in the head (if anywhere at all) and pains are located in body parts (e.g., toes), pains are not experiences. David Bain (2007) states the location problem in the following way:

---

<sup>25</sup> This suggests that F can be an object of awareness in one of two ways: (i) F can be a *way* an occurrent transparent experience *represents* an item, or (ii) F can be an occurrent opaque experience.

I know of no philosopher who thinks both that we have pains in limbs and that the experience view explains what it is for us to do so. That would strike most as madness. But the experience view is nevertheless worth registering partly because it can seem to be no more than a consequence of common sense: pains are experiences, it can seem natural to say, hence if we have pains in our limbs, we must have experiences in our limbs, and what could it mean to say an experience is *in* a limb if not that the limb *undergoes* the experience? (185; italics in the original).

Because it is absurd to think that our limbs have experiences, Bain thinks that the experiential view—the view that pains are experiences—must be rejected.<sup>26</sup>

If we let L stand for a distal body part (e.g., a toe or a limb), then the location problem has the following argument structure:

- (1) I have a pain in L.
- (2) The pain is *located* in L. [inferred from 1]
- (3) L has the pain. [inferred from 2]
- (4) If pains are experiences, then L has an experience. [inferred from 3]
- (5) L cannot have experiences.
- (6) Pains are not experiences. [inferred from 4 and 5]

Alternatively, we might say that (2) straightaway implies that pains aren't experiences since experiences can't be located in distal body parts.

The key issue here is whether (2) follows from (1). Does 'I have a pain in L' imply 'the pain is *located* in L'? In my view, it does not since 'I have a pain in L' can be true regardless of whether you open up L and find a pain *located* there. Expressions like 'I have a pain in L' do not portend to *locate* pains in body parts. Instead, they are self-attributions of tokens of *types* of experiences.<sup>27</sup> To make the discussion in this paper flow smoother, I've been using the expression 'pain' to talk about pains; but the expression 'pain' doesn't fully capture the

---

<sup>26</sup> See Harman (1990, pp. 39-40) and Byrne (2009b, p. 270) for similar kinds of arguments.

<sup>27</sup> Armstrong (1968) has a view similar to this. See also Aydede (2005, 2013).

different *types* of pains you might have. 'Pain' is a general term under which different types of pain fall under; and one thing (not the only thing) that identifies these different types is the expression 'in L' added to 'pain'. To identify a certain type of pain then, one of the things we do is use the expression 'pain in L'.<sup>28</sup> And when I have a pain in L, this just means that I have a token of a certain type of experience.

Consequently, when I say, 'I have a pain in L', I'm not saying that a pain is *located* in L. I'm saying that I'm having a token of a type of experience, the type *pain in L*. The same applies to other statements that seem to locate pains, e.g., 'there is a pain in L'. This statement makes the existentially quantified claim that there exists a token of the type of experience *pain in L*, and I am having a token of that type of experience. The statement does not imply that a pain is *located* in L, so the inference from (1) to (2) present in the location problem is invalid.

Why is it easy to think that (2) follows from (1) even though 'the pain is located in L' is (pretty clearly) not a truth condition for 'I have a pain in L'? I think again the dominance of the transparent framework is to blame. When I have an experience of an F (e.g., an apple) on a table, my (transparent) experience suggests that the F is *located* on the table. This leads to following tendency: when I attribute to myself an experience and I am aware of an F at (or on or in) such-and-such place, I should conclude that the F is *located* at (or on or in) such-and-such place.

Applied to an opaque experience, e.g., pain, this yields: when I have a pain in L and I am aware of a pain in L, I should conclude that the pain is located in L. But note that my experience

---

<sup>28</sup> How does someone having a phantom pain but not having an L denote the type of experience they are having? They say, 'I have a pain where my L used to be', where 'pain where my L used to be' denotes the same type of experience as 'pain in L' and does *not* imply that the pain is *located* where their L used to be.

of F on a table doesn't suggest that the *experience* is located anywhere. Nor is 'the F is located on the table' a truth condition for me having the experience. The expression 'of an F on a table' only serves to type-identify the experience; it does not locate the experience. Similarly, the expression 'in L' only serves to type-identify, not locate, the pain. And me having the pain doesn't depend in any way on the pain being located in L.

More generally, we can say the following. The reason why it is reasonable to conclude 'the F is located at L' when one has a (transparent) experience of F at L is because, when one has the experience, one is (possibly) aware of F and its location. However, in the case of an opaque experience, e.g., a pain, when one has the pain in L, one is merely aware of a token of a type of experience (not the location of the token experience you have). Thus, it is not reasonable to conclude that the pain is located in L. The general tendency to do so is due to the influence of the transparent framework.

The transparent framework may also be responsible for the inference of (2) from (1) in the following way. If terms like 'see' are success terms—as some philosophers believe—then it is plausible that 'I see an F at L' has 'the F is located at L' as one of its truth conditions. This leads to the following tendency: I should take 'the F is located at L' to be the truth condition for 'I  $\phi$  an F at L'. Now we already know that, because of the dominance of the transparent framework, the infelicitous expression 'I feel pain' is often used in place of 'I have pain'. Because of this, it is easy to slide from 'I have a pain in L' to 'I feel a pain in L', and once this slide is made, the tendency mentioned above leads to the following thought: 'the pain is located in L' is a truth condition for 'I feel a pain in L', i.e., 'I have a pain in L'. But this thought is a result of the transparent framework being misapplied to an opaque experience, and again we

can see that the location problem is due to the undue influence of the transparent framework. It is not a real problem for the experiential view.

I conclude that none of the supposed problems with the experiential view are actual problems.

## 6. Two Hard Problems

In this paper, I have argued that four problems (the semantic problem, the content problem, the awareness problem, and the location problem) that supposedly dog the experiential view of pain (the view that pains are sensations and sensations are experiences) can be solved by recognizing: (i) experiences come in two distinct types: transparent experiences and opaque experiences; (ii) there are two linguistic frameworks apt for these two types: the transparent framework and the opaque framework; and (iii) the transparent framework is dominant. In my view, an unwarranted commitment to the uniformity thesis (the thesis that experiences are alike with respect to, e.g., transparency) is mostly to blame for why these facts about the heterogeneity of experiences are often set aside and ignored.

There is another reason why we may be reluctant to accept that experiences are heterogeneous with respect to transparency. If we have transparent and opaque experiences, then there is likely no single 'hard problem of consciousness'. There would be two hard problems: a transparent hard problem associated with transparent experiences, and an opaque hard problem associated with opaque experiences. What these problems are and whether they

can be solved are topics I address in other papers.<sup>29</sup> But the heterogeneity of experiences may be seen as bad news because it doubles our problems with consciousness.

---

<sup>29</sup> I examine the opaque hard problem in “Knowing What Pain Is” and the transparent hard problem in “Does Transparency Solve the Mind-Body Problem?”

## Chapter 3

## Knowing What Pain Is

### 1. The Hard Problem of Consciousness and the Phenomenal Concept Strategy

The **hard problem of consciousness** (henceforth, the **hard problem**) comprises three arguments against physicalism: the conceivability argument, the knowledge argument, and the explanatory gap argument. According to the conceivability argument, because zombies are conceivable—e.g., it is conceivable that a non-conscious physical duplicate of me could exist—zombies are possible; and because zombies are possible, phenomenally conscious mental states—or **experiences**—can't be brain states. The knowledge argument states that Mary—a fictional neuroscientist who, despite knowing all the physical facts there are to know about color and the experience of color, hasn't seen color herself—lacks knowledge of one crucial fact about the experience of color: what it is like to have the experience. That fact, then, can't be a physical fact. Finally, the explanatory gap argument states that, because it is not possible to explain why being in a certain brain state feels (or seems) a certain way, an explanatory gap exists between experiences and brain states. The best explanation for this gap is that experiences are not brain states.

In one way or another, all of the arguments making up the hard problem turn on a special feature of **phenomenal concepts**, i.e., our concepts of experiences.<sup>30</sup> Specifically, phenomenal concepts are associated with **non-contingent modes of presentation** of their referents, and this feature of phenomenal concepts, e.g., the concept of pain or 'pain'<sup>31</sup>, makes

---

<sup>30</sup> I explain later what the items I mention in this paragraph—phenomenal concepts, modes of presentation, and manifest concepts—are.

<sup>31</sup> I use single quotes to denote concepts.



them very different from **manifest concepts**, e.g., the concept of water or ‘water’. This difference between (say) ‘water’ and ‘pain’ is at the root of the hard problem.

Prior to the development of the **phenomenal concept strategy**, physicalists often responded to the hard problem by making special exceptions for experiences like pain. For example, though in nearly all cases where a manifest item is identified with a scientific item, the manifest concept of the item is associated with a contingent mode of presentation (henceforth, I’ll abbreviate ‘mode of presentation’ to ‘**MOP**’), in the case of pain, physicalists would claim, often without explanation, that pain is a brain state even though ‘pain’ is associated with a non-contingent MOP. This kind of *ad hoc* response to the hard problem struck many people as unsatisfying.

Beginning with the work of Brian Loar (1997) and continuing forward with the works of Brian McLaughlin (1999, 2012), Ned Block (2007), David Papineau (2007, 2011), Katalin Balog (2009, 2012a, 2012b), and others<sup>32</sup> this problem with the physicalist response to the hard problem was rectified. A new strategy for rebutting the three arguments was developed, and that strategy has come to be known as the **phenomenal concept strategy**<sup>33</sup> (henceforth, the **PCS**). What the PCS is and how it works is something that I will discuss later. For now, what is important to know about the PCS is that it solves the hard problem by appealing to a particular theory of phenomenal concepts: the **constitutional theory**.<sup>34</sup> According to PCS proponents, the

---

<sup>32</sup> For other examples, see Tye (1999) (though Tye is no longer a proponent of the PCS), Perry (2001), Carruthers and Veillet (2007), and Levin (2008).

<sup>33</sup> Stoljar (2005) coined this term.

<sup>34</sup> In truth, not all versions of the PCS appeal to the constitutional theory, though this is by far the most popular version of the strategy. What unifies all versions of the PCS is an appeal to a theory of phenomenal concepts that affords a direct experiential and conceptual awareness of experiences. As will be made clear later in the paper, it is

constitutional theory has bi-partisan support: both dualists and physicalists find the theory appealing. What's more, the constitutional theory is compatible with physicalism. Most important, the constitutional theory explains why, e.g., 'pain' is associated with a non-contingent MOP, and this fact about the constitutional theory lies at the heart of the PCS. In short, the PCS solves the hard problem by putting forward an independently plausible theory of phenomenal concepts that explains—in a manner consistent with physicalism—why phenomenal concepts are associated with non-contingent MOPs. The PCS, then, can't be accused of being *ad hoc* since it *explains why*, e.g., pain could be a brain state even though 'pain' is associated with a non-contingent MOP.

The PCS has gained widespread support among philosophers of mind and many think that it satisfactorily addresses the hard problem. I will argue, however, that the PCS has a problem: it incurs a commitment that is at odds with physicalism. Specifically, the constitutional theory—to which the PCS is committed—implies that, when I have pain, I am directly aware of pain. And what I will argue in this paper is the following: my direct awareness of pain allows me to know something about pain that is at odds with the claim that pain is a brain state. The structure of this paper is as follows. In §2, I discuss some preliminary matters about concepts, modes of presentation, and claims of the form 'x is y', and I also explain what phenomenal concepts are according to the constitutional theory. In §3, I explain what the PCS is and how it solves the hard problem. And in §4, I argue that, given the constitutional theory, my direct

---

this direct awareness of experiences that is central to my argument. In addition, I believe the constitutional theory is the best theory of phenomenal concepts, and I explain why later in the paper. For these reasons—and to simplify the presentation of my argument—I choose to focus exclusively on the version of the PCS based on the constitutional theory.

awareness of pain puts me in a position to know that pain is not a brain state. Ultimately, what my argument shows is that the tool the PCS uses to solve the hard problem simply exchanges one hard problem for another.

## 2. Preliminaries

In order to understand how the PCS works and why it faces a problem, we have to get clear on some preliminary matters. In this paper, my focus will be on pain and how the PCS defends the view that pain is a brain state. Two preliminary matters then need to be discussed. First, we have to understand what claims of the form 'x is y' say since our claim of interest is 'pain is B' (B is some brain state). Second, in order to know what we even mean by 'pain', we have to get some independent grip on pain and the concept 'pain'. Let's start with the first issue.

### 2.1 Modes of Presentation and 'Water is H<sub>2</sub>O'

Think about a glass of water. In that glass, there is some water and there are a bunch of H<sub>2</sub>O molecules. Does this mean that there are two things in the glass: some water *and* a bunch of H<sub>2</sub>O molecules? No, there is only one thing in the glass since water is H<sub>2</sub>O.<sup>35</sup> Call that one thing F. F is both water and H<sub>2</sub>O, which is just to say that the concept 'water' and the concept 'H<sub>2</sub>O' both pick out F (i.e., F is the semantic value of both 'water' and 'H<sub>2</sub>O').

Even though 'water' and 'H<sub>2</sub>O' have the same referent, the thoughts 'water is water' and 'water is H<sub>2</sub>O' seem different. In chemistry class, I might learn about H<sub>2</sub>O without ever learning that water is H<sub>2</sub>O. So I might believe that water is water and H<sub>2</sub>O is H<sub>2</sub>O without

---

<sup>35</sup> More properly, the property of being water is the property of being composed of H<sub>2</sub>O molecules. For the sake of simplicity, I will stick with the simpler 'water is H<sub>2</sub>O'.

believing that water is H<sub>2</sub>O; and in having and not having these beliefs, I might be perfectly rational. This suggests that 'water' and 'H<sub>2</sub>O' are individuated by more than just their semantic value. Call this additional element individuating 'water' and 'H<sub>2</sub>O' a **mode of presentation** or **MOP**, for short.

MOPs can be thought of as descriptions, theories, phenomenal appearances (or images), etc., of items *associated with* concepts of those items. To figure out what an MOP associated with a concept is, you need only think about the *way or guise under which* you think about the referent of the concept when you deploy the concept. Alternatively, you can think about the subjective application conditions of the concept. For example, when I think about water, I think about an item having a certain look or appearance. Moreover, I tend to call some item 'water' only if it has that look or appearance. Call that look the 'water-look'. The water-look is an appearance of water that is plausibly part of the MOP of water associated with my concept of water ('water'). More generally, the MOP of water associated with 'water' likely consists of manifest descriptions of water (e.g., 'is in the oceans', 'is in plastic bottles', etc.) and manifest appearances of water (e.g., the look of water, the taste of water, etc.). Call the MOP of water associated with 'water' the **manifest description<sup>36</sup> of water**.

Note that because, e.g., what water looks like (i) changes depending on the surrounding environment, and (ii) depends on properties of my visual system, the water-look is *contingent* and *extrinsic* to water. More generally, nearly all of the descriptions, appearances, etc., making up the manifest description of water are either contingent or extrinsic to water. This entails

---

<sup>36</sup> Clearly, since manifest appearances are included in the MOP, my use of the term 'description' here is fairly loose. The manifest "description" of water includes things other than descriptions, e.g., appearances or images.

that the manifest description of water is a **contingent description of water**. So it is possible for something to *satisfy* the manifest description of water—e.g., to look like water, to be in bottles, etc.—and *not be* water; and it is possible for something to *not satisfy* the manifest description of water—e.g., to not look like water, to not be in bottles, etc.—and *be* water.

How about the MOP associated with 'H<sub>2</sub>O'? What is it and is it contingent too? Unlike 'water', 'H<sub>2</sub>O' is a scientific concept coming out of and embedded within a scientific theory. Its associated MOP, then, will likely be fixed by the theory, and, for the purposes of this paper, we can think of its associated MOP as being a scientific description of some kind (something like 'is the molecule that consists of two hydrogen atoms and one oxygen atom in a bent molecular configuration, has a dipole moment of 1.85D, etc.'). Call this MOP of H<sub>2</sub>O associated with 'H<sub>2</sub>O' the **scientific description of water**.

Note that, if water is H<sub>2</sub>O, then the scientific description of water is not a contingent description of water (= H<sub>2</sub>O). If something has two hydrogen atoms and one oxygen atom, then that item is H<sub>2</sub>O (= water). So the scientific description of water is a **non-contingent description of water**.

The fact that 'water' is associated with the manifest description of water and 'H<sub>2</sub>O' is associated with the scientific description of water helps explain why, despite 'water' and 'H<sub>2</sub>O' having the same semantic value, I can believe that water is water and H<sub>2</sub>O is H<sub>2</sub>O without believing that water is H<sub>2</sub>O and still be perfectly rational. The reason I am rational is that the manifest description of water and the scientific description of water are inferentially isolated, i.e., it is not *a priori* that they are descriptions of the same thing. Thus, if water is H<sub>2</sub>O, then

'water is H<sub>2</sub>O' is **necessary** (since 'water' and 'H<sub>2</sub>O' have the same semantic value) and **a posteriori**.

## 2.2 'Pain is B'

Physicalists claim that mental states and brain states are related in just the same way that water and H<sub>2</sub>O are related. Let's stipulate that mental states and brain states are nomologically correlated with each other. So, in the case of pain, there is some brain state **B** such that: S has pain (or is in pain) iff S has B (or is in B) (S is an arbitrary subject of experience). Note that S has pain iff S has B resembles the biconditional 'G has water iff G has H<sub>2</sub>O' (G is a glass). In the case of 'G has water iff G has H<sub>2</sub>O', we said that, if G has water and G has H<sub>2</sub>O, it isn't the case that G has two items. G has only one item since water is H<sub>2</sub>O. The physicalist says that something similar is true for 'S has pain iff S has B'. When S has pain and S has B, S doesn't have two items. S has only one item since pain is B.

Given what we said in the previous section about 'water is H<sub>2</sub>O', we know that 'pain is B', if true, implies the following: (i) 'pain is B' is necessary, i.e., 'pain' and 'B' have the same semantic value; and (ii) 'pain is B' is *a posteriori*, i.e., the MOP associated with 'pain' and the MOP associated with 'B' are inferentially isolated. In order to evaluate 'pain is B' then, we have to examine the concept 'pain' and the concept 'B'. Let's start with 'B'.

Like 'H<sub>2</sub>O', 'B' is a scientific concept coming out of and embedded within a scientific theory. Its associated MOP, then, will be fixed by the theory. For the purposes of this paper, we can think of its MOP as being a complicated scientific description of ions and molecules moving and interacting in lawlike ways. Call this MOP associated with 'B' the **scientific description of pain**.

What about 'pain'? What is its associated MOP? The concept 'pain' is more difficult to discuss than previous concepts since 'pain' is a special kind of concept. 'Pain' is what is known as a **phenomenal concept**, and in order to understand 'pain', we have to consider some basic matters about pain.

### 2.2.1 Pain, the Phenomenal Concept 'Pain', and the Constitutional Theory

Suppose my doctor asks me: 'Are you in pain?' What do I do? I introspect and "search for" or "look for" a sensation (or feeling) of some kind. When I say "search for" or "look for" I mean that I check to see if I'm aware of the sensation. If I am aware of it, then I say 'I am in pain' or 'I have pain'. If I am not aware of it, then I say 'I am not in pain' or 'I don't have pain'. In fact, something stronger than this seems true. It is hard to imagine a case where I could be aware of the sensation but not be in pain; and it is hard to imagine a case where I could be in pain but not be aware of the sensation. This suggests that the concept 'pain' tracks the sensation across conceivable scenarios; and this, in turn, suggests that 'pain' picks out that sensation. So the referent of 'pain' is a sensation or feeling of some kind. Pain, therefore, is a sensation or feeling.<sup>37</sup>

To say that pain is a sensation (or feeling) is important because sensations have several salient characteristics. I will highlight four. First, sensations are *necessarily had*. That is, it's hard to make sense of there being free-floating or unbound sensations. Second, sensations are *subjective*. This just extends the point made earlier: it's hard to imagine someone having a sensation but not being aware of it, and it's hard to imagine someone being aware of a

---

<sup>37</sup> See Lewis (1980): "Pain is a feeling. Surely that is uncontroversial. To have pain and to feel pain are one and the same" (222).

sensation but not having it. Sensations occur iff subjects are aware of them. Third, sensations are *private*. That is, it's hard to imagine someone else being aware of a sensation you are having in just the same way that you are aware of the sensation. Finally, sensations are *had rather than experienced*. What do I mean by this? Well, we don't have experiences of our experiences; we just *have* experiences. So we don't have perceptions of our perceptions, or sensations of our sensations, or feelings of our feelings. We just have perceptions, sensations, and feelings. Sensations, then, are things we *have* rather than *experience*.

Let's say that  $x$  is **internal** iff  $x$  is subjective, private, necessarily had, and had not experienced. Since pain is a sensation, and since sensations are internal, pain is internal too.<sup>38</sup> The internality of pain entails that it is a mental state. This is because non-mental or **external** items do not exhibit, e.g., subjectivity. You could have a cut on your hand but not be aware of it (if, for example, your hand is anesthetized); but you couldn't have a pain you are not aware of. This not only tells us that the truth conditions for 'S has a pain' and 'S has a cut' are different (thus, pains aren't cuts even though they can be caused by cuts); but it also tells us that cuts are external (to the mind) whereas pains are internal (to the mind).

So pain is a state that your mind can be in—it is a **mental state**—or, alternatively, it is an event that occurs in your mind—it is a **mental event** (I will stick with calling pain a 'mental

---

<sup>38</sup> It is true that we sometimes say: 'I am having an experience of pain', 'I am having a sensation of pain', 'I am sensing pain', 'I am having a feeling of pain', and 'I am feeling pain'. If pain is a sensation or feeling, then these statements can't be taken at face value. And, in fact, they *shouldn't* be taken at face value since all of them are equivalent to: 'I am in pain' and 'I have pain' (see Lewis's comment in note 37). As it turns out, all of the ways of talking about pain other than 'I am in pain' or 'I have pain' are infelicitous. Why do we talk about pain in these infelicitous ways? I address this question in "Puzzles of Pain", but, briefly, the reason is this: we talk in these infelicitous ways about pain because (i) experiences come in (at least) two distinct types: transparent experiences and non-transparent experiences; and (ii) for reasons of economy and reproductive fitness, the main linguistic framework we employ to talk about experiences is apt only for transparent experiences. Non-transparent experiences (like pain) are shoe-horned into this transparent linguistic framework, and this leads to infelicities when we talk about non-transparent experiences.



state'). In addition, because there is something it is like to have (or be in) pain, pain is a special kind of mental state: it is an experience.<sup>39</sup>

Because pain is an experience, and because I have had pain before, my concept of pain—'pain'—is a **phenomenal concept**. Phenomenal concepts play an important role in the hard problem since they are the concepts we use to think about our experiences subjectively or from a first-person point of view (from the inside, so to speak). To see why phenomenal concepts are important, consider how I think about pain using my phenomenal concept 'pain'.

What do I think about when I think about pain? Three answers seem possible. The first and most obvious is: I think about a feeling of some kind. But we might also say: I think about what it *feels* like to have pain. And: I think about what pain *feels* like. It might seem that these are three different answers to the same question, but, in fact, they are the same answer. This is because, when I think about what it *feels* like to have pain, I think about a *feeling* of some kind; and when I think about what pain *feels* like, I also think about a *feeling* of some kind. What *feeling* do I think about when I think about these two things? I think about the very same *feeling* mentioned in the first answer. So all three answers amount to the same thing: when I think about pain using 'pain' I think about a feeling.

---

<sup>39</sup> Some philosophers, e.g., Harman (1990, pp. 39-40), have claimed that pains cannot be mental states because pains are located in (say) toes and mental states are located only in the head. More generally, the seeming location of pain has seemed to some in tension with pain's mentality. I discuss why there is no tension here in "Puzzles of Pain", but, briefly, the reason is this: statements like 'there is a pain in my toe' and 'I have a pain in my toe' are self-attributions of a token of a type of experience (see, e.g., Armstrong (1968)). The type of experience is *pain in my toe*. Thus, when I say, 'I have a pain in my toe', I am saying that I am having a certain type of experience, one denoted by the expression 'pain in my toe'. And when I say 'there is a pain in my toe', I am making the existential claim that there exists a token of the type of experience *pain in my toe*, and I am having that token experience. In neither case am I saying 'a pain is *located* in my toe' since I would still have a pain in my toe even if you didn't find a pain *located* in my toe. So, pains aren't *located* in body parts.

Of course, this feeling that I think about is pain (pain, remember, is a sensation or feeling). This suggests that, when I think about pain, I think about it directly in terms of itself (i.e., in terms of the feeling that is pain). If that is right, then the MOP of pain associated with ‘pain’ *is* pain itself. Call this MOP the **phenomenal description of pain**. According to what is said above, the phenomenal description of pain is pain.

This is starkly different from the manifest description of water. The manifest description of water did not *include* water. The manifest description included things like the *appearance* of water, but the appearance of water is distinct from, contingent to, and extrinsic to water. This is part of the reason why the manifest description is a contingent description.

If the manifest description of water includes the appearance of water, then, presumably, the phenomenal description of pain includes the appearance of pain too. What is the appearance of pain? Well, if the appearance of water is what water looks like (the look of water), then the appearance of pain is what pain *feels* like (the *feel* of pain). But, according to what we said earlier, what pain *feels* like is a *feeling* of some kind, specifically, the *feeling* that is *pain*. So the appearance of pain *is* pain—pain is not distinct from its appearance—and because the phenomenal description of pain includes the appearance of pain, this implies that the phenomenal description of pain is pain (which conforms with what we said earlier).<sup>40, 41</sup>

There are two important consequences of pain being the phenomenal description of pain. First, unlike the case with the manifest description of water, the phenomenal description

---

<sup>40</sup> It is perhaps better to say that it *includes* pain rather than it *is* pain. However, for the purposes of this paper, this detail won’t matter. What is important is that pain is a constituent of the phenomenal description of pain.

<sup>41</sup> The idea that pain is its own appearance is sometimes expressed by saying that pain lacks an ‘appearance-reality distinction’. See Nagel (1974, p. 444), Kripke (1980, pp. 152, 154), and Searle (1994, pp. 121-122) for discussion of pain—and other sensations—lacking an appearance-reality distinction.

of pain is a *non-contingent* description of pain. This is because anything that satisfies the description—anything *feeling* like or having the appearance of pain—is pain.<sup>42</sup> Second, in order for S to possess the phenomenal concept ‘pain’, S must have had pain before. This is because the phenomenal description of pain is pain. So, if S has never had pain before, S cannot think about pain in the same way that I can think about pain. I can think about pain directly in terms of the feeling that is pain (or, equivalently, I can think about what it *feels* like to have pain and what pain *feels* like) because I have had pain before and so have ‘pain’. But because S has never had pain before and does not have ‘pain’, S cannot think about pain in the ways that I can.

Why does ‘pain’ have these distinguishing features? In recent years, the **constitutional theory of phenomenal concepts** has emerged as a favored account of ‘pain’'s distinguishing features. Proponents of the constitutional theory include both dualists and physicalists<sup>43</sup>, and at the heart of the constitutional theory is the idea that pain is a constituent of ‘pain’. Katalin Balog (2012a) describes the constitutional theory as follows:

[I]n the case of phenomenal concepts, e.g., the concept PAIN, constitution matters for reference, both in terms of reference fixing, and in terms of how the concept cognitively presents its referent [i.e., its associated MOP]. More precisely, on this view, every token of a phenomenal concept applied to current experience is (partly) constituted by *that token experience* . . . (7; italics in the original)

Ned Block (2007) says: “We could take the form of a phenomenal CMoP [i.e., a phenomenal concept] to be ‘the experience: \_\_\_\_\_,’ where the blank is filled by a phenomenal property,” (264) and “[i]n . . . fundamental uses a simultaneous actually occurring experience is used to

---

<sup>42</sup> For example, I couldn’t have an itch that feels like (or appears as) a pain since if what I have feels like a pain, then I have (or am in) pain.

<sup>43</sup> See, e.g., Hill and McLaughlin (1999), Chalmers (2003), Block (2007), Papineau (2007, 2011), and Balog (2009, 2012a, 2012b).

think about that very experience” (252). Similarly, David Papineau (2007) says, “phenomenal concepts have the very peculiar feature of *using* the experiences they refer to” (131; italics in the original).

The constitutional theory defended by Balog, Block, Papineau and others could be summarized as follows. According to the constitutional theory: (i) the phenomenal concept ‘pain’ refers to a sensation (or feeling); (ii) that sensation (pain) is an experience; (iii) that sensation (pain) is the phenomenal description of pain (the MOP associated with ‘pain’); and (iv) in order to possess ‘pain’, you must have had that sensation (pain) before. These claims state nothing that we haven’t already said. What is novel about the constitutional theory though is how it explains (i)-(iv): it explains it by positing that pain (the sensation) is a constituent of ‘pain’ (the phenomenal concept).<sup>44</sup> That, according to the constitutional theory, is why the phenomenal description of pain is pain and you must have had pain before in order to have ‘pain’.

Having now discussed pain, the phenomenal concept ‘pain’, and the constitutional theory, let’s return to ‘pain is B’.

### 2.3 Back to ‘Pain is B’

If ‘pain is B’ is true, then two things follow. First, ‘pain is B’ is necessary: ‘pain’ and ‘B’ have the same semantic value. Because pain (the referent of ‘pain’) is an experience, this

---

<sup>44</sup> More accurately, proponents of the constitutional theory distinguish between direct phenomenal concepts and indirect phenomenal concepts. Direct phenomenal concepts are phenomenal concepts that are applied to occurrent experiences. Tokens of direct phenomenal concepts are constituted by the tokens of the occurrent experiences they refer to. Indirect phenomenal concepts refer to past, future, or imagined experiences. Accounting for indirect phenomenal concepts is a bit more complicated than accounting for direct phenomenal concepts, but most theorists think that indirect phenomenal concepts are constituted by “faint images” of experiences. I will set this subtlety aside in the paper as it is peripheral to my argument. For more on direct and indirect phenomenal concepts, see Chalmers (2003), Balog (2009), and Sundström (2011).

implies that 'B' refers to an experience (specifically, the sensation or feeling that is pain).

Second, 'pain is B' is *a posteriori*: the phenomenal description of pain and the scientific description of pain are inferentially isolated. Because the scientific description of pain (the MOP associated with 'B') is a description of ions and molecules moving and interacting, and because the phenomenal description of pain is pain, these two descriptions are inferentially isolated.

Now, in the case of 'water is H<sub>2</sub>O', we said that, though the manifest description of water is a contingent description of water, the scientific description of water is a non-contingent description of water (= H<sub>2</sub>O). Similarly, if pain is B, then the scientific description of pain will be a non-contingent description of pain (= B) (since the scientific description of pain is a non-contingent description of B). But we know that the phenomenal description of pain is *not* a contingent description of pain: it is a non-contingent description of pain. And it is this fact about pain and its phenomenal description that gives rise to the hard problem.

### 3. The Hard Problem and the Phenomenal Concept Strategy

Why is it problematic for 'pain is B' that both the phenomenal description of pain and the scientific description of pain are non-contingent descriptions of pain? Well, consider Kripke's (1980) conceivability argument against 'pain is B' (144-155). According to Kripke, I can imagine being in pain but not being in B. This suggests that it is possible for me to be in pain but not be in B. If that is right, then pain can't be B since they are contingently related (i.e., 'pain is B' is not necessary).

But can't a similar argument be given against 'water is H<sub>2</sub>O'? After all, I can imagine some substance being water but not being H<sub>2</sub>O. This suggests that it is possible for something to be water but not be H<sub>2</sub>O. If that is right, then water can't be H<sub>2</sub>O since they are contingently

related. Kripke, however, thinks there is a problem with this argument. When I imagine some substance being water but not being H<sub>2</sub>O, I imagine some substance satisfying the manifest description of water—e.g., it looks like water—but not satisfying the scientific description of water. But this really isn't a case where water is not H<sub>2</sub>O since the manifest description of water is a contingent description of water. Just because something satisfies the manifest description doesn't mean that it is water, and this entails that I haven't succeeded in imagining something to be water but not be H<sub>2</sub>O. The appearance of contingency between water and H<sub>2</sub>O can be explained away.

Is the same true for pain and B? Well, when I imagine being in pain but not being in B, I imagine being in a state that satisfies the phenomenal description of pain (it feels like pain) but not the scientific description of pain. However, in this case, because the phenomenal description of pain is a non-contingent description of pain (since it is pain), anything that satisfies it (anything that feels like pain) is pain. So if I am in a state that satisfies it, then that state *is* pain. Thus, when I imagine being in pain but not being in B, I *really* am imagining this. And this suggests that the appearance of contingency between pain and B can't be explained away. Kripke concludes that pain is not B because 'pain is B' is contingent.

Clearly, the problem in the 'pain is B' case is that the phenomenal description of pain is a non-contingent description of pain. And, in one way or another, all of the arguments that go into the hard problem—the conceivability argument, the knowledge argument, and the explanatory gap argument—trade on this fact about the phenomenal description of pain.

The **phenomenal concept strategy** (the **PCS**) is a strategy for solving (or at least defusing) the hard problem. The PCS claims that, by appealing to the constitutional theory, it

can rebut the three arguments making up the hard problem. How exactly does the strategy work? Consider again Kripke's conceivability argument. What exactly has Kripke's argument shown? The only thing that Kripke has shown is that you cannot explain away the appearance of contingency between pain and B. Why can't you do that? Kripke's answer is 'because pain is not B'.<sup>45</sup> But this is too quick. The real reason you cannot explain away the appearance of contingency between pain and B is because the phenomenal description of pain is a non-contingent description of pain. If the phenomenal description were a contingent description—just as the manifest description of water is a contingent description of water—then you could explain away the appearance of contingency between pain and B. But because the phenomenal description is a non-contingent description, you cannot.

The key question then is: *why* is the phenomenal description of pain a non-contingent description of pain? Again, Kripke says that it is because pain is not B but this is too quick. The real reason the phenomenal description is a non-contingent description is given to us by the constitutional theory: it is because pain is a constituent of 'pain'. *That* is why the phenomenal description is a non-contingent description since it entails that the phenomenal description is pain.

In light of this, the PCS, by appealing to constitutional theory, stands poised to solve the hard problem. Here is how Balog (2012a) describes the solution:

[V]ia phenomenal concepts we are acquainted with phenomenal properties: we grasp them directly and in a way that appears to reveal their essence. The constitutional account explains this by suggesting that an experience instantiating the phenomenal property *q* serves as mode of presentation of a particular token of the phenomenal concept *C<sub>q</sub>* by *being constitutive of the relevant token of C<sub>q</sub>*. There is a delicate issue regarding the nature of acquaintance here. If phenomenal properties are, as the

---

<sup>45</sup> More accurately: because it is *possible* for you to be in pain but not be in B, so 'pain is B' is not necessary.

physicalist claims, physical or functional properties, then there is a clear sense in which acquaintance *doesn't* reveal their nature . . . [I]t would be better for the physicalist to analyze acquaintance and the substantiality of phenomenal belief in terms of the phenomenal *presence* of the introspected properties in phenomenal judgments; and not in terms of our direct grasp of the *essence* of phenomenal properties. (14-15; italics in the original)

In another paper (2012b), she says:

If phenomenal concepts are partly constituted by phenomenal states, our knowledge of the presence of these states (in the first-person, subjective way of thinking of them) is not mediated by something distinct from these states. Rather the state itself serves as its own mode of presentation . . . Is this a problem for physicalism? You can see why not by focusing on . . . [t]he important point that this kind of direct insight (via shared phenomenality of thought and experience) does not reveal anything about the metaphysical nature of phenomenality. (30-31)

Claims similar to these are made by other proponents of the PCS, e.g., Papineau (2007):

[A]ntiphysicalists . . . conclude that pain must refer to something nonphysical, something with which the possessors of the concept are indeed directly acquainted. But type-B physicalists can respond that, however it is with other concepts, this combination of semantic stability and ignorance of essence is just what we should expect, given the use-mention feature characteristic of phenomenal concepts. Even if phenomenal concepts don't involve direct knowledge of real essences, they will still come out semantically stable, for the simple reason that the use-mention feature leads us to think of the referent as "built into" the concept itself . . . If this is right, then the semantic stability of phenomenal concepts provides no reason to think that they must refer to nonphysical properties with which their possessors are directly acquainted. For the use-mention feature of phenomenal concepts yields an independent explanation of why they should be semantically stable, even while their possessors remain ignorant of the real physical essences of their referents. (131-132)

How exactly do these comments yield a solution to the hard problem? The basic idea behind the PCS is this. The constitutional theory (Balog: 'pain' is "partly constituted" by pain; Papineau: 'pain' has a "use-mention feature") explains why the phenomenal description of pain is a non-contingent description of pain (Balog: pain "serves as its own mode of presentation"; Papineau: 'pain' is "semantically stable"). It, therefore, explains why we can't explain away the appearance of contingency between pain and B.



Is the constitutional theory *incompatible* with pain being B? No, say the proponents of the PCS. The constitutional theory does not imply that pain is not B. It only implies that (in Balog's words): we grasp pain "directly"; our awareness of pain "is not mediated by something distinct from these states"; we have a "direct insight (via shared phenomenality of thought and experience)" into pain; and pain has a "phenomenal presence" in 'pain' and thoughts about pain. What does all of this mean? Here is how I interpret these consequences of the constitutional theory.

Because pain is an experience, and because we don't have experiences of our experiences (e.g., sensations of our sensations), our awareness of pain (when we have it) isn't mediated by a distinct experience of pain. We simply have pain and, by virtue of that, are aware of it.<sup>46</sup> We can therefore say that we are **directly experientially aware of pain**. Moreover, because pain is a constituent of 'pain', even our conceptual awareness of pain isn't mediated by a wholly distinct mental entity. We can therefore say that we are **directly conceptually aware of pain**.<sup>47</sup> If we use the term 'directly aware' to mean 'directly experientially aware AND directly conceptually aware', then we can say:

**(DA)** When we have pain, we are **directly aware** of pain

---

<sup>46</sup> This explains the subjectivity of pain, e.g., why it is hard to imagine S having pain (or being in pain) but not being aware of the pain.

<sup>47</sup> This explains incorrigibility with respect to pain, e.g., why it hard to imagine S having an occurrent pain (or occurrently being in pain) yet sincerely believing 'I am having an itch'. Incorrigibility with respect to pain seems to be rooted in the nested, constitutional relation between (i) pain and the concept 'pain', and (ii) the concept 'pain' and the thought 'I am having pain'. See Chalmers (2003) and Balog (2012b) for more on the nested, constitutional relation between experiences, phenomenal concepts, and phenomenal thoughts, and how this relation leads to incorrigibility.

The constitutional theory, then, implies (DA). But is (DA), in any way, problematic for pain being B? In particular, does our direct awareness of pain give us any insight into pain that allows us to rule out pain being B (or any other brain state)? No, say the PCS proponents. Just as S's awareness of water does not give S any insight into water—or allow S to know anything about water—that allows them to rule out water being H<sub>2</sub>O, my direct awareness of pain does not give me any insight into pain—or allow me to know anything about pain—that allows me to rule out pain being B (Balog: my direct awareness of pain “does not reveal anything about the metaphysical nature of” pain; Papineau: my direct awareness of pain does not give me “direct knowledge of [the] real essence” of pain). If that is right, then the PCS can rebut, not just Kripke's conceivability argument, but all of the arguments making up the hard problem. This is because all of these arguments trade, in one way or another, on the phenomenal description of pain being a non-contingent description of pain. But the PCS, by appealing to the constitutional theory, can explain why the phenomenal description is a non-contingent description; and **provided that (DA) leads to no problem for 'pain is B'**, the PCS's explanation will be compatible with pain being B. The PCS, then, can resolve the hard problem by rebutting the antiphysicist conclusion of the three arguments.

#### 4. The Problem with the PCS

Whether or not the PCS succeeds depends on (DA) not being problematic for 'pain is B'. In particular, the PCS hinges on it being true that my direct awareness of pain doesn't allow me to know things about pain that are at odds with pain being B. I will argue, however, that this isn't the case. By virtue of being directly aware of pain when I have it, I can learn things about pain that aren't consistent with pain being B.

How will I argue for this? Broadly speaking, my argument is divided into two parts. First, I will describe a test case where a subject S's direct awareness of an item x puts them in a position to know something about x that rules out x being B. Second, I will argue that, with respect to me and pain, my situation is epistemically indistinguishable from test case. This, I claim, allows me to know that pain is not B by virtue of being directly aware of pain.

#### 4.1 The Test Case: Eden, Perfect Red, and Perfect Pain

The test case I will describe is inspired by David Chalmers's (2006) discussion in "Perception and the Fall from Eden" (all forthcoming quotes from Chalmers are from this article). Eden, as Chalmers uses the term, has nothing to do with the biblical Eden. Instead, it is a possible world having special features. The best way to understand what Eden is is to think about color on Eden. Here is how Chalmers describes color on Eden:

[On Eden] properties were revealed to us in their true intrinsic glory. When an apple in Eden looked red to us, the apple was gloriously, perfectly, and primitively red . . . [T]he perfect redness of the apple was simply revealed to us. The qualitative redness in our experience derived entirely from the presentation of perfect redness in the world. Eden was a world of perfect color. (49)

On Eden, something called 'perfect red' is instantiated, and perfect red is *revealed* to the subjects on Eden. What is perfect red? Well, there are three crucial facts we need to know about Eden. First, objects and properties on Eden are revealed to the subjects on Eden in the sense that these subjects know what these objects and properties are.<sup>48</sup> Second, there is, on Eden, a *perfect match* between the phenomenology of experience (how things phenomenally seem) and the way that things are (how things are) (Chalmers: "[t]he qualitative redness in our

---

<sup>48</sup> Chalmers says that "in Eden, an especially strong perceptual relation obtains, one we might call *perfect perception*" (94; italics in the original). Perfect perception "reveal[s] the intrinsic nature of things" (95).

experience derive[s] entirely from the presentation of perfect redness in the world”). So, on Eden, how things are matches how they seem.

Finally, how things seem on Eden matches how things seem here on Earth:

Phenomenologically, it seems that visual experience presents the world to us as an Edenic world. Taking the phenomenology completely at face value, visual experience presents a world where perfect redness and perfect blueness are instantiated on the surface of objects, as they were in Eden . . . For the world to be *exactly* the way that my phenomenology seems to present it as being, the world would have to be an Edenic world in which these properties are instantiated. (66; italics in the original)

So, from a phenomenological perspective, there is a *perfect match* between how things seem on Eden and how things seem on Earth.

In light of this, if we want to answer the question, ‘what is perfect red?’, we first have to figure out how red seems to us here on Earth. We will then know how perfect red seems on Eden, and knowing this will allow us to know what perfect red is (since perfect red will match how it seems). So: how does red seem to us here on Earth?

Suppose that I am looking at a patch of red on a white wall and I visually focus on the color I am seeing. How does that color seem? The first thing we can say is that red seems to have a *distinct qualitative and chromatic nature*. What do I mean by this? It is difficult to describe—using non-color vocabulary—what distinct qualitative and chromatic nature red seems to have. In fact, it is difficult to describe what having a distinct qualitative and chromatic nature *even means* using non-color terms. In order for you to know what I am talking about here, you have to have seen a color yourself. And what this suggests is that colors seem *sui generis*, i.e., they seem to be *sui generis* properties. Colors seem to have qualitative and chromatic natures that mark them off from other *kinds* of properties, and these natures can

only be described using color vocabulary. Non-color terms won't do, and this suggests that colors—e.g., red—seem *sui generis*.

In fact, something stronger than this seems true. Alex Byrne and David Hilbert (2007) claim that, merely by reflecting on color experience, we would think: “[C]olors are *not* reflectance types . . . [and] colors are not dispositions to affect perceivers, microphysical properties, chemical properties, ectoplasmic properties, and so on” (77; italics in the original). This is because: “The only nature or essence of the colors apparent from reflection on color experience is chromatic through and through” (77). Similarly, Chalmers claims:

When I have a phenomenally red experience of an object, the object seems to be simply, primitively, *red*. The apparent redness does not seem to be a microphysical property, or a mental property, or a disposition, or an unspecified property that plays an appropriate causal role. Rather, it seems to be a simple qualitative property, with a distinctive sensuous nature. (66; italics in the original)

Byrne, Hilbert, and Chalmers are each claiming that red doesn't just seem to have a distinct qualitative and chromatic nature: it seems to be qualitative and chromatic through and through. That is, red doesn't seem to have any non-qualitative or non-chromatic components; it seems to be *exhausted* by its qualitative and chromatic nature. Red then, in addition to seeming *sui generis*, seems **primitive** as well.

Finally, colors (like red), unlike sensations (like pain), are notable for seeming objective and public. Chalmers, for instance, says: “[T]he phenomenology of pain is quite different from the phenomenology of color, where we have no trouble conceiving of an object being perfectly colored even though no one ever experiences its color” (113-114). And Ned Block (2005) says: “[W]e have to recognize that colors are objective in a way that Subjective Qualities [i.e., sensations] are not. There is an appearance–reality distinction for red but not for a Subjective

Quality such as achiness . . . [T]here can be unseen red but not unfelt achiness” (140-141). The thought here is that, because I experience red as inhering in external objects (i.e., objects that are distinct from and external to my mind), red seems to be an **objective and public property**. I can imagine red being instantiated even if no one sees or is aware of it; and I can imagine others seeing and being aware of the very same instance of red that I see and am aware of. If we use the term ‘external’ to denote ‘objective and public’, then we can say that red seems to be an **external** item.

Altogether, this yields: red seems to be a primitive, *sui generis*, and external item having a distinct qualitative and chromatic nature.

As I mentioned earlier, specifying the distinct qualitative and chromatic nature that red seems to have isn’t easy. Evidently, the only way to do it is to look at an instance of red and think ‘*that* is the distinct qualitative and chromatic nature that red seems to have’.<sup>49</sup> Thus, the distinct qualitative and chromatic nature that red seems to have is ‘*that* nature’. This gives: red seems to be a primitive, *sui generis*, and external item having *that* nature.

Now that we know how red seems, we can say what perfect red is. Since everything on Eden is exactly as it seems, and since how things seem on Eden matches how things seem on Earth, perfect red is red *exactly as it seems*. Thus, perfect red is a primitive, *sui generis*, and external item having *that* nature.

---

<sup>49</sup> *That*, deployed in this thought, is a mental demonstrative. Mental demonstration is a capacity we have to single out or highlight items in our experiential field by thinking *that* (or *this*). To mentally demonstrate *x*, I must be directly aware of *x*, I must mentally attend to (or focus on) *x*, and I must pick *x* out by thinking *that* of it. See Levine (2010) for more on mental demonstration.

Perfect red isn't quite the test case I want to discuss. The test case concerns another perfect item on Eden: **perfect pain**. The discussion of perfect red, however, was necessary since it helps us understand what Eden is (and I will return to perfect red later in the paper). The question we now want to ask is: what is perfect pain?<sup>50</sup> Fortunately, what we said for perfect red applies to perfect pain. In order to figure out what perfect pain is, we first have to figure out how pain seems. Suppose that I have pain and I focus my attention on the pain that I am having. How does that pain seem?

The first thing we can say is that pain seems to have a distinct qualitative and (this time) *sensory* nature. For reasons paralleling those in the case of red, this entails that pain seems (and the other sensations seem) *sui generis*. Also, because pain seems qualitative and sensory through and through, pain seems primitive. In these respects, pain and red seem similar. But an important difference exists with respect to externality. Unlike red, pain doesn't seem external. Pain, remember, is a sensation (or feeling), and sensations (according to what was said earlier) seem internal (i.e., subjective, private, necessarily had, and had not experienced). So pain seems to be a primitive, *sui generis*, and internal item having a distinct qualitative and sensory nature; and, again, for reasons paralleling those expressed for red, the distinct qualitative and sensory nature that pain seems to have can be specified only by having pain, focusing on it, and

---

<sup>50</sup> In what follows, I depart from Chalmers in my understanding of perfect pain. Chalmers doesn't say much about perfect pain in his paper, and when he discusses it, it is clear that he is attempting to explicate perfect pain in a way that *exactly* mirrors perfect red. That project fails as he concludes: "[P]erfect pain cannot be instantiated: there is no possible world in which there is perfect pain, and on reflection it is not even conceivable that there is perfect pain. In effect, the instantiation of perfect pain places incoherent requirements on the world" (114). I think that the mistake Chalmers makes is that he fails to recognize that perfect pain and perfect red are fundamentally different kinds of things. In particular, perfect pain is an *experience* whereas perfect red is an *external* property that is an *object* of experience.

thinking ‘*this* is the distinct qualitative and sensory nature that pain seems to have’. Thus, pain seems to be a primitive, *sui generis*, and internal item having *this* nature.

Just as perfect red is red exactly as it seems to be, perfect pain is pain exactly as it seems to be. So perfect pain is a primitive, *sui generis*, and internal item having *this* nature. Note that this description of perfect pain delimits what perfect pain could be. Consider, for example, whether perfect pain could be a brain state—say, brain state B. B, recall, is described by the MOP associated with ‘B’—the scientific description of pain—and this description is a complicated scientific description of ions and molecules moving and interacting in lawlike ways. B, evidently, is neither primitive nor *sui generis*, so if perfect pain is a primitive, *sui generis*, etc., then it couldn’t be B. In fact, perfect pain couldn’t be *any item* specified in terms different than ‘is a primitive, *sui generis*, and internal item having *this* nature’.<sup>51</sup>

Let’s now discuss the test case. The test case, remember, is an instance where S’s direct awareness of an item x puts them in a position to know something about x that rules out x being B. Let’s say that S is a subject on Eden who is cognitively just like us, and x is perfect pain. Perfect pain, like pain, is an experience; and given the phenomenological correspondence between Eden and Earth, when we have pain, S has an experience that corresponds to our having pain, and that experience is perfect pain. So when we have pain, S has perfect pain.

Because S is on Eden, S knows what perfect pain is. In particular, S knows that perfect pain is a primitive, *sui generis*, and internal item having *this* nature. But why exactly does S

---

<sup>51</sup> See, e.g., Byrne and Hilbert (2007): “According to Minimal Primitivism, the colors are *sui generis* properties, which we earlier expressed by saying that the colors are not identical with properties specified in other terms” (78). One caveat to this: when we say that a primitive, *sui generis* property F can’t be identified with a property G specified in other terms, we are assuming that G is specified by a non-contingent, intrinsic description that differs from the primitive, *sui generis* description of F. This assumption is necessary since, even if F is primitive and *sui generis*, it could always be identified with a property specified by a contingent and/or extrinsic description.



know this? In other words, even if we grant that S knows this, we still need to explain why S knows this about perfect pain on Eden. And the question I am now asking is: what best explains S's knowledge that perfect pain is a primitive, *sui generis*, etc.?

In my view, S knows that perfect pain is a primitive, *sui generis*, and internal item having *this* nature because: (i) perfect pain seems to be a primitive, *sui generis*, and internal item having *this* nature; (ii) S can think 'perfect pain is a primitive, *sui generis*, and internal item having *this* nature'; and (iii) S has no defeaters to this belief (i.e., S has no reasons to think that it is not true—this is how I am using the term 'defeater'). The reason I think that (i)-(iii) are sufficient for S to know that perfect pain is a primitive, *sui generis*, etc., is that S doesn't know that they are Eden—i.e., S doesn't know that they are on a world where things are as they seem—so even if Eden plays an external role in S's knowledge that is not something that S is internally aware of. The only internal evidence that S has regarding what perfect pain is is how perfect pain seems. There is nothing else informing S about what perfect pain is.

If that is right, and if we grant (i) and (iii), then S's knowledge that perfect pain is a primitive, *sui generis*, etc., depends on S being able to think 'perfect pain is a primitive, *sui generis*, and internal item having *this* nature'. Why can S think this proposition? Well, S can think this proposition iff S has the relevant concepts in the proposition (e.g., 'primitive', '*sui generis*', 'internal', etc.), and S can mentally demonstrate perfect pain's distinct qualitative and sensory nature. Since S is cognitively similar to us, we can assume that S has the concepts in question. Moreover, since we have the powers of mental attention and demonstration, we can assume that S has these powers too. Note that, because perfect pain is primitive, it is exhausted

by its distinct qualitative and sensory nature. So mentally demonstrating that nature *just is* mentally demonstrating perfect pain.

In light of all this, we have the following: S can think 'perfect pain is a primitive, *sui generis*, etc.' iff S is directly aware of perfect pain.<sup>52</sup> Is S directly aware of perfect pain? As mentioned above, perfect pain, like pain, is an experience; and given that we don't have experiences of our experiences, S is not aware of perfect pain by perceiving, sensing, or feeling it. S has perfect pain and, by virtue of that, is aware of it. So, S is directly experientially aware of perfect pain. Moreover, the constitutional theory implies that S's concept of perfect pain is partly constituted by perfect pain. So even S's conceptual awareness of perfect pain isn't mediated by a wholly distinct mental entity: S is directly conceptually aware of perfect pain.

This implies that S is directly aware of perfect pain when S has perfect pain. As a result, S can think and come to know that perfect pain is a primitive, *sui generis*, and internal item having *this* nature. Knowing this allows S to know that perfect pain is not B.

In sum, S's knowledge that perfect pain is a primitive, *sui generis*, etc., is rooted in all of the following. **First**, perfect pain seems to be a primitive, *sui generis*, etc. **Second**, S can think 'perfect pain is a primitive, *sui generis*, etc.' Note that S can think this proposition because S has the concepts involved in it, S has the powers of mental attention and demonstration, and, *crucially*, S is directly aware of perfect pain when S has perfect pain. Direct awareness of perfect pain, then, plays an essential role in S knowing what perfect pain is. **Finally**, S has no defeaters to their belief (i.e., no reasons not to believe) that perfect pain is a primitive, *sui generis*, etc.

---

<sup>52</sup> Recall (see note 49) that mentally demonstrating *x* requires being directly aware of *x*, attending to *x*, and picking *x* out by thinking *this*. Since S has the powers of mental attention and demonstration, they can mentally demonstrate perfect pain iff they are directly aware of perfect pain.

S having perfect pain on Eden is the test case because S's direct awareness of perfect pain plays an essential role in allowing S to know something about perfect pain—that it is a primitive, *sui generis*, etc.—that is at odds with perfect pain being, e.g., a brain state. It is true that other conditions have to be met in order for S to know what perfect pain is. But that does not negate the role played by direct awareness, and, in the next section of the paper, we will see that direct awareness has another important role to play (at least, in the case of me having pain on Earth).

#### 4.2 Having Pain on Earth

I now want to argue that, if we consider (in isolation from the other occurrences on Earth and Eden) the state of affairs consisting of me having pain on Earth and the state of affairs consisting of S having perfect pain on Eden, the two situations are epistemically indistinguishable. Thus, if S knows that perfect pain is a primitive, *sui generis*, etc., I know that pain is a primitive, *sui generis*, etc. Put another way, I know that pain is perfect pain. In addition, if S knows that perfect pain is not B, then I know that pain is not B (since I know that pain is perfect pain and perfect pain is not B).

In order to argue that me having pain on Earth is epistemically indistinguishable from S having perfect pain on Eden, let's consider how the two states of affairs are similar.

**(1)** Both pain and perfect pain are experiences (sensations or feelings). **(2)** Both pain and perfect pain seem to be primitive, *sui generis*, and internal items having *this* nature. **(3)** My concept of pain is just like S's concept of perfect pain: pain is a constituent of my concept of pain, and perfect pain is a constituent of S's concept of perfect pain. **(4)** My awareness of pain is just like S's awareness of perfect pain: I am directly aware of pain when I have it, and S is

directly aware of perfect pain when they have it. **(5)** Both S and I have the concepts in the description ‘is a primitive, *sui generis*, and internal item having *this* nature’, and we both have the powers of mental attention and demonstration. **(6)** Both S and I can think ‘perfect pain/pain is a primitive, *sui generis*, and internal item having *this* nature’.

If (1)-(6) describe how the two states of affairs are epistemically similar, then how are they different? One possible difference is that S and I are different. To control for this, we can run the thought experiment by imagining S and I to be so qualitatively alike that an epistemic difference won’t emerge simply because we are numerically different (remember that, from a phenomenological perspective, we are the same).

A more relevant difference is that S is on Eden and I am on Earth. Eden, we know, is a world where things are as they seem, while Earth is not such a place. Does this introduce an important epistemic difference between the two situations? In particular, does it introduce a difference that suggests that I don’t know that pain is perfect pain (i.e., that pain is a primitive, *sui generis*, etc.) while S knows that perfect pain is a primitive, *sui generis*, etc.? In my view, it does not. First, as mentioned earlier, S doesn’t know nor are they aware that they are on Eden. There isn’t, for instance, a sign that flashes whenever S has an experience that lets them know that things are as they seem (again, S and I are phenomenologically alike). So both S and I seem to be the same with respect to knowing whether things are as they seem.

Second, in the particular case of pain and perfect pain, if we assumed that the difference between the two cases is that pain is not what it seems to be while perfect pain is, we would be begging the question. It is an open question whether pain is as it seems—i.e., whether pain is a primitive, *sui generis*, etc.—and, in fact, I am trying to argue that pain *is* what

it seems to be (pain is perfect pain). So we can't say that the situations differ epistemically because pain is not what it seems to be while perfect pain is.

You might think that, even though at the local level (surrounding pain and perfect pain) we can't specify an external difference between the two cases, globally there is a difference. Eden, *in general*, is a place where things are as they seem while Earth is not. Does this introduce an epistemic difference between the two cases? As I see things, unless one adopts a strong—and, in my view, implausible—form of externalism about justification, the above consideration can at most imply that, whereas S knows that perfect pain is a primitive, *sui generis*, etc., I am justified in believing that pain is a primitive, *sui generis*, etc. (assuming all other things are equal). Assuming that justification is closed under known entailment, this would imply that I am justified in believing that pain is not B, which is just as damaging to the PCS. Thus, whether we are talking about knowing that pain is perfect pain or being justified in believing that pain is perfect pain, the damage to the PCS is just the same. So for the purposes of arguing that (DA) poses a problem for the PCS, the external fact that Eden, *in general*, is a place where things are as they seem while Earth is not does not introduce a relevant epistemic difference.<sup>53</sup>

As I see things then, the only source for an epistemic difference between the two cases capable of implying that I don't know that pain is perfect pain is if I have a defeater for my belief that S lacks for their corresponding belief. In other words, only if I have a reason to not

---

<sup>53</sup> Given that, whether we are talking about knowledge or justification, the problem I want to highlight for the PCS is the same, I will talk in terms of knowing that pain is perfect pain as opposed to being justified in believing that pain is perfect pain. The ensuing arguments can be made either in terms of knowledge or justification to the same effect.

believe that pain is a primitive, *sui generis*, etc., can it be said that there is an epistemically relevant difference between the two cases that implies that (i) I do not know that pain is perfect pain, and so (ii) I do not know that pain is not B.

If this is right, then the PCS must argue that I have a defeater to the belief (a reason to not believe) that pain is perfect pain. But is this something that the PCS can do? In the next section, I will argue that it cannot.

#### 4.3 Reducing Red and Reducing Pain

In order for the PCS to succeed, it must be the case that my direct awareness of pain does not allow me to know something about pain that is at odds with pain being B. In the case of S on Eden, S's direct awareness of perfect pain allows them to know something about perfect pain that is at odds with perfect pain being B. The fact that I am directly aware of pain implies that, if I have no defeaters to the claim that pain is perfect pain, then, in all the ways that matter, my epistemic situation is indistinguishable from S's. Thus, my direct awareness of pain, under the assumption that I have no defeaters to 'pain is perfect pain', should allow me to know that pain is not B (since I know that pain is perfect pain). The PCS, therefore, must argue that there is reason to *not* believe that pain is perfect pain (or there is reason to believe that pain is *not* perfect pain).

In order to investigate whether the PCS can argue this, it will help to consider how someone might make the corresponding argument for red and perfect red. That is, we should consider how someone might argue that there is reason to believe that red is not perfect red. We can then investigate whether this argument translates to the case of pain.

Perfect red, remember, is a primitive, *sui generis*, and external item having *that* nature. Since red seems to be a primitive, *sui generis*, and external item having *that* nature, red seems to be perfect red. Suppose someone told me that red is not what it seems to be. Red, they say, is **F**, where F is (say) a specific reflectance type (i.e., a specific *way* that an object can alter or reflect light).

Let's stipulate that the concept 'red' and the concept 'F' are nomologically co-extensive. Even under that assumption, red will not seem to me to be F. Why? Because red seems to be perfect red, and perfect red and F are clearly different properties (e.g., perfect red is primitive and *sui generis* while F is not). Does this mean that one cannot argue that red is not what it seems to be, i.e., that red is not perfect red? Not necessarily. One may be able to still make this argument given that perfect red and F, though being different with respect to primitiveness and *sui generosity*, are similar with respect to *externality*.

Why is that important? Well, the argument for red being F (not perfect red) is an argument that the concept 'red' picks out F (not perfect red). How can this be given that 'red' seems to pick out perfect red? The thought here is that F is a property out in the world that red objects have; and when I stand in the right causal relation to these objects and F, I have an **experience of red**. When I have this experience, red seems to me to be perfect red ('red' seems to pick out perfect red). But, in fact, red is not what it seems to be. Red actually is F ('red' picks out F).

At first glance, this view of red—call it the **reductive view of red**—seems implausible. The reductionist about red concedes that 'red' seems to pick out perfect red; they also concede that we have some concept of perfect red (they, after all, talk about perfect red). But why think

that the concept 'red' and the concept 'perfect red' are different? Aren't they just the same concept? And doesn't that imply that 'red' picks out perfect red (not F)?

The reductionist says no, and they defend their view by pointing to the semantics of color terms. The semantics of color terms appear to allow for the possibility of color illusions. For example, it is conceivable that, though being aware of perfect red, I might judge an object to be green (I might say something like, 'that object is green but looks red'). Moreover, I might be correct in my judgment. The reductionist claims that, in order to make sense of the truth conditions for judgments like these, i.e., judgments about the colors of objects when we knowingly undergo color illusions, 'red' and 'perfect red' must have different semantic values. In particular, 'red' must refer to F and 'perfect red' must refer to perfect red.

The reductionist also claims that they can explain why color concepts like 'red' behave in the way they do across conceivable scenarios. Recall the external phenomenology of color and our external intuitions about color. We experience colors (like red) as being external, i.e., objective and public, in nature. Colors seem to inhere in external objects, and our intuitions about color place them externally too. This, the reductionist claims, gives color experiences (like the experience of red) and color concepts (like 'red') an "external character". More specifically, it bestows upon color experiences and color concepts the function of tracking items in the external world. If the experiences and concepts actually succeed in referring, then they must refer to items that actually *are* in the external world—specifically, the items they have the function of tracking.

Given the kinds of properties science attributes to objects, properties like perfect red don't seem to have a place in the world. Objects don't instantiate perfect red, and this suggests



that perfect red is not the referent of the experience of red or the concept 'red'. Objects, however, do instantiate F; and F—given that 'F' is nomologically co-extensive with 'red'—stands in the right kind of causal-nomological relation to the experience of red and the concept 'red'. Thus, if the experience and concept do refer to something, their referent must be F.

In essence, the reductionist about red claims that 'red' is just like 'water'. You might think that the reduction of red faces a problem that the reduction of water lacks. Evidently, the reductionist about red has managed to reduce red to F but only at the cost of ignoring perfect red. What does the reductionist have to say about perfect red? Surely, they can't just leave matters with perfect red dangling.

In point of fact though, reductions of items routinely leave the subjective appearances of those items—or the *ways* those items appear—dangling. David Chalmers and Frank Jackson (2001) say:

[S]cientists are prepared to put the explanation of the subjective aspects of a phenomenon (such as color and heat) on hold and settle for an explanation of the objective aspects. This "carving off" strategy arguably does not yield a complete reductive explanation of these phenomena; but at least we know just what we are not explaining. (351)

So even in the case of water, the reduction of water "carves off" the *way* water appears—i.e., the *appearance property* (or properties) of water—from the reduction. If that is right, and if perfect red is the *way* F (red) appears, i.e., the *appearance property* of F (of red), then the reductionist about red should be allowed to "carve it off" too. The upshot is: perfect red is not a *special* problem for the reductionist about red. Whatever problem the reductionist about red faces here, the reductionist about water faces it too.

Regardless of whether we think the reductionist about red has succeeded<sup>54</sup> in arguing that 'red' refers to F (and not perfect red), I think, given the external phenomenology of red and our external intuitions about red, it must be conceded that the reductionist has at least given us a reason to think that red is not perfect red (i.e., they've provided a defeater for the belief that red is perfect red). This is precisely what the PCS must do for the belief that pain is perfect pain. But can you give the same kind of argument to suggest that pain is not perfect pain? Can you argue that 'pain' picks out, not perfect pain, but B instead?

To see whether this is possible, let's start by listing the similarities between the two cases. **(1)** The item seems to be the perfect item: red seems to be perfect red, and pain seems to be perfect pain. **(2)** The item is being reduced to an item distinct from the perfect item: red is being reduced to F and F is distinct from perfect red, and pain is being reduced to B and B is distinct from perfect pain. **(3)** The perfect item and the reducing item differ with respect to primitiveness and *sui generis*: perfect red is, but F is not, primitive and *sui generis*, and perfect pain is, but B is not, primitive and *sui generis*.

There are, however, two important differences between the cases. **(4)** Perfect red is external (i.e., not an experience) but perfect pain is internal (i.e., an experience). **(5)** F stands in a causal relation to the experience of red, but B is supposed to be identical to pain (= the experience of pain since the experience of pain and pain are one). These differences are important because they entail that you cannot argue that pain is B in the same way that you

---

<sup>54</sup> In my view, the semantics of 'red' and the reference of experiences of red are more complicated than the reductionist allows (I discuss these issues and the mind-body problem for color experience in "Does Transparency Solve The Mind-Body Problem?"). Though I don't think the reductionist's story is wholly accurate, I do think that they have succeeded in giving us a reason to doubt that red is perfect red, and that is what is at issue here.

can argue that red is F. To see why this is the case, consider the following three essential features of the argument that red is F ('red' picks out F).

The first essential feature is that perfect red has to be an appearance property of F. In light of (4), we know that perfect red is not an experience. But if perfect red were an experience, then it could not be an appearance property of F. Though experiences may represent items in certain *ways*—and though those *ways* may be the *appearance properties* of those items—the ways that items are represented by experiences can't be confused with the experiences themselves. Experiences are distinct from the ways they represent items—from the appearance properties of items—and so, if something is an experience, then it cannot be an appearance property of an item (a way an item appears).

This implies that perfect red is an appearance property of F (a way that F appears) only if perfect red is not an experience (which it is not). This necessary condition on an item being an appearance property will clearly create a problem in the case of perfect pain. Because perfect pain is an experience, it cannot be an appearance property of B (a way that B appears). So the first essential feature of the argument that red is F (not perfect red) can't be satisfied for the claim that pain is B (not perfect pain).

The second essential feature is that F has to stand in a causal relation to the experience of red and the concept 'red'. The reason that the experience and concept refer to F is that (i) F is a property actually instantiated by objects, and (ii) F stands in the right causal-nomological relation to the experience of red and 'red'. In the case of pain, this translates to the following requirement: B must stand in a causal relation to the experience of pain and the concept 'pain'. But because the experience of pain is pain (to have an experience of pain is to have pain—see

note 38), the first part of this requirement becomes: B must stand in a causal relation to pain. But B isn't supposed to *cause* pain; B is supposed to *be* pain. Moreover, given the constitutional theory, if pain is B, then B must be a *constituent* of 'pain'. B, therefore, can't stand in a causal relation to 'pain', so the second essential feature can't be satisfied.

Finally, the argument that red is F (not perfect red) relied heavily on the conceivability of red illusions, the external phenomenology of red, and our external intuitions about red. Clearly, in the case of pain, these conditions won't be satisfied since pain illusions aren't conceivable, pain phenomenology is internal, and our intuitions about pain put them inside the mind.

It appears then that the argument that we have reason to think that red is not perfect red does not translate into an argument that we have reason to think that pain is not perfect pain. If that is right, then it is hard to see how else the PCS can argue that we have reason to believe that pain is not perfect pain (or to not believe that pain is perfect pain). Note the important role played by direct awareness in blocking translation of the red argument. Because I am directly experientially aware of pain and directly conceptually aware of pain, the distance between pain, the experience of pain, and the concept of pain needed to make the argument work doesn't exist. My direct awareness of pain collapses the experience of pain and pain into one, and it also ties together constitutively my concept of pain and pain. This implies that we don't have the needed room to argue that pain is not perfect pain. Thus, my direct awareness of pain plays a key role in making it the case that I have no defeaters for the claim that pain is a primitive, *sui generis*, and internal item having *this* nature.

Because of this, my epistemic situation with respect to pain on Earth matches S's epistemic situation with respect to perfect pain on Eden. Thus, if S knows that perfect pain is a primitive, *sui generis*, etc., I know that pain is a primitive, *sui generis*, etc. (that pain is perfect pain). And if S knows that perfect pain is not B, then I know that pain (= perfect pain) is not B. Contrary to what proponents of the PCS claim then, my direct awareness of pain allows me to something about pain at odds with the claim that pain is a brain state.

## 5. Conclusion

In this paper, I have argued that the PCS faces a problem. In order to solve the hard problem, the PCS appeals to the constitutional theory. While the constitutional theory does explain why the phenomenal description of pain is a non-contingent description of pain, it has a salient consequence: it implies (DA), i.e., when I have pain, I am directly aware of pain.

Proponents of the PCS claim that (DA) poses no problem for the claim that pain is B. I have argued that it does. (DA) (in conjunction with other conditions that are satisfied) implies that a subject on Eden having perfect pain can come to know that perfect pain is a primitive, *sui generis*, and internal item having *this* nature. Such a subject, then, can come to know that pain is not B. (DA) (in conjunction with other conditions that are satisfied) also implies that I have no defeaters for the claim that pain is perfect pain. Thus, it implies that my epistemic situation is indistinguishable from S's. If that is right, then I too can know that pain is a primitive, *sui generis*, etc., and that pain is not B. The PCS, then, will not work.

Let me conclude this paper by making three comments. First, as may already be clear, the argument I have given against the PCS generalizes to other kinds of view. I have argued that I can know that pain is a primitive, *sui generis*, etc. (that pain is perfect pain). If that is right,

then I can know that pain is not *x* for any *x* (non-contingently and intrinsically) described in terms different than 'is a primitive, *sui generis*, etc.' Thus, my knowledge of pain severely delimits what pain can be.

Second, you might think that, even if a physicalist can't defend the view that pain is B, they can still come to terms with pain being perfect pain. Take, for example, the case of red. Even if red is perfect red, that does not necessarily pose a problem for the physicalist. This is because the physicalist can adopt what Sydney Shoemaker (1994) calls "figurative projectivism" (295). According to figurative projectivism, though you seem to see red, i.e., have experiences of red, red (= perfect red) is not actually instantiated in the world. No objects are red—no objects instantiate the property of being red—so red is an uninstantiated property. Though it might be hard to explain why you seem to see red (have an experience of red),<sup>55</sup> red itself, under this view, is no more problematic than any other uninstantiated property (e.g., the property of being a unicorn). So the physicalist may be able to accommodate red (= perfect red) into their worldview by adopting figurative projectivism and taking red to be an uninstantiated property.

Will the same tactic work for pain under the assumption that pain is perfect pain? No, it will not. In the case of red, one can take red (= perfect red) to be an uninstantiated property because perfect red is not an experience (or even a mental state). So though I may seem to see perfect red (have an experience of perfect red) that does not imply that perfect red (= red) is instantiated. In the case of pain (= perfect pain) though, perfect pain is an experience. Having an experience of perfect pain is just having perfect pain, so perfect pain (= pain) is an actually

---

<sup>55</sup> I examine the difficulty explaining this in "Does Transparency Solve the Mind-Body Problem?"

occurring item (it is a mental state I have). As a result, figurative projectivism is not possible for pain.

Finally, I believe that my account of direct awareness of pain and the knowledge about pain that it leads to is of value because it can help explain what some people mean when they say we are acquainted with pain and our acquaintance with pain gives us insight into pain's nature (what pain is). It is sometimes said that pains are utterly silent on what they are (what their nature is). If I am right, this isn't the case. Our direct experiential awareness and our direct conceptual awareness of pain yields knowledge about what pain is. And what we learn is that pain is unlike anything else.

## Chapter 4



## Does Transparency Solve The Mind-Body Problem?

### 1. Transparency-Intentionalism

An electric car has an engine. It also has a machine that converts electric energy to mechanical energy. Does this mean that an electric car has two things: an engine *and* a machine that converts electric energy to mechanical energy? No, an electric car has only one thing since a machine that converts electric energy to mechanical energy *is* an engine.

You don't have an engine but you do have a mind and mental states. For example, if you are looking at a blue patch on a white wall and you focus on the color you are seeing, you have a mental state we can call an '**experience of blue**'. You also have a brain and brain states; and let's stipulate that your mental states and your brain states are tightly correlated. So, for an experience of blue, there is some brain state B such that: *you* have an experience of blue iff *you* have B (or an instance of B—I'll leave 'an instance of' implicit in what follows).

This biconditional resembles the biconditional 'an *electric* car has an engine iff it has a machine that converts *electric* energy to mechanical energy'. And just as we said that electric cars don't have two things since machines that convert electric energy to mechanical energy are engines, we might say something similar for your experiences of blue and B. That is, we might say that, when you have an experience of blue and when you have B, you don't have two things: an experience of blue *and* B. You have only one thing since B *is* an experience of blue. That is why *your* experiences of blue and B are tightly correlated.

Although few people have difficulty accepting that a machine that converts electric energy to mechanical energy is an engine, the thought that B is an experience of blue is much harder to accept. Consider what an experience of blue is. Suppose that, right now, you were

having an experience of blue. You might think: ‘an experience of blue is just *this*’ where *this* would be you pointing to, to borrow a phrase from McGinn (1989), the “technicolor phenomenology” surrounding you (349). But B is a brain state; it is just a bunch of ions and molecules moving and interacting in a specific way. It is therefore hard to see how B could be an experience of blue. After all, how could that (pointing to B, as it were) maelstrom of ionic and molecular activity be *this* technicolor phenomenology? The two seem utterly different.

The problem of explaining why B is an experience of blue is **the mind-body problem**,<sup>56</sup> and to many the mind-body problem has seemed **hard** (i.e., impossible to solve—in this literature, ‘hard problems’ are impossible to solve while ‘**easy** problems’ are difficult but solvable). In “The Intrinsic Quality of Experience,” Gilbert Harman argued that the belief that the mind-body problem is hard is based on a mistake. It is true, he said, that the mind-body problem is the problem of explaining why B is an experience of blue; but this problem is not the problem explaining why that maelstrom of ionic and molecular activity is *this* technicolor phenomenology. In fact, if this is what you think the mind-body problem is, then you are conflating the “the intentional object of [the] experience . . . [with] the experience itself” (39). Why is this important? Well, according to Harman, if you “[l]ook at a tree and try to turn your attention to intrinsic features of your visual experience . . . you will find that the only features there to turn your attention to will be features of the presented tree . . .” (39). And though Harman did not use the word ‘**transparency**’ to describe this feature of visual experience, it has

---

<sup>56</sup> Strictly speaking, this is the mind-body problem *for color experience*. In my view, there is more than one mind-body problem for experience since experiences differ with respect to transparency (see “Puzzles of Pain” and “Knowing What Pain Is”). In this paper, I will focus exclusively on color experience and the mind-body problem for color experience. Therefore, ‘mind-body problem’ in this paper means ‘mind-body problem for color experience’, in particular, ‘mind-body problem for experiences of blue’.

since come to be known by that name. What exactly transparency is is something that I will discuss later; for now, the important point is that Harman thinks that this feature of, e.g., an experience of blue *solves* the mind-body problem. He states: “All . . . arguments [i.e., all arguments that might go into the mind-body problem] can be *defused* by distinguishing properties of the object of experience from properties of the experience of an object” (31; italics added for emphasis).

Why does this solve the mind-body problem? Well, if experiences of blue are transparent, then you cannot attend to your experience when you have it. You can only attend to what your experience *represents*—what Harman calls ‘the intentional object of the experience’—and in the case of an experience of blue, what your experience represents is (let’s say) blue. If that is right, then when you say, ‘an experience of blue is *this*’, you are wrong. What *this* points to is not an experience of blue; *this* points to blue. And though there may be a **color-body problem**, i.e., a problem explaining why *this* (blue) is a physical property of an object, there is no mind-body problem that is the problem of explaining why that maelstrom of ionic and molecular activity is *this* technicolor phenomenology (since such a problem doesn’t exist).

A number of philosophers<sup>57</sup> have found Harman’s approach to the mind-body problem compelling, though they have found it necessary to slightly refine his solution. According to these philosophers, the transparency of experiences of blue does indeed solve the mind-body problem, though the way it solves it is slightly more complicated than Harman allowed. In “Color and the Mind-Body Problem,” Alex Byrne lays out this revised approach. He argues that,

---

<sup>57</sup> In addition to Byrne (2006), see, e.g., Tye (2002, 2015) and Jackson (2007).

if experiences of blue are transparent, then “[t]here is *no* ‘mind-body problem’, or ‘*hard* problem of consciousness” (223; italics added for emphasis). He says:

[O]nce we recognize the source of the puzzlement [i.e., we are puzzled over the claim that B is an experience of blue and the source of our puzzlement is a failure to recognize that experiences of blue are transparent], the mind-body problem *disappears*. There may an unbridgeable gulf between color and ways of reflecting light; there is *none* between consciousness and brain process. (243; italics added for emphasis)

The reason why Byrne thinks that the transparency of experiences of blue *solves* the mind-body problem is that:

- (I) The transparency of experiences of blue entails that the mind-body problem is the problem of naturalizing intentionality.
- (II) The problem of naturalizing intentionality is easy.<sup>58</sup>

Here, ‘**the problem of naturalizing intentionality**’ refers to the problem of explaining why one item *represents* another item. In the case of experiences of blue, if experiences of blue represent blue and B is an experience of blue, then the problem of naturalizing intentionality for experiences of blue is the problem explaining why B represents blue (henceforth, ‘the problem of naturalizing intentionality’ means ‘the problem of naturalizing intentionality for experiences of blue’). So to say that the mind-body problem is the problem of naturalizing intentionality is to say that the problem of explaining why B is an experience of blue is the problem of explaining why B represents blue. And given the way that the terms ‘hard’ and ‘easy’ are used in this literature, to say that the problem of naturalizing intentionality is easy is to say that the problem of explaining why B represents blue is difficult but solvable.

---

<sup>58</sup> See p. 224, note 3, where Byrne says that, though the problem of naturalizing intentionality is “on the difficult side”, it does not face an “especially powerful objection” and so is not impossible to solve (i.e., it is easy).

Let's call this view of the mind-body problem '**transparency-intentionalism**'. Given that **transparency-intentionalists** think that the mind-body problem is the problem of naturalizing intentionality *because* experiences of blue are transparent, we can say that transparency-intentionalism consists of two claims:

**Transparency-Intentionalism**

**(Intentionalism)** The mind-body problem is the problem of naturalizing intentionality.

**(Easy)** The problem of naturalizing intentionality is easy.

As previously mentioned, a number of philosophers think that transparency-intentionalism solves the mind-body problem by rendering it an easy problem. In this paper, I will argue against these philosophers. I will argue that transparency-intentionalists face a **dilemma** in ensuring that their two components claims—(Intentionalism) and (Easy)—are satisfied. In light of this dilemma, transparency-intentionalism is false, and this entails that, even if we *assume* that experiences of blue are transparent, the mind-body problem is a hard problem.

This paper has three parts. In §2, I explain what it means to say that experiences of blue are transparent. In §3, I discuss what the identity conditions for experiences of blue are (i.e., what condition an item, e.g., a brain state, has to meet in order for it to be an experience of blue). And in §4, I present the **transparency-intentionalist's dilemma**, and I argue that mind-body problem can't be solved.

## 2. The Transparency of Experiences of Blue

Though many philosophers talk of the transparency of experience and think that it is important to the mind-body problem, there is little consensus on what transparency actually means. Some philosophers (e.g., Harman in the passage quoted above) define transparency in terms of attention (in particular, *what* you can attend to when you have an experience), while

other philosophers define it in terms of awareness (in particular, *what* you are aware of when you have an experience).<sup>59</sup> Still others go a step further defining transparency in terms of theoretical notions like ‘phenomenal character’, ‘qualitative character’, ‘essence’, or ‘representation’ that are themselves contentious and, to some extent, poorly understood.<sup>60</sup> This variability in how transparency is defined presents us with a problem since our goal is to understand how the transparency of experiences of blue impacts the mind-body problem. Our first order of business, then, is to clear up this confusion and say precisely what it means to say that an experience of blue is transparent.

Let’s start with a general characterization of transparency. What is a good example of a transparent item? In my view, contact lenses are a paradigmatic transparent item, and if we can understand why contact lenses are transparent, then this will help us say what it means to say that experiences of blue are transparent.<sup>61</sup> I wear contact lenses, and they seem transparent to me. Why are they transparent? At first pass, we might say: it is because, when I wear my contact lenses, I don’t *see* my contact lenses. I see straight through them to the external world.

I don’t wear my experiences of blue like I wear my contact lenses, but I do have these experiences. This suggests the following *preliminary* definition of transparency: experiences of blue are transparent iff I do not *see* them when I *have* them. Since my aim in this paper is to (i)

---

<sup>59</sup> See Kind (2003, p. 230) and Stoljar (2004, p. 356) for characterizations of transparency in terms of attention, and Tye (2002, p. 139; 2015, p. 484) and Mendelovici (2014, p. 151) for characterizations in terms of awareness.

<sup>60</sup> Chalmers (2006, p. 62) defines transparency in terms of phenomenal character; Shoemaker (2006, p. 475) in terms of qualitative character; Jackson (2007, p. 55 ) in terms of essence; and Byrne (2006, pp. 223-224) in terms of representation.

<sup>61</sup> Though no philosopher I am aware of explicates transparency by reference to contact lenses, items similar to contact lenses, e.g., transparent window panes, are often used.

grant the transparency of experiences of blue to the transparency-intentionalist, and then (ii) argue that, even under that assumption, the mind-body problem is hard, I will assume that, when I have an experience of blue, I *do not* see the experience of blue. What then do I see when I have the experience?

Suppose that I am looking at a patch of blue on a white wall and I visually focus on the patch. What am I seeing when I have this experience of blue? The most obvious answer is that I am seeing blue. But what exactly is blue? Given the assumption that experiences of blue are transparent, we can say that blue is not the experience of blue nor is it an intrinsic property of the experience. This is because, if I could see an intrinsic property of the experience, then it wouldn't be like contact lenses. Contact lenses are transparent because, when I wear them, I can't see their, e.g., shape, color, or material constitution (if I could see these things, then, at least partly, I could see my contact lenses). So if the experience of blue is transparent, and if I see blue when I have the experience, then blue cannot be either the experience nor an intrinsic property of the experience. Blue must be extrinsic to the experience.

That still, however, doesn't tell me what exactly blue is. When I reflect on the instance of blue I am visually focused on and seeing, blue seems to be a property having a distinct **qualitative and chromatic nature**. Describing what that qualitative and chromatic nature (**QC nature**) is isn't easy to do if the person you are describing it to hasn't ever seen blue or had an experience of blue. In fact, describing what 'having a distinct QC nature' even means isn't easy to do if the person you are describing it to hasn't ever seen color or had an experience of color. Colors (like blue) seem to have a QC nature that sets them apart from other *kinds* of properties and can't be described in non-color terms. What's more, the distinct QC natures that colors

seem to have seems to *exhaust* them. That is, colors seem qualitative and chromatic through and through: they seem to lack non-qualitative or non-chromatic components.<sup>62</sup> These considerations suggest that colors (like blue) seem to be **primitive, *sui generis* properties having distinct QC natures**. So blue seems to be a primitive, *sui generis* property having a distinct QC nature.

Let's call this property that blue seems to be '**perfect blue**'.<sup>63</sup> Though blue *seems to be* perfect blue, there is a problem with thinking that blue *is* perfect blue. The problem is that, given the way that science characterizes objects, there seems to be no place for a property like perfect blue in the world. At least, it is far from obvious how perfect blue fits into or relates to the scientific characterization of objects, and this suggests that the claim that perfect blue is actually instantiated by objects is out of step with our scientific understanding of the world. From the perspective of science then, perfect blue is an uninstantiated property (like the property of being phlogiston).

Why does this present a problem for the view that blue is perfect blue? Well, to say that blue is perfect blue is to say that our concept (or term) 'blue' picks out perfect blue. But there are at least two reasons to think that 'blue' does *not* pick out perfect blue. First, it seems that color illusions are possible and happen. For example, we routinely say things like, 'that object is

---

<sup>62</sup> See, e.g., Byrne and Hilbert (2007): "The only nature or essence of the colors apparent from reflection on color experience is chromatic through and through" (77). Also, Chalmers (2006): "When I have a phenomenally red experience of an object, the object seems to be simply, primitively, *red*. The apparent redness . . . seems to be a simple qualitative property, with a distinctive sensuous nature. (66; italics in the original)

<sup>63</sup> This terminology comes from David Chalmers's (2006) discussion of Eden and perfect colors in "Perception and the Fall from Eden". Since we can talk about perfect blue, form a clear conception of it, and characterize it as a primitive, *sui generis* property having a distinct QC nature (the distinct QC nature that blue seems to have), there is good reason to think that there is such a property (though it may not be instantiated) and we can form a concept of it.



blue but it looks green to me'. If objects don't instantiate perfect blue, and if, in this circumstance, I don't even seem to see perfect blue, then truth of 'that object is blue but it looks green to me' suggests that 'blue' does not refer to perfect blue.

Second, some philosophers think that the term 'see' is a success term. Frank Jackson (1977), for instance, says: "'Jones sees the tree' . . . entails that the tree exists . . . The one thing that I think cannot be said is that 'Macbeth did see something, and that something did not exist'" (4). One reason philosophers like Jackson think that 'see' is a success term comes from the following consideration. Suppose that I am, unknowingly, hallucinating an apple, and you know that I am hallucinating. If I were to say, 'I see an apple', it seems that you could truthfully respond, 'You are not *really* seeing an apple. It just seems to you that you are.' This suggests that 'see' is a success term, and if that is right, then I see blue iff blue is instantiated. If I see blue then, this implies that 'blue' does not refer to perfect blue.

What then does 'blue' refer to? Let's say that **F** is a property that external objects actually have (i.e., F is included in the scientific description of objects), and let's stipulate that, in this world, the concept 'F' is nomologically co-extensive with the concept 'blue' (i.e., objects are blue iff objects have F). If claims like 'that object is blue but it looks green to me' are true, and if 'see' is a success term, then this suggests that 'blue' refers to F.

What complicates the semantics of 'blue', though, is that there are reasons to think (in addition to 'blue' seeming to pick out perfect blue) that 'blue' refers to perfect blue (and not F). Go back to the case of color illusion. Suppose that instead of saying what you said before you said: 'that object is green but it looks blue to me'. When you say 'it looks blue to me', what exactly do you mean? Clearly, one thing you don't mean is that the object is blue, so whatever

you mean here, success is not at issue (i.e., 'looks blue' doesn't mean 'is blue'). Apparently, we need some *non-success* term to capture what you mean by 'it looks blue to me', and the term I find most appealing is 'subjectively aware'.<sup>64</sup> So when you say 'that object is green but it looks blue to me' you mean the following: when you look at the object, though the object is green, the color you are *subjectively aware* of is blue.

But, in this case, F is not instantiated by the object. So there is little reason or motivation here for thinking that 'blue' in 'it looks blue to me' refers to F. In fact, what is most reasonable to say here is that what you are subjectively aware of is *perfect blue*; so 'blue' in 'it looks blue to me' seems to refer to perfect blue.

There is another reason to think that 'blue' picks out perfect blue. Many philosophers think that, if token experience E<sub>1</sub> and token experience E<sub>2</sub> are phenomenally identical (i.e., what it is like to have E<sub>1</sub> matches what it is like to have E<sub>2</sub>), then E<sub>1</sub> and E<sub>2</sub> are tokens of the same type of experience. Moreover, many philosophers think that one of these experiences could be hallucinatory (say, E<sub>2</sub>). So E<sub>1</sub> and E<sub>2</sub> could be tokens of the same type of experience even though E<sub>2</sub> is hallucinatory but E<sub>1</sub> is not.

Suppose that E<sub>1</sub> and E<sub>2</sub> are experiences of blue. Since E<sub>2</sub> is hallucinatory, this again suggests that 'blue' does not refer to F. In fact, the only thing that seems to unify E<sub>1</sub> and E<sub>2</sub> is subjective awareness of *perfect blue*; and this again suggests that 'blue' refers to perfect blue. More generally, if you think about the conditions under which you would attribute to yourself an experience of blue, it is highly plausible that those conditions *match* the conditions under

---

<sup>64</sup> Harman (1990) uses 'see+' (36) while others use 'seem to see'. It doesn't what term (or expression) we use here so long as that term is not a success term and it does the work that I describe in the text.

which you are subjectively aware of perfect blue. After all, could you be having an experience of blue but not be subjectively aware of perfect blue? And could you be subjectively aware of perfect blue but not be having an experience blue? It seems to me that the answer to these questions is no. If that is right, then you have an experience of blue iff you are subjectively aware of perfect blue. And this (again) suggests that 'blue' refers to perfect blue.

With respect to the semantic value of 'blue' then, we are pulled in two different directions. Locutions like 'S sees blue' or 'x is blue' suggest that 'blue' refers to F (if one makes the assumptions mentioned earlier, e.g., 'see' is a success term); while locutions like 'x looks blue' and 'S has an experience of blue' suggest that 'blue' refers to perfect blue.

For the purposes of explicating transparency and understanding how transparency bears on the mind-body problem, there are two important points to take from this discussion. First, given the complications associated with locutions involving 'blue' and the semantic value of 'blue', our earlier characterization of transparency should be emended. Specifically, transparency must be characterized relative to 'seeing' *and* 'being subjectively aware of'. In the case of contact lenses and their intrinsic properties, it is clear that I neither see them nor am I subjectively aware of them. For experiences of blue and their intrinsic properties then, it must be the case that I neither see them nor am I subjectively aware of them. More generally, we can say that an experience E is **transparent** iff when I have E, (i) I cannot see E or its intrinsic properties, and (ii) I cannot become subjectively aware of E or its intrinsic properties.<sup>65</sup> Since we are assuming that experiences of blue are transparent, this implies that both F and perfect

---

<sup>65</sup> This doesn't rule out becoming *de dicto* aware of E or its intrinsic properties. I can, for instance, become aware *that* I am having E. See Dretske (1995, chapter 2).

blue are extrinsic to these experiences (since they exhaust what I see and am subjectively aware of).

Second, while there is room for debate over the issue of 'see' being a success term or F being the referent of 'blue' in the locutions 'S sees blue' and 'x is blue', there is, I think, little room for debate over two points: (i) 'blue' in the locutions 'x looks blue' and 'S has an experience of blue' (and, for that matter, 'S is subjectively aware of blue') refers to perfect blue; and (ii) S has an experience of blue iff S is subjectively aware of perfect blue. The second point will play a central role in the rest of the paper since it forms the basis for my explication of the mind-body problem. Before we get to that explication though, let me point out that, if 'see' is not a success term and perfect blue is the referent of 'blue' in 'S sees blue', then the properties *seeing blue* and *having an experience of blue* are not distinct. Thus, in the ensuing explication, we could solely focus on the biconditional 'S has an experience of blue iff S is subjectively aware of perfect blue'. Since I will be focusing on this biconditional anyway if 'see' is a success term and F is the referent of 'blue' in 'S sees blue', I will move forward assuming that: (i) 'sees' is a success term; (ii) F is the semantic value of 'blue' in 'S sees blue'; (iii) perfect blue is the semantic value of 'blue' in 'S has an experience of blue'; and (iv) S has an experience of blue iff S is subjectively aware of perfect blue.

### 3. The Identity Conditions for Experiences of Blue

In §1, I claimed that the mind-body problem arises from the thought that, when S has an experience of blue and when S has B, S is having only one item since B is an experience of blue. And the mind-body problem is simply the problem of explaining why B is an experience of blue.

Clearly, before we consider how to explain why B is an experience of blue, we have to know what the **identity conditions** for experiences of blue are. That is, we have to know what conditions an item has to meet in order for it to be an experience of blue. Given what was said in §2, it seems that we already have an answer to this question. We said that S has an experience of blue iff S is subjectively aware of perfect blue. Why is this biconditional true? The best answer seems to be: what makes something an experience of blue is that it *makes* the subject who has it subjectively aware of perfect blue. If that is right, then some item *x* is an experience of blue iff *x* makes S subjectively aware of perfect blue (i.e., iff *x* instantiates the state-property *making S subjectively aware of perfect blue*). This entails that the identity conditions for experiences of blue are: ‘makes S subjectively aware of perfect blue’. Since this description is a description of what experiences of blue *do*—or what *function* they perform—relative to perfect blue, the identity conditions for experiences of blue are captured by a functional description.<sup>66</sup>

Let ‘**the functional description**’ mean ‘the functional description that captures the identity conditions for experiences of blue’ (so whenever I talk about ‘the functional description’ I mean the latter expression). Given that terminology, we can say this about the mind-body problem: since the functional description is ‘makes S subjectively aware of perfect blue’, and since the mind-body problem is the problem of explaining why B is an experience of blue, to solve the mind-body problem we need only explain why B makes S subjectively aware

---

<sup>66</sup> Because of this, the concept (or term) ‘experience of blue’ is not rigid, i.e., it is not the case that, in all possible worlds, the items that fall under the extension of ‘experience of blue’ are tokens of a single kind (or type) of item. Put another way, since experiences of blue are characterized by what they *do*—not by what they *are*—‘experience of blue’ can pick out different kinds of items (e.g., states of a biological nervous system, states of an artificial nervous system) both in the actual world and in different possible worlds.

of perfect blue. At this point, however, the transparency-intentionalist will raise an objection. Though they agree that the identity conditions for experiences of blue are captured by a functional description, they disagree that the functional description is ‘makes S subjectively aware of perfect blue’. In their view, what makes something an experience of blue is that it *represents* blue. So the functional description is ‘represents blue’ (note that ‘blue’, not ‘perfect blue’, appears in the transparency-intentionalist’s characterization of the functional description). Alex Byrne (2006), for instance, says: “[T]he claim of transparency . . . [implies] that all one knows about experiences of red, simply from undergoing them, is that they are experiences of this color” (240; italics in the original and underline added for emphasis). He also says: “[That] experience is transparent . . . [entails that] [i]ntrospection of one’s experience of blue . . . merely yields what the experience is *of or about*” (223-224; italics in the original). Statements similar to these are expressed by other transparency-intentionalists.<sup>67</sup>

In the view of transparency-intentionalists then, because experiences of blue are transparent, the only thing we know about them is that they are *of or about*—or, more simply, they *refer to*—blue. This suggests that the functional description is ‘refers to blue’. Why then do I say that transparency-intentionalists think the functional description is ‘represents blue’? It is because, as transparency-intentionalists use these expressions, ‘refers to blue’ and ‘represents blue’ mean the same thing. Consider the following passage from Byrne (2006):

If we like, we can say experiences of blue have a ‘qualitative character’, but that is simply because they *represent* that objects have a ‘qualitative’ property—namely, blueness. The experiences [of color] are, in this respect, like the *words* ‘blue’, ‘purple’, ‘yellow’, and so forth. We may say that ‘blue’ is more similar in a salient qualitative respect to ‘purple’ than to ‘yellow’, but that can only mean that ‘blue’ *represents* a property that is more similar in a salient qualitative respect to the property *represented*

---

<sup>67</sup> See, e.g., Harman (1990) and Tye (2002, 2015).

by 'purple' than it is to the property *represented* by 'yellow'. Likewise for the experiences of blue, purple and yellow. They too *inherit* their 'qualitative character' from the qualitative nature of the properties they *represent* . . . [I]f we can provide a satisfying naturalistic explanation of the qualitative nature of the *colors*, there will be no mysterious qualitative residue left in *experience*. (224; italics in the original and added for emphasis)

This passage makes clear three key components of transparency-intentionalism. First, because transparency-intentionalists see experiences as being like words, 'refers to blue' and 'represents blue' are (in their view) synonymous. Thus, the functional description is 'represents blue'.

Second, because (in their view) the functional description is 'represents blue', (Intentionalism)—i.e., the first of the two claims making up transparency-intentionalism—follows. (Intentionalism), remember, said that the mind-body problem is the problem of naturalizing intentionality, where the problem of naturalizing intentionality is the problem of explaining why B represents blue. If the functional description is 'represents blue', and if the mind-body problem is the problem of explaining why B is an experience of blue, then the mind-body problem is the problem of explaining why B represents blue.

Finally, because transparency-intentionalists think that—under the assumption that the functional description is 'represents blue'—the only hard problem is accounting for the qualitative nature of colors, they think that (Easy) is true, i.e., the problem of naturalizing intentionality (= the mind-body problem) is easy (Byrne: "there will be no mysterious qualitative residue left in *experience*").<sup>68</sup> With (Intentionalism) and (Easy) in place, we have

---

<sup>68</sup> Recall that Harman (1990) claimed that the transparency of experience "defuses" (31) the mind-body problem, and Byrne (2006) claimed that the mind-body problem "disappears" (243) once we recognize that experience is transparent (see §1). Byrne (2006) also said that, if experience is transparent, "[t]here is no 'mind-body problem', or 'hard problem of consciousness'" (223), and he and other transparency-intentionalists make this claim because they think (i) the mind-body problem is the problem of naturalizing intentionality, and (ii) the problem of

both component claims of transparency-intentionalism and, therefore, a solution to the mind-body problem.

Does transparency-intentionalism solve the mind-body problem? Let's focus on the transparency-intentionalist's claim that the functional description is 'represents blue'. Is this the functional description? The answer is that it depends. Specifically, it depends on what 'represents' means and what 'blue' refers to in the locution 'represents blue'.

Consider, for instance, the semantic value of 'blue'. As discussed in §2, with respect to the reference of 'blue', we are pulled in two different directions. If we consider statements like 'S sees blue' and 'x is blue', then 'blue' seems to refer to F. But if we consider statements like 'x looks blue' and 'S has an experience of blue', then 'blue' seems to refer to perfect blue. So, before we can say whether the functional description is 'represents blue', we need to know what 'blue' refers to in the expression 'represents blue'.

What's more, we need to know what 'represents' *means* in 'represents blue', and this is not a simple matter. The problem here is that 'represents' has no clear, unambiguous pre-theoretical meaning.<sup>69</sup> Pre-theoretically, we may say that visual experiences, beliefs, sub-personal states, words, pictures, thermometers, etc., all represent things, but it is far from clear that 'represents', used in all these various ways, has a single (or even unified) sense or meaning. So we can't just talk of 'x represents y' and assume that we know what we mean by this. Some

---

naturalizing intentionality is easy. Note that the second claim is essential to transparency-intentionalism since transparency-intentionalists think that the first claim *solves* the mind-body problem.

<sup>69</sup> See, e.g., Block (2005): "To understand representationism, we need to know what to make of the phrase 'representational content' as applied to an experience. There is no clear *pretheoretical* notion of representational content as applied to an experience, certainly none that will be of use to the representationist" (137; italics in the original). Note that Block uses the term 'representationism' in place of 'intentionalism'.



explication of ‘represents’, then, is needed before we can evaluate the transparency-intentionalist’s claim that the functional description is ‘represents blue’.

Now you might think that a transparency-intentionalist can respond to this by making a similar criticism of my view. I claim that the functional description is ‘makes S subjectively aware of perfect blue’. But what does ‘subjective awareness’ mean and what does ‘perfect blue’ refer to? As I see things though, I have already answered these questions. Perfect blue is a primitive, *sui generis* property having a distinct QC nature (the distinct QC nature that blue seems to have). And I claimed that perfect blue is not instantiated by objects in the world since it is not included in—and does not have any clear connection to—the scientific conception of objects and the world.

What is ‘subjective awareness’? I specified subjective awareness by pointing to certain examples where it is manifest. For example, if you are undergoing a color illusion where a green object looks blue, you are subjectively aware of perfect blue. If I am not undergoing the color illusion, I am subjectively aware of perfect green. Also, I claimed that a veridical experience of blue and a hallucinatory experience of blue are tokens of the same type of experience.<sup>70</sup> Moreover, what makes them tokens of the same type is that, in both instances, you are subjectively aware of perfect blue. And I could point to other examples to illustrate subjective awareness. If a color-blind person and I am looking at a blue object, the difference between us is that I am subjectively aware of something the color-blind person is not. What am I

---

<sup>70</sup> Not all views of perception will agree with this. Some disjunctivists, for instance, don’t think that a hallucination and a veridical experience can be tokens of the same type of experience. There isn’t space in this paper to take up this issue. Suffice it to say, in this paper, I assume that (i) experiences are type-identified by their phenomenology, and (ii) a hallucinatory experience of blue and a veridical experience of blue can be phenomenally indistinguishable.

subjectively aware of? I am subjectively aware of perfect blue. And if me and my invert are looking at the same blue object, the difference between us is that I am subjectively aware of perfect blue while they are subjectively aware of perfect yellow.

You might complain that these examples are too simple, naïve, and loose to specify the meaning of ‘subjective awareness’. You might say that something more is needed, e.g., a theoretical or reductive definition. The problem however is that, when it comes to talking about consciousness, in particular, the mind-body problem, we have to start with what we have an independent, non-theoretical grip on with respect to, e.g., having an experience of blue and *then* move on from there. If we don’t start from such a point, then it won’t be clear that our theorizing about “experiences of blue” is really about experiences of blue and the problems they present (i.e., the mind-body problem). Indeed, part of the way we ground ourselves in our discussions of the mind-body problem is by getting clear on our starting points, i.e., by getting clear on what we have an independent, non-theoretical grip with respect to having experiences of blue. And it is hard to see how we could do that other than by pointing to the kinds of examples I discuss above (illusions, hallucinations, the color-blind, inverts, etc.).

In my view then, a notion like ‘subjective awareness’ is essential to properly characterizing the mind-body problem. And given its fundamentality, it is likely that it can only be specified by pointing to certain kinds of illustrative examples. These considerations also suggest that the term ‘represents’, given the ambiguity surrounding its meaning, has no role to play in helping us get clear on our starting points. It will not be involved in specifying what we have an independent, non-theoretical grip on with respect to experiences of blue and the mind-body problem. It, therefore, will not be involved in our specification of the mind-body problem.

The upshot is that we can't start by saying that the functional description is 'represents blue'. We have to start, instead, by saying that the functional description is 'makes S subjectively aware of perfect blue'. Only by starting here will we get our initial characterization of the mind-body problem right. And if the functional description is 'makes S subjectively aware of perfect blue', then the mind-body problem is: the problem of explaining why B makes S subjectively aware of perfect blue.

It is important to distinguish my view from another view that it may be confused for. Byrne, in the passage quoted above, says: "If we like, we can say experiences of blue have a 'qualitative character', but that is simply because they represent that objects have a 'qualitative' property—namely, blueness . . . [Experiences] inherit their 'qualitative character' from the qualitative nature of the properties they represent." It's noteworthy here that Byrne uses the term 'qualitative character' because, in using this term, he is taking aim at what transparency-intentionalists generally take to be the main rival—and only plausible alternative—to their view. That rival view is (what I'll call) 'the qualitative character view'.<sup>71</sup>

According to the qualitative character view, experiences *have* a qualitative character. So an experience of blue *has* a qualitative character Q. What is Q? Well, if you think about two tokens of an experience of blue, they seem to share a certain qualitative property. What qualitative property do they share? The best answer seems to be: perfect blue. Perfect blue, in other words, seems to capture how the two experiences are qualitatively (or phenomenally) alike. So, Q = perfect blue.

---

<sup>71</sup> This view could also be called the 'phenomenal character view' since 'qualitative character' and 'phenomenal character' are used interchangeably in this literature.

If experiences of blue have Q (i.e., have a qualitative character Q), then this suggests that experiences of blue have perfect blue. And according to a proponent of the qualitative character view, what makes something an experience of blue is having perfect blue (i.e., ‘having perfect blue’ captures the identity conditions of experiences of blue). But this can’t be right, say transparency-intentionalists. Because experiences of blue are transparent, perfect blue is extrinsic to experiences of blue. So experiences of blue don’t *have* perfect blue; they *represent* perfect blue. More accurately, transparency-intentionalists respond to qualitative character theorists by (i) taking the semantic value of ‘blue’ to be perfect blue, and then (ii) saying *having perfect blue = representing blue*.<sup>72</sup>

Whatever the merits of this argument against the qualitative character view, I think it is clear that it has no force against my view since I am not claiming that experiences of blue have perfect blue. I am claiming that the functional description is ‘makes S subjectively aware of perfect blue’, and this is different from saying—as the qualitative character view says—‘having perfect blue’ captures the identity conditions of experiences of blue. It can help here to think of the dialectic between me, the transparency-intentionalist, and the qualitative character theorist as follows.

The transparency-intentionalist and the qualitative character theorist both think that “qualitative stuff” is the problematic element in this dispute. In the case of experiences of blue, the qualitative stuff is perfect blue, and accounting for perfect blue is what is hard. Qualitative character theorists put perfect blue (the qualitative stuff) in the mind—specifically, in an

---

<sup>72</sup> Note that, in order for Byrne’s response to the qualitative character theorist to have force, ‘blue’ must refer to perfect blue. This kind of constraint on transparency-intentionalism will be important when I discuss the dilemma transparency-intentionalism faces.

experience of blue—and this makes the mind-body problem hard since, in order to solve it, we must explain why B has perfect blue. Transparency-intentionalists, on the other hand, put perfect blue (the qualitative stuff) out in the world—specifically, onto an external object *O*—and this makes the color-body problem, i.e., the problem of explaining why perfect blue is a physical property of *O*, the hard problem. The mind-body problem, however, is (according to transparency-intentionalists) easy since it is just the problem of explaining why B represents perfect blue (or blue assuming ‘blue’ picks out perfect blue).

Now I agree with transparency-intentionalists that experiences of blue do not have perfect blue (the qualitative stuff). And I think that perfect blue is extrinsic to experiences of blue since experiences of blue are transparent. I happen to think that, even under these assumptions, the color-body problem is not hard, though that is not a claim I will defend here.<sup>73</sup> Instead, my aim is to argue that, even if you put the qualitative stuff (perfect blue) out into the world, the mind-body problem is still hard since we must explain why B makes us subjectively aware of the qualitative stuff (perfect blue) and we can’t do this. That is something I will argue in §4.3 and §4.4. For now, the important point is that my view is distinct from the qualitative character view, and whatever arguments transparency-intentionalists have against that view don’t translate into arguments against my view.

What we have then is this. I, like the transparency-intentionalist, put the qualitative stuff (perfect blue) out of the mind. And I, like the transparency-intentionalist, think that the mind-body problem has something to do with explaining why B performs some function relative to perfect blue. While the transparency-intentionalist claims that the functional

---

<sup>73</sup> See “Knowing What Pain Is” for resolving the color-body problem.

description is 'represents blue', I have argued that that can't be where we start. In order to properly characterize the mind-body problem (at least initially), we must start from the claim that the functional description is 'makes S subjectively aware of perfect blue'. Under this characterization of the functional description, the mind-body problem is the problem of explaining why B makes S subjectively aware of perfect blue.

Does this mean that (Intentionalism), i.e., the mind-body problem is the problem of naturalizing intentionality, is false? Not necessarily. Though we have to *start* by saying the functional description is 'makes S subjectively aware of perfect blue', that doesn't mean that this is where we must *end*. As mentioned previously, the meaning of 'represents' and the semantic value of 'blue' in 'represents blue' are vague.<sup>74</sup> So the transparency-intentionalist can still argue for (Intentionalism) by specifying the meaning of 'represents' and the semantic value of 'blue' so that the property of making S subjectively aware of perfect blue and the property of representing blue come out the same (i.e., they can be seen to be the same property). In fact, the transparency-intentionalist *must* do this if their view is to survive.

Of note, there is a further restriction on the transparency-intentionalist. In order for (Intentionalism) to be true, a certain property identity must hold: *making S subjectively aware of perfect blue = representing blue*. This identity must clearly be **necessary** since the functional description 'makes S subjectively aware of perfect blue' captures the identity conditions of experiences of blue. But, in addition to being necessary, the identity must be **a priori** as well.

Why is that the case?

---

<sup>74</sup> Though in the argument against the qualitative character view, the transparency-intentionalist takes a stand on the semantic value of 'blue'—it refers to perfect blue—that doesn't mean that, in responding to my view, they need to say the same thing. So we shouldn't hold them to that just yet.

Well, think about Byrne's argument against the qualitative character view. Byrne argued that the property *having qualitative character Q (= having perfect blue)* is identical to the property *representing blue (= representing perfect blue)*. That argument did not rely on making observations, running experiments, or formulating empirically-confirmed theories. The argument, instead, was *a priori*; it was supposed to follow from our concepts of *having qualitative character Q* and *representing blue*. By the same token, the identity *making S subjectively aware of perfect blue = representing blue* must be, if true, *a priori*. It must follow from our concepts of *making S subjectively aware of perfect blue* and *representing blue*.

More generally, intentionalists—of all stripes—are committed to the view that the properties that make something a token of a type of experience are, necessarily, representational properties. But this thesis about what experiences are is not one that they think we discover and justify by empirically investigating the world (e.g., by investigating brains). Instead, we discover and justify this thesis by *a priori* reflection on our concept of experience (in particular, what it is to have an experience) and our concept of representation (in particular, what it is to represent something). This is why, according to intentionalists, their thesis *reveals* what consciousness *is*; it tells us, e.g., what *it is* to have an experience of blue.

So intentionalists aren't in the business of asserting necessary, *a posteriori* identities between the properties that make something a token of a type of experience (e.g., *making S subjectively aware of perfect blue*) and representational properties (e.g., *representing blue*). In their view then, the property identity *making S subjectively aware of perfect blue = representing blue* isn't like the property identity *being water = being composed of H<sub>2</sub>O molecules*. Whereas the latter identity is necessary and *a posteriori*—since we discover what water is by empirically

investigating samples of water—the former identity is necessary and *a priori*. We don't discover its truth by empirically investigating something (e.g., a brain). Instead, we discover—and justify—the identity by *a priori* reflection, and this entails that, if *making S subjectively aware of perfect blue = representing blue* is true, then it is both necessary and *a priori*.<sup>75</sup>

So the question we must ask is: can the transparency-intentionalist specify a meaning for 'represents' and a semantic value for 'blue' in 'represents blue' such that *making S subjectively aware of perfect blue = representing blue* is necessary and *a priori*? It is important here to highlight the fact that the transparency-intentionalist is under yet another constraint. Remember that, in addition to satisfying (Intentionalism), the transparency-intentionalist must also satisfy (Easy), i.e., the problem of naturalizing intentionality is easy. So, in point of fact, the transparency-intentionalist must specify a meaning for 'represents' and a semantic value for 'blue' in 'represents blue' such that the following claims are true:

**(Property Identity):** *making S subjectively aware of perfect blue = representing blue* is necessary and *a priori* [this is needed in order to satisfy (Intentionalism)]

**(Easy)** The problem of naturalizing intentionality is easy

I will argue that the transparency-intentionalist cannot do this because they face a **dilemma**: (i) any specification of the meaning of 'represents' and the semantic value of 'blue' in 'represents blue' that makes (Easy) true makes (Property Identity) false [thereby making (Intentionalism)

---

<sup>75</sup> This does not mean that *a posteriori* information won't be relevant to determining the truth of the identity. Remember that 'represents' and 'blue' in 'represents blue' are both vague. So, prior to evaluating the identity, we must be told by the transparency-intentionalist what 'represents' means and what 'blue' refers to in 'represents blue'. What 'represents' means will be an *a priori* matter since it will be explicated theoretically. But what 'blue' refers to may be something that science determines. The key point though is that, once we are told what 'represents' means and what 'blue' refers to in 'represents blue', the identity must then be necessary and *a priori*. Intentionalists, after all, are (like functionalists) type-A materialists. They do not think that zombies are conceivable, Mary doesn't know what red looks like in her black-and-white-room, and there is an explanatory gap. This could only be if the identity in the text is necessary and *a priori*.



false], and (ii) any specification of the meaning of ‘represents’ and the semantic value of ‘blue’ in ‘represents blue’ that makes (Property Identity) true [thereby making (Intentionalism) true] makes (Easy) false. Call this the **transparency-intentionalist’s dilemma**. The transparency-intentionalist’s dilemma is why transparency-intentionalism is false and the mind-body problem is a hard problem (even under the assumption that experiences of blue are transparent).

#### 4. The Transparency-Intentionalist’s Dilemma

Let’s go down both horns of the dilemma starting with the first horn.

##### 4.1 Satisfying (Easy)

Suppose that the transparency-intentionalist specifies the meaning of ‘represents’ and the semantic value of ‘blue’ in ‘represents blue’ so that (Easy) is true. For example, suppose that, in ‘represents blue’, ‘represents’ means ‘makes S see’ and ‘blue’ refers to F (F, remember, is a property that objects actually instantiate, and it is included in the scientific characterization of objects). If ‘see’ is a success term (as we are assuming), then ‘blue’ likely refers to F in ‘S sees blue’. In that case, *seeing blue* is likely a causal relation of some kind (e.g., a causal-nomological or causal-historical relation). More accurately, S sees blue (i.e., sees F) by virtue of having an internal state that stands in the appropriate causal relation to blue (F)—say, the relation of being causally correlated with blue. So S sees blue (instantiates the subject-property *seeing blue*) by virtue of having an internal state that makes S see blue (instantiates the state-property *making S see blue*); and because the state-property *making S see blue* is identical to the state-property *being causally correlated with blue*, this implies that S sees blue by virtue of having an internal state that is causally correlated with blue (instantiates the state-property *being causally correlated with blue*).

No doubt, there are pre-theoretical uses of 'represents' that support the view that 'represents' means something like 'is causally correlated with'. And if 'blue' refers to F in 'represents blue', then, plausibly, one could hold that the state-property *making S see blue* = the state-property *representing blue* since one could claim that state property *representing blue* = the state-property *being causally correlated with blue*.

This specification of the meaning of 'represents' and the semantic value of 'blue' makes the problem of naturalizing intentionality easy. In order to solve the problem of naturalizing intentionality, we need only explain why B represents blue, i.e., why B is causally correlated with F (= blue). And that is something that seems possible to do since we can explain why a certain brain state is causally correlated with a certain physical property of objects.

However, even though the above specification makes (Easy) true, it entails that (Property Identity) is false. This is because there are many things that might be causally correlated with F without being such that they make the subjects who have them subjectively aware of perfect blue (for some of these items, it may not even make sense to talk about subjects having them). Moreover, given that S can have experiences of blue even though S does not see blue (= F), the functional description cannot be 'makes S see blue'. Since the property *making S see blue* is identical to the property *representing blue*, this implies that the functional description is not 'represents blue'. If that is right, then (Property Identity) is false: *making S subjectively aware of perfect blue* ≠ *representing blue*.

If (Property Identity) is false, then (Intentionalism) is false: the mind-body problem is not the problem of naturalizing intentionality. The mind-body problem is the problem of explaining why B makes S subjectively aware of perfect blue, and, in §4.3 and §4.4, I will explain why this

problem is hard. For now, the important point is that, even though the above specification of ‘represents’ and ‘blue’ makes (Easy) true, it entails that (Property Identity)—and, so, (Intentionalism)—is false.

The problem with the above specification is that it is not guided by (Property Identity). It specifies the meaning of ‘represents’ and the semantic value of ‘blue’ in ‘represents blue’ with an eye towards making (Easy) true. Typically, such accounts will have to specify ‘represents’ in terms of causal notions or tracking; and they will have to take ‘blue’ to refer to F. Accounts like these make it easy to explain why B represents blue since we can explain why B, e.g., is causally correlated with or tracks F. But the price we have to pay for solving the problem of naturalizing intentionality is that (Property Identity) is not plausible: it is not plausible that *making S subjectively aware of perfect blue = representing blue* is necessary and *a priori*. Put another way, the price we have to pay is that the mind-body problem is not the problem of naturalizing intentionality.

#### 4.2 Satisfying (Property Identity)

Suppose then we specified ‘represents’ and ‘blue’ in ‘represents blue’ with an eye towards satisfying (Property Identity). Note that, if ‘blue’ does not refer to perfect blue, then it will not be plausible that *making S subjectively aware of perfect blue = representing blue* is necessary and *a priori*. This need for the arguments in *making S subjectively aware of \_\_* and *representing \_\_* to agree was apparent, in a different form, in Byrne’s argument against the qualitative character view (see note 72). Byrne argued that the state-property *having perfect blue* = the state-property *representing blue*; but this was plausible only if we assumed, as Byrne himself did, that ‘blue’ referred to perfect blue. So, in order for (Property Identity) to hold, the

arguments in *making S subjectively aware of* \_\_ and *representing* \_\_ must agree. This implies that 'blue' refers to perfect blue in 'represents blue'.

What about 'represents'? How should we specify its meaning? Because 'represents' has no clear, unambiguous pre-theoretical meaning, it will have to be specified theoretically; in particular, it will have to be specified using expressions that do not themselves involve the term 'represents'. Needless to say, there have been many ways 'represents' has been theoretically explicated in the literature. The notion of representation has been specified in all of the following ways: having accuracy conditions, having the function of indicating F, being caused by F under ideal conditions, sensorily (or experientially) entertaining (or zeeing) a proposition, it subjectively seeming that p (p is a proposition), among many others.<sup>76</sup> Despite this lack of uniformity in theoretical specifications of 'represents', for our purposes, what is important is this: however the transparency-intentionalist chooses to specify 'represents', it must be true that *making S subjectively aware of perfect blue = representing perfect blue* is necessary and *a priori* (note that what is represented is perfect blue since, down this horn, 'blue' must refer to perfect blue). This requirement functions as a constraint on specifications of 'represents', in particular, it constrains the way we understand the expressions making up the specification.

Why is this important? Well, suppose that  $\Delta$  is a specification of 'represents' that entails that *making S subjectively aware of perfect blue = representing perfect blue* is necessary and *a priori*. Though  $\Delta$  specifies the meaning of 'represents', it will contain expressions that

---

<sup>76</sup> Siegel (2016) discusses an explication of and Burge (2010) actually explicates 'representing' in terms of having accuracy conditions; Dretske (1995) explicates 'representing' in terms of having the function of indicating F; Tye (2005b) explicates 'representing' in terms of being caused by F under ideal conditions; Pautz (2007) explicates 'representing' in terms of sensorily entertaining, Siegel (2016) discusses an explication of 'representing' in terms of experientially entertaining, and Byrne (2009a) explicates 'representing' in terms of zeeing a proposition; and Byrne (2001) and Tye (2009) explicate 'representing' in terms of it subjectively seeming that p.

themselves need explication. For example,  $\Delta$  will contain expressions like ‘accuracy conditions’, ‘function’, ‘indicating’, ‘ideal’, ‘sensorily entertaining’ (or ‘experientially entertaining’ or ‘zeeing’), or ‘subjectively seems’; and each of these expressions will need to be explicated—i.e., the transparency-intentionalist will need to tell us what each of these expressions mean—so that we know what it means to, e.g., have the function of indicating F or sensorily entertain a proposition.<sup>77</sup> More to the point, the transparency-intentionalist must explicate each of these expressions under the constraint that *making S subjectively aware of perfect blue = representing perfect blue* be necessary and *a priori*. After all, if we are not told what, e.g., having the function of indicating F or sensorily entertaining a proposition means, then how can we determine whether *making S subjectively aware of perfect blue = representing perfect blue* is necessary and *a priori*?

Given this constraint on the expressions making up  $\Delta$ , it is hard to see how we can avoid explicating these expressions in such a way that they just end up meaning ‘making S subjectively aware of perfect blue’ (at the state-level) and ‘being subjectively aware of perfect blue’ (at the subject-level). Suppose, for instance, that ‘S represents perfect blue’ is explicated in terms of the expression ‘it subjectively seems to S that something is perfect blue’. What does this explicating expression mean? Well, given that S represents perfect blue by virtue of having an internal state that represents perfect blue<sup>78</sup>, it follows that it subjectively seems to S that something is perfect blue because S has an internal state that *makes it* subjectively seem to S

---

<sup>77</sup> I assume it is clear that ‘having the function of indicating F’, ‘sensorily entertaining a proposition’, etc., have no clear, pre-theoretical meaning.

<sup>78</sup> Note that even though ‘represents’ occurs both at the subject-level and at the state-level, the subject-level ‘represents’ is derivative of the state-level ‘represents’. That is, S represents perfect blue only insofar as S has an internal state that represents perfect blue.

that something is perfect blue. We can then ask: what does 'makes it subjectively seem to S that something is perfect blue' mean? Well, under the constraint that this expression be interpreted so that *making S subjectively aware of perfect blue = making it subjectively seem to S that something is perfect blue (= representing perfect blue)* is necessary and *a priori*, there must be an *a priori*, conceptual connection between 'S being subjectively aware' and 'it subjectively seems to S'. If that is right, then it is hard to see how 'it subjectively seems to S that something is perfect blue' and 'makes it subjectively seem to S that something is perfect blue' won't just end up meaning 'being subjectively aware of perfect blue' (at the subject-level) and 'makes S subjectively aware of perfect blue' (at the state-level).

Here is another way of making the same point. In §3, I said that the notion of 'subjective awareness' was essential to properly characterizing the mind-body problem. I said that what we had an independent, non-theoretical grip on with respect to having an experience of blue was being subjectively aware of perfect blue; and I defended this view by pointing to certain kinds of examples, e.g., the difference between me and a color blind person is that I am subjectively aware of perfect blue while they are not. If 'subjectively seems' is to play the same role as 'subjectively aware', and if the identity *making S subjectively aware of perfect blue = making it subjectively seem to S that something is perfect blue* is to be necessary and *a priori*, then when we consider the kinds of cases that establish our starting point for a discussion of mind-body problem, e.g., the difference between me and a color-blind person, we will be forced to say something like this: when I and a color-blind person look at a blue object, the difference between us is that it subjectively seems to me, but not to the color-blind person, that something is perfect blue.

But putting things this way makes clear that ‘it subjectively seems to S that something is perfect blue’ and ‘being subjectively aware of perfect blue’ (at the subject-level) and ‘makes it subjectively seem to S that something is perfect blue’ and ‘makes S subjectively aware of perfect blue’ (at the state-level) are simply different ways of saying the same thing. And the same could be said for all the other ways of explicating ‘represents’. Put simply, in order to ensure that *making S subjectively aware of perfect blue = representing perfect blue* is necessary and *a priori*, expressions like ‘has the function of indicating perfect blue’, ‘caused by perfect blue under ideal conditions’, ‘sensorily entertains that something is perfect blue’, etc., all have to be interpreted to either mean ‘is subjectively aware of perfect blue’ or ‘makes S subjectively aware of perfect blue’, depending on whether we are operating at the level of the subject (S) or the state (the experience of blue).

In light of this, I think it is clear that any specification of the meaning of ‘represents’ and the semantic value of ‘blue’ in ‘represents blue’ that is under the constraint that (Property Identity), i.e., *making S subjectively aware of perfect blue = representing blue* is necessary and *a priori*, be true will entail that ‘representing blue’ means ‘being subjectively aware of perfect blue’ (at the level of the subject) and ‘representing blue’ means ‘making S subjectively aware of perfect blue’ (at the level of the state). So if we want to satisfy (Property Identity)—and, thereby, satisfy (Intentionalism) (the mind-body problem is the problem of naturalizing intentionality)—it must be the case that, in ‘represents blue’, ‘represents’ means ‘makes S subjectively aware of’ and ‘blue’ refers to ‘perfect blue’. The question now is: is the problem of naturalizing intentionality easy under these specifications? Is (Easy) true down this horn?

#### 4.3 Solving Neural Functional Problems

In §4.1, I argued that, if the transparency-intentionalist specifies ‘represents’ and ‘blue’ in ‘represents blue’ in such a way that the problem of naturalizing intentionality is easy, then (Property Identity) is false. This implies that the mind-body problem is the problem of explaining why B makes S subjectively aware of perfect blue, and so the mind-body problem is not the problem of naturalizing intentionality.

In §4.2, I argued that, if the transparency-intentionalist specifies ‘represents’ and ‘blue’ in ‘represents blue’ in such a way that (Property Identity) is true, then ‘represents blue’ means ‘makes S subjectively aware of perfect blue’. Under this specification, the mind-body problem is the problem of naturalizing intentionality since the mind-body problem is the problem of explaining why B makes S subjectively aware of perfect blue and the problem of naturalizing intentionality is the problem of explaining why B makes S subjectively aware of perfect blue (= the problem of explaining why B represents blue).

Note that, in either case, **the mind-body problem is the problem of explaining why B makes S subjectively aware of perfect blue**. I now want to argue that **this problem is hard** (i.e., can’t be solved). If this problem is hard, then transparency-intentionalism is false because it faces a dilemma. Any specification of ‘represents’ and ‘blue’ that satisfies (Easy) will not satisfy (Property Identity). Therefore, (Intentionalism) will be false; moreover, the mind-body problem will be hard. Down the other horn, any specification of ‘represents’ and ‘blue’ that satisfies (Property Identity)—though it will satisfy (Intentionalism)—will not satisfy (Easy) since the problem of naturalizing intentionality (= the mind-body problem) will be hard. Thus, if the problem of explaining why B makes S subjectively aware of perfect blue is hard, transparency-



intentionalism will be false since its two components claims—(Intentionalism) and (Easy)—can't be jointly satisfied.

But why exactly is the mind-body problem hard? Why can't we explain why B makes S subjectively aware of perfect blue? In order to see what difficulty we face here, we must recognize what kind of problem the mind-body problem is. The mind-body problem is the problem of explaining why B makes S subjectively aware of perfect blue because the identity conditions for experiences of blue are captured by the functional description 'makes S subjectively aware of perfect blue'. In order to solve the mind-body problem, we must explain why B does what this functional description describes (or why B performs the function described by this functional description).

This kind of problem—i.e., explaining why a certain item *satisfies* a functional description—is a **functional problem**. And because B is a brain state, the kind of functional problem the mind-body problem presents is a **neural functional problem**. It is a problem having to do with explaining why a certain *brain state* does a certain thing or performs a certain function. In order to figure out how to solve the mind-body problem then, we must answer the following question: how, in general, do we solve neural functional problems? How do we explain why certain brain states satisfy certain functional descriptions?

Clearly, our concern here is with *brain states* and their satisfaction of functional descriptions, and this means that we must look to *neuroscience* in order to answer the questions raised above. This is because we solve neural functional problems by doing neuroscience. When we explain why a given brain state does a certain thing or performs a certain function, the kind of explanation we give is a neuroscientific explanation. So to answer

the above questions, we must determine how neuroscientists explain why a given brain state satisfies a given functional description.<sup>79</sup>

It goes without saying that a complete and satisfactory account of neuroscientific explanation is beyond the scope of this paper.<sup>80</sup> For our purposes though, a complete and satisfactory account is not needed. We need only focus on one essential element of neuroscientific explanations since it is this element that plays the biggest role in the mind-body problem. The element I am referring to is the specification of a **physical mechanism**. When neuroscientists explain why a brain state satisfies a functional description, there are many things they do. They posit hypotheses and causal mechanisms, run and report on experiments, measure things using sophisticated instruments, make observations and calculations, and do many other things. Importantly, they do all this from within the context of a vast background theory that “anchors” these hypotheses, posits, experiments, measurements, observations, and calculations.<sup>81</sup> In addition to all this, there is one other important thing that neuroscientists do from within the background theory: they specify a *physical mechanism* that shows how the brain state satisfies the functional description (i.e., how the brain state does what the

---

<sup>79</sup> I think it is clear that our explanations of why brain states do certain things or perform certain functions are *scientific* explanations. Brain states are items postulated to exist and studied by science. But you might wonder why the scientific explanations in question are neuroscientific explanations as opposed to, say, psychological explanations. In my view, this is because, though psychologists study the psychological states of subjects, they remain neutral on the metaphysical nature of these states. If, for example, a psychologist attributes an internal psychological state to a subject in order to explain its behavior, the attributed state, so far as psychology goes, could just as well be a state of a Cartesian substance as a state of the brain. Neuroscientists, however, are not similarly neutral with respect to the internal states they study. They investigate brain states and their properties, and this means that, in order to figure out how we solve neural functional problems, we have to examine how neuroscientists explain why brain states satisfy functional descriptions.

<sup>80</sup> For a thorough account of neuroscientific explanation see Craver (2007).

<sup>81</sup> More precisely, these things are formulated within the context of the background theory, and they are given meaning and sense by being embedded in the theory.

description describes).<sup>82</sup> It is this element of neuroscientific explanations that I want to focus on.

The importance of physical mechanisms to neuroscientific explanations is an instance of a more general feature of biological explanation. In general, biological explanations are distinguished by the importance they place on structure-function relations and physical mechanisms. For example, in order to explain why a protein does what it does—e.g., catalyzes a reaction, transports a molecule, provides structural support to a part of the cell—scientists typically first determine the structure and conformation of the protein, and then they show—via a physical mechanism—how this structure and conformation allows the protein to perform its function. The structure and conformation of the protein, coupled to a physical mechanism, explains why the protein satisfies the functional description in question, and this way of solving functional problems is a hallmark of biological explanations.

To see how this works, suppose that a biologist wants to explain why a protein P is a catalyst of a reaction R, where R is the reaction  $X-Y \rightarrow X + Y$  (X and Y are atoms and X—Y is the molecule consisting of X and Y bonded together by a covalent bond). P, we know, is a catalyst of R iff P lowers the activation energy of R (the activation energy is the energy required to make the reaction happen); and P will lower this energy iff P lowers the energy required to break the bond between X and Y (since the reaction happens iff the bond between X and Y is broken). Hence, P is a catalyst of R iff P makes it easier to break the bond between X and Y (i.e., P lowers

---

<sup>82</sup> Polger and Sufka (2005) and Craver (2007) emphasize the importance of physical mechanisms in neuroscientific explanations. The picture of biological explanation, in general, and neuroscientific explanation, in particular, I discuss in this paper is heavily influenced by Polger and Sufka (2005) and Craver (2007).

the energy required to break the X—Y bond), and what the biologist wants to do is explain why P does this.

How does the biologist proceed? First, the biologist determines (by using, say, x-ray crystallography) the structure and conformation of P. Suppose that the biologist discovers that P is a very large protein containing two adjacent surface indentations into which X and Y could fit. If that were the case, then the biologist could explain why P lowers the activation energy of R by providing a physical mechanism that shows how P makes it easier to break X—Y bond. The mechanism would be as follows. When X—Y and P are mixed together, X—Y falls into the two adjacent surface indentations (X falls into one indentation while Y falls into the other). Because P is flexible (proteins are flexible molecules), it jostles back-and-forth in response to the turbulent flow of fluid in the cell. Thus, when X—Y is inserted into P's indentations, P's movements will pull X and Y further apart. Since what is holding X and Y together is the electromagnetic force, and since this force decreases as the distance between them increases, it should now be easier—i.e., less energy should be required—to break the bond between them. Thus, P lowers the activation energy of R, and this entails that P is a catalyst of R.

Admittedly, this catalyst example is simple and purely hypothetical; nonetheless, it illustrates, on a small scale, the important role structure-function relations and physical mechanisms play in biological explanations—in particular, biological explanations that explain why a biological item (e.g., protein P) satisfies a functional description (e.g., 'lowers the activation energy of R'). Neuroscientific explanations share this general feature of biological explanations, though in the case of neuroscience, the biological item in question may be a complicated brain state in addition to a simple protein. Regardless, structure-function relations

and physical mechanisms play the same important explanatory role. In order to explain why a brain state B does what a functional description describes—e.g., ‘contracts a muscle’—neuroscientists (at the very least) appeal to structure-function relations (e.g., the synaptic connections the neurons making up B make to themselves and other neurons in the central and peripheral nervous system, the line-tracked pathways in the central and peripheral nervous system that the axons of the neurons making up B define, etc.) and specify a physical mechanism (e.g., depolarization events consisting of ion channels opening and closing and ions flowing across membranes, vesicle fusion events consisting of receptors on the vesicle and membrane coming into contact, etc.) that shows how B contracts the muscle. These structure-function relations and the physical mechanism are proposed and appealed to from within the context of a background, anchoring scientific theory; and they are surely more complicated than the simple relations and mechanism present in the case of protein P and reaction R. But the basic idea is the same in the two cases. Put simply, in order to explain why brain states satisfy functional descriptions, neuroscientists (at the very least) appeal to structure-function relations and specify a physical mechanism that shows how the brain state does the thing (or performs the function) the functional description describes.

This provides at least a partial answer to the question we started with. We said that the mind-body problem is the problem of explaining why brain state B satisfies the functional description ‘makes S subjectively aware of perfect blue’. We then wanted to know how to solve this problem (i.e., what is required to solve it). Since the mind-body problem is a neural functional problem, and since neural functional problems are solved (in part) by appealing to structure-function relations and specifying physical mechanisms that show how brain states

satisfy functional descriptions, **to solve the mind-body problem, we must specify a physical mechanism that shows how B makes S subjectively aware of perfect blue.**<sup>83</sup>

#### 4.4 No Physical Mechanism

Can we specify a physical mechanism that shows how B makes S subjectively aware of perfect blue? To answer this question, consider again the simple example involving protein P and reaction R ( $R = X-Y \rightarrow X + Y$ ). In this example, we moved atoms and molecules around, and we posited interactions between these atoms and molecules, and the result of all this was a physical mechanism that showed how P made it easier to break the bond between X and Y. Notably, the reason why it is possible to specify a physical mechanism here is that we have a background theory that enables us to do two things. First, our background theory—along with our knowledge of the structure and conformation of P—allows us to imagine a scenario where P pulls X and Y further apart. Second, our background theory allows us to establish a *conceptual connection* between X and Y being pulled further apart and it being easier to break the bond between X and Y.<sup>84</sup> Because of these two things, it is possible to move atoms and molecules around, and posit interactions between them, so that we end up with a physical mechanism that shows how P lowers the activation energy of R.

---

<sup>83</sup> To make things simple, I'll assume that (i) any physical mechanism that shows how B makes S subjectively aware of perfect blue will appeal to structure-function relations present in B, and (ii) the main obstacle to explaining why B makes S subjectively aware of perfect blue is specifying a physical mechanism that shows how B makes S subjectively aware of perfect blue.

<sup>84</sup> Specifically, our background theory tells us that what holds X and Y together is an electromagnetic bond, and because (according to our background theory) the electromagnetic force between items decreases as the distance between them increases, pulling X and Y further apart will decrease the energy required, i.e., make it easier, to break the bond between them.

Now think about the atoms, ions, molecules, ion channels, membranes, synaptic vesicles, neurotransmitters, etc., making up B, and consider whether there is any way to move these items around and posit interactions between them so that we end up with a physical mechanism that shows how B makes S subjectively aware of perfect blue. It is hard to see what kinds of movements or interactions could work here. Indeed, it is so hard to see what could work here that it might seem like a mistake to even be asking for a physical mechanism. But if the mind-body problem is a neural functional problem, and if neural functional problems are solved (at least in part) by specifying physical mechanisms, then it is no more a mistake to be asking for a physical mechanism here than it is for any other neural functional problem. The only difference is that, in this case, no physical mechanism suggests itself.

A natural response to this line of thought is to say: even though no physical mechanism is currently available, that does not mean that, in the future, science (or philosophy) won't formulate a background theory and physical mechanism capable of showing how B makes S subjectively aware of perfect blue. Though, in general, this kind of deferment to future science to solve current unsolved problems is reasonable, in case of the mind-body problem, there are reasons to be skeptical that a future background theory and physical mechanism will emerge. The problem is that the future background theory must allow us to establish a *conceptual connection* between the state-property *making S subjectively aware of perfect blue* and the torrent of ionic and molecular activity making up B (just as we had a conceptual connection between X and Y being pulled further apart and it being easier to break the bond between X and Y). But it is hard to imagine how we could *conceptually* bridge the gap between making a subject *subjectively aware* of something and particles *moving and interacting* in lawlike ways.

Given that there seems to be no connection at all between the ionic and molecular activity making up B and the state-property of making a subject subjectively aware of perfect blue, there appears to be nothing to base a physical mechanism on. If that is right, then it is hard to accept the claim that a future science (or philosophy) will provide the requisite background theory and physical mechanism.

In light of this, I think it's reasonable to conclude that we can't specify a physical mechanism that shows how B makes S subjectively aware of perfect blue. If that is right, then we can't explain why B makes S subjectively aware of perfect blue. Since the functional description is 'makes S subjectively aware of perfect blue', this implies that we can't explain why B is an experience of blue. As a result, the mind-body problem can't be solved. It is a hard problem.

## 5. Conclusion

Transparency-intentionalists think that, because experiences of blue are transparent, the mind-body problem is the problem of naturalizing intentionality. Moreover, they claim that the mind-body problem is easy because the problem of naturalizing intentionality is easy. I have argued against these philosophers. I have argued that, because the identity conditions of experiences of blue are captured by the functional description 'makes S subjectively aware of perfect blue', transparency-intentionalists face a dilemma. On one hand, in order to ensure that the mind-body problem is the problem of naturalizing intentionality, they must specify a meaning for 'represents' and a semantic value for 'blue' such that 'represents blue' means 'makes S subjectively aware of perfect blue'. But any such specification makes the problem of naturalizing intentionality (= the mind-body problem) hard since we can't specify a physical



mechanism that shows how a brain state makes a subject subjectively aware of perfect blue. On the other hand, any specification of ‘represents’ and ‘blue’ that makes the problem of naturalizing intentionality easy—i.e., that allows us to explain why a brain state represents blue—entails that the mind-body problem is not the problem of naturalizing intentionality since ‘represents blue’ doesn’t mean ‘makes S subjectively aware of perfect blue’. Thus, transparency-intentionalism is false, and the mind-body problem is a hard problem even if we assume that experiences of blue are transparent.

In my view, one reason why this argument against transparency-intentionalism is significant is that it highlights the importance of centering our discussions of the mind-body problem on what we have an independent, non-theoretical grip on with respect to consciousness.<sup>85</sup> Part of the reason why it is easy to be swayed by the thought that, if experiences of blue are transparent, then the mind-body problem is easy is that theoretical terms with unclear meanings—e.g., ‘represents’, ‘qualitative character’, ‘essence’—are often used to characterize experiential properties relevant to the mind-body problem. So, for example, theoretical terms like these are often used to explain what it means for experiences of blue to be transparent (see the examples mentioned in note 60). But this could only lead to a confused understanding of transparency; more important, it will lead to a confused articulation of what the mind-body problem *is*. If we want to solve the mind-body problem, we have to accurately characterize it *and* the experiential properties we think are relevant to it—e.g., *being transparent*. I think that this can only be done by getting clear on what we have an independent, non-theoretical grip on with respect to, e.g., experiences of blue. In my view,

---

<sup>85</sup> See Speaks (2015) and Tye (2015) for further discussion of this issue.

what we have an independent grip on is that these experiences make us subjectively aware of perfect blue, and I have tried to explain what that means by pointing to certain key examples. As I see things, this is the only way that we can get clear on what the mind-body problem is.

## References

- Armstrong, David. *A Materialist Theory of the Mind*. New York, Humanities Press: 1968.
- Aydede, Murat. "The Main Difficulty with Pain," in *Pain: New Essays on Its Nature and the Methodology of Its Study*. Ed. Murat Aydede (MIT Press, 2005): 123-136.
- Aydede, Murat. "Pain." *The Stanford Encyclopedia of Philosophy* Ed. Edward N. Zalta (Spring 2013 Edition), URL = <<http://plato.stanford.edu/archives/spr2013/entries/pain/>>.
- Bain, David. "The Location of Pains." *Philosophical Papers* 36:2 (2007): 171-205.
- Balog, Katalin. "Phenomenal Concepts," in *The Oxford Handbook of Philosophy of Mind*. Eds. Brian P. McLaughlin, Ansgar Beckermann, and Sven Walter (OUP, 2009): 292-312.
- Balog, Katalin. "In Defense of the Phenomenal Concept Strategy." *Philosophy and Phenomenological Research* Vol. LXXXIV, No. 1 (January 2012a): 1-23.
- Balog, Katalin. "Acquaintance and the Mind-Body Problem," in *New Perspectives on Type Identity: the Mental and the Physical*. Eds. Simone Gozzano and Christopher S. Hill (Cambridge University Press, 2012b): 16-42.
- Batty, Clare. "A Representational Account of Olfactory Experience." *Canadian Journal of Philosophy* 40:4 (2010): 511-538.
- Block, Ned. "Bodily Sensations as an Obstacle for Representationism," in *Pain: New Essays on Its Nature and the Methodology of Its Study*. Ed. Murat Aydede (MIT Press, 2005): 137-142.
- Block, Ned. "Max Black's Objection to Mind-Body Identity," in *Phenomenal Concepts and Phenomenal Knowledge: New Essays on Consciousness and Physicalism*. Eds. Torin Alter and Sven Walter (OUP, 2007): 249-306.
- Block, Ned and Stalnaker, Robert. "Conceptual Analysis, Dualism, and the Explanatory Gap." *The Philosophical Review* Vol. 108, No. 1 (January 1999): 1-46.
- Burge, Tyler. *Origins of Objectivity*. New York, OUP: 2010.
- Byrne, Alex. "Intentionalism Defended." *The Philosophical Review* Vol. 110, No. 2 (April 2001): 199-24
- Byrne, Alex. "Color and the Mind-Body Problem." *dialectica* Vol. 60, No. 3 (2006): 223-244.
- Byrne, Alex. "Experience and Content." *The Philosophical Quarterly* Vol. 59, No. 236 (July 2009a): 429-451.

- Byrne, Alex. "Sensory Qualities, Sensible Qualities, Sensational Qualities," in *The Oxford Handbook of Philosophy of Mind*. Eds. Brian P. McLaughlin, Ansgar Beckermann, and Sven Walter (OUP, 2009b): 268-280.
- Byrne, Alex and Hilbert, David R. "Color Primitivism." *Erkenntnis* 66 (2007): 73-105.
- Carruthers, Peter and Veillet, Bénédicte. "The Phenomenal Concept Strategy." *Journal of Consciousness Studies* 14, No. 9-10 (2007): 212-236.
- Chalmers, David. "Facing up to the Hard Problem of Consciousness." *Journal of Consciousness Studies* 2, No.3 (1995): 200-219.
- Chalmers, David. "The Content and Epistemology of Phenomenal Belief," in *Consciousness: New Philosophical Perspectives*. Eds. Quentin Smith & Aleksandar Jokic (OUP, 2003): 220-272.
- Chalmers, David. "Perception and the Fall from Eden," in *Perceptual Experience*. Eds. Tamar Szabó Gendler and John Hawthorne (OUP, 2006): 49-125.
- Chalmers, David. "Phenomenal Concepts and the Explanatory Gap," in *Phenomenal Concepts and Phenomenal Knowledge: New Essays on Consciousness and Physicalism*. Eds. Torin Alter and Sven Walter (OUP, 2007): 167-194.
- Chalmers, David and Jackson, Frank. "Conceptual Analysis and Reductive Explanation." *Philosophical Review* 110, 3 (2001): 315-361.
- Crane, Tim. "Intentionalism," in *The Oxford Handbook of Philosophy of Mind*. Eds. Brian P. McLaughlin, Ansgar Beckermann, and Sven Walter (OUP, 2009): 474-492.
- Craver, Carl. *Explaining the Brain: Mechanisms and the Mosaic Unity of Neuroscience*. New York, OUP: 2007.
- Descartes, Rene. *Meditations on First Philosophy: Selections from the Objections and Replies*. Ed. John Cottingham. New York, Cambridge University Press: 1996.
- Dretske, Fred. "If You Make One, You Don't Know How it Works." *Midwest Studies in Philosophy*. Eds. P. French, T. Uehling, and H. Wettstein, Vol. 19 (1994): 468-482.
- Dretske, Fred. *Naturalizing the Mind*. Cambridge, MIT Press: 1995.
- Eddington, Arthur. *The Nature of the Physical World*. New York, Macmillan: 1928.  
<<https://archive.org/details/natureofphysical00eddi>>

- Harman, Gilbert. "The Intrinsic Quality of Experience." *Philosophical Perspectives*, Vol. 4, Action Theory and Philosophy of Mind (1990): 31-52.
- Hill, Christopher S. "Ow! The Paradox of Pain," in *Pain: New Essays on Its Nature and the Methodology of Its Study*. Ed. Murat Aydede (MIT Press, 2005): 75-98.
- Hill, Christopher S. and McLaughlin, Brian P. "There are Fewer Things in Reality Than are Dreamt of in Chalmers's Philosophy." *Philosophy and Phenomenological Research* Vol. 59, No. 2 (June 1999): 445-454.
- Jackson, Frank. *Perception: A Representative Theory*. New York: Cambridge University Press, 1977.
- Jackson, Frank. "Epiphenomenal Qualia". *The Philosophical Quarterly* Vol. 32, No. 127 (April 1982): 127-136.
- Jackson, Frank. "Finding the Mind in the Natural World" in *Philosophy and the Cognitive Sciences*. Eds. Roberto Casati, B. Smith & Stephen L. White, (Holder-Pichler-Tempsky, 1994): 227-49.
- Jackson, Frank. "The Knowledge Argument, Diaphanousness, Representationalism," in *Phenomenal Concepts and Phenomenal Knowledge: New Essays on Consciousness and Physicalism*. Eds. Torin Alter and Sven Walter (OUP, 2007): 52-64.
- Kind, Amy. "What's So Transparent About Transparency?" *Philosophical Studies* 115 (2003): 225-244.
- Kind, Amy. "The Case Against Representationalism About Moods," in *Current Controversies in Philosophy of Mind*. Ed. Uriah Kriegel (Routledge, 2014): 113-134.
- Kripke, Saul. *Naming and Necessity*. Cambridge, HUP: 1980.
- Levin, Janet. "Taking Type-B Materialism Seriously." *Mind & Language* Vol. 23, No. 4 (September 2008): 402-425.
- Levine, Joseph. "On Leaving What it's Like Out," in *Consciousness: Psychological and Philosophical Essays*. Eds. Martin Davies & Glyn W. Humphreys (MIT Press, 1993): 543-557.
- Levine, Joseph. "Color and Color Experience: Colors as Ways of Appearing." *dialectica* Vol. 60, No. 3 (2006): 269-282.

- Levine, Joseph. "Phenomenal Concepts and the Materialist Constraint," in *Phenomenal Concepts and Phenomenal Knowledge: New Essays on Consciousness and Physicalism*. Eds. Torin Alter and Sven Walter (OUP, 2007): 145-166.
- Levine, Joseph. "Demonstrative Thought." *Mind & Language* Vol. 25, No. 2 (April 2010): 169-195.
- Lewis, David. "Mad Pain and Martian Pain," in *Readings in the Philosophy of Psychology*. Ed. Ned Block, Vol. 1 (HUP, 1980): 216-22.
- Loar, Brian. "Phenomenal States," in *The Nature of Consciousness: Philosophical Debates*. Eds. Ned Block, Owen Flanagan & Güven Güzeldere (MIT Press, 1997): 597-616. URL = <<http://www.nyu.edu/gsas/dept/philo/courses/consciousness97/papers/loar.html>>
- McGinn, Colin. "Can We Solve the Mind-Body Problem?" *Mind* New Series, Vol. 98., No. 391 (July 1989): 349-366.
- McLaughlin, Brian. "Phenomenal Concepts and the Defense of Materialism." *Philosophy and Phenomenological Research* Vol. LXXXIV, No. 1 (January 2012): 206-214.
- Mendelovici, Angela. "Pure Intentionalism About Moods," in *Current Controversies in Philosophy of Mind*. Ed. Uriah Kriegel (Routledge, 2014): 135-157.
- Nagel, Thomas. "What is it Like to be a Bat?" *Philosophical Review* Vol. 83, No. 4 (October 1974): 435-450.
- Papineau, David. *Thinking About Consciousness*. New York, OUP: 2002.
- Papineau, David. "Phenomenal and Perceptual Concepts," in *Phenomenal Concepts and Phenomenal Knowledge: New Essays on Consciousness and Physicalism*. Eds. Torin Alter and Sven Walter (OUP, 2007): 111-144.
- Papineau, David. "What Exactly is the Explanatory Gap." *Philosophia* 39 (2011): 5-19.
- Pautz, Adam. "Intentionalism and Perceptual Presence." *Philosophical Perspectives* 21, *Philosophy of Mind* (2007): 495-541.
- Perry, John. *Knowledge, Possibility, and Consciousness*. Cambridge, MIT Press: 2001.
- Polger, Thomas W. and Sufka, Kenneth J. "Closing the Gap on Pain: Mechanism, Theory, and Fit," in *Pain: New Essays on Its Nature and the Methodology of Its Study*. Ed. Murat Aydede (MIT Press, 2005): 325-350.
- Russell, Bertrand. *The Analysis of Matter*. Spokesman: 1927/2007.

- Russell, Bertrand. "Mind and Matter," in *Portraits from Memory*. New York, Simon and Schuster: 1950.
- Searle, "Minds, Brains, and Programs." *Behavioral and Brain Sciences* Volume 3, Issue 3 (September 1980): 417-457.
- Searle, John R. *The Rediscovery of Mind*. Cambridge, MIT Press: 1994.
- Sellars, Wilfrid. "Philosophy and the Scientific Image of Man," in *Empiricism and the Philosophy of Mind*. (Routledge & Kegan Paul Ltd, 1963a): 1-40.
- Sellars, Wilfrid. "Empiricism and the Philosophy of Mind," in *Science, Perception and Reality*. (Routledge & Kegan Paul, 1963b): 127-196.
- Shapiro, Lisa. "Elisabeth, Princess of Bohemia." *The Stanford Encyclopedia of Philosophy* Ed. Edward N. Zalta (Winter 2014 Edition) URL = [<https://plato.stanford.edu/archives/win2014/entries/elisabeth-bohemia/>](https://plato.stanford.edu/archives/win2014/entries/elisabeth-bohemia/).
- Shoemaker, Sydney. "Self-Knowledge and 'Inner Sense': Lecture III: The Phenomenal Character of Experience." *Philosophy and Phenomenological Research* Vol. 54, No. 2 (June 1994): 291-314.
- Shoemaker, Sydney. "On the Way Things Appear," in *Perceptual Experience*. Eds. Tamar Szabó Gendler and John Hawthorne (OUP, 2006): 461-480.
- Siegel, Susanna. "The Contents of Perception." *The Stanford Encyclopedia of Philosophy* Ed. Edward N. Zalta (Winter 2016 Edition) URL = [<https://plato.stanford.edu/archives/win2016/entries/perception-contents/>](https://plato.stanford.edu/archives/win2016/entries/perception-contents/).
- Speaks, Jeff. "Is Phenomenal Character Out There in the World?" *Philosophy and Phenomenological Research* Vol. XCI, No. 2 (September 2015): 465-482.
- Stoljar, Daniel. "Two Conceptions of the Physical." *Philosophy and Phenomenological Research* 62 (2001): 253-281.
- Stoljar, Daniel. "The Argument from Diaphanousness." *Canadian Journal of Philosophy* 34:sup1 (2004) 341-390.
- Stoljar, Daniel. "Physicalism and Phenomenal Concepts." *Mind & Language* Vol. 20, No. 5 (November 2005): 469-494.
- Stoljar, Daniel. "The Argument from Revelation" in *Conceptual Analysis and Philosophical Naturalism*. Eds. David Braddon-Mitchell and Robert Nola (MIT Press, 2009): 113-138.

Stoljar, Daniel. "The Semantics of 'What it's like' and the Nature of Consciousness." *Mind* Vol 125.500 (October 2016): 1161-1198.

Stubenberg, Leopold. "Austria vs. Australia: Two Versions of the Identity Theory," in *Austrian Philosophy Past and Present: Essays in Honor of Rudolf Haller*. Eds. Keith Lehrer & Johann Christian Marek (Kluwer Academic Publishers, 1997): 125-146.

Sundström, Pär. "Phenomenal Concepts." *Philosophy Compass* 6/4 (2011): 267-281.

Tye, Michael. "Phenomenal Consciousness: The Explanatory Gap as a Cognitive Illusion." *Mind* New Series, Vol. 108, No. 432 (October 1999): 705-725.

Tye, Michael. "Representationalism and the Transparency of Experience." *NOÛS* 36:1 (2002): 137-151.

Tye, Michael. "Another Look at Representationalism About Pain," in *Pain: New Essays on Its Nature and the Methodology of Its Study*. Ed. Murat Aydede (MIT Press, 2005a): 99-120.

Tye, Michael. "In Defense of Representationalism: Reply to Commentaries," in *Pain: New Essays on Its Nature and the Methodology of Its Study*. Ed. Murat Aydede (MIT Press, 2005b): 163-176.

Tye, Michael. "Representationalist Theories of Consciousness," in *The Oxford Handbook of Philosophy of Mind*. Eds. Brian P. McLaughlin, Ansgar Beckermann, and Sven Walter (OUP, 2009): 253-267.

Tye, Michael. "Yes, Phenomenal Character Really is Out There in the World." *Philosophy and Phenomenological Research* Vol. XCI, No. 2 (September 2015): 483-488.