



**Language on Holiday and the Philosophy of Mind: A  
Linguistically Sensitive Approach to Phenomenal  
Consciousness, Pain, and Psychological Predicates**

Erlend Winderen Finke Owesen

Robinson College,  
University of Cambridge

July 2021

This thesis is submitted for the degree of Doctor of Philosophy

## **Declaration**

This thesis is the result of my own work and includes nothing which is the outcome of work done in collaboration except as declared in the preface and specified in the text. It is not substantially the same as any work that has already been submitted before for any degree or other qualification except as declared in the preface and specified in the text. It does not exceed the prescribed word limit for the Department of History and Philosophy Degree Committee.

## **Statement of Length**

This thesis, excluding the bibliography, contains 65,994 words.

## Abstract

The aim of this thesis is to contribute to debates about three topics in the philosophy of mind: phenomenal consciousness, pain, and the extension of psychological predicates. Chapter 1 outlines my metaphilosophical views and gives a roadmap to the thesis. Chapter 2 argues that the ‘what-it’s-like’ phrase as it is characteristically used in the literature on phenomenal consciousness has a technical meaning. I argue that this has the consequence that the phrase says nothing informative about phenomenal consciousness and that lay people’s use of the phrase does not show that they believe in phenomenal consciousness. Chapter 3 argues that eliminativism about phenomenal consciousness is a view worth taking seriously. First, I clarify the eliminativist position, then I explain its motivation. Finally, I draw on the discussion in chapter 2 and argue that there is no compelling evidence for lay belief in phenomenal consciousness, hence that we are currently not in a position to say whether common sense counts in favour of eliminativism or realism. Chapter 4 addresses the main problem for eliminativism – the problem of explaining why people believe in phenomenality. I discuss four recently proposed theories and conclude that we are currently not in a position to tell which of these is the most promising eliminativist-friendly explanation of belief in phenomenality, but that the potential of (at least some of) these theories confirms the claim that eliminativism is a view worth taking seriously. Chapter 5 defends the Bodily Theory of pain, according to which pains are bodily occurrences located in an extra-cranial body part, which contrasts with the Experiential Theory, according to which pains are experiences located in the mind or brain. Chapter 6<sup>1</sup> discusses what a defender of the Bodily Theory should say about the pain-in-mouth argument, i.e. the step from (1) There is a pain in my finger, and (2) My finger is in my mouth, to (3) There is a pain in my mouth. Several accounts have been offered to explain why (1)-(3) sounds wrong. In chapter 6 I offer a novel account – the mereological view – that entails the Bodily Theory of pain. Chapter 7 discusses cognitive scientists’ ascription of psychological predicates to the brain, i.e. when cognitive scientists say things like ‘The brain thinks’. I propose the Synecdoche View, according to which the locutions of cognitive scientists are figurative, with ‘the brain’ referring to the human being, such that ‘The brain thinks’ reports the thinking of the human being, not the thinking of the brain. One consequence of this is that the locutions of cognitive scientists offer no reason to believe that psychological predicates extend to brains. Chapter 8 concludes the thesis by offering a speculative error theory, according to which some of my opponents’ views are based on conceptual conflation.

---

<sup>1</sup> This chapter is based on an unpublished manuscript co-written by Arif Ahmed and me.

## **Acknowledgements**

I am indebted to many people for this thesis. Thanks to Marta Halina, my supervisor, for insightful guidance, feedback, support, encouragement and kindness. Her academic rigour was invaluable to get this thesis to its present shape.

Thanks to Arif Ahmed, my co-supervisor, for taking me seriously and objecting to virtually every claim I made, for collaborating on the pain-in-mouth paper, and for teaching me how to do philosophy. There is no philosopher I have learned more from than him, and his intellectual influence on this thesis is beyond what I can say.

Thanks to Anna Alexandrova, Murat Aydede, Lukas Beck, Adam Bradley, Luke Cash, David Chalmers, Tim Crane, Ståle Finke, Ludvig Fæhn Fuglestad, Peter Hacker, Richard Holton, Colin Klein, Michelle Liu, Tom McClelland, Daniel Ott, Guy Sechrist, Henry Taylor, Oscar Westerblad, audiences at the Joint Session 2019 and 2020, participants at the Mental Sciences Club (Cambridge), and participants at Oslo Mind Group for helpful comments on, and discussion of, ideas in this thesis.

Finally, thanks to Kjell Inge Røkke and Aker Scholarship for generously funding my PhD.

# Table of Contents

<b>Chapter 1: Introduction</b>	<b>9</b>
1. Semantic Externalism, Conceptual Freedom and Conceptual Pluralism	10
2. Conceptual Confusions	11
2.1. Verbal Disputes	11
2.2. Conceptual Conflations	14
3. Monism	15
4. The Linguistically Sensitive Approach	17
5. Phenomenal Consciousness	17
5.1. Chapter 2: What-it's-Like Talk is Technical Talk	18
5.2. Chapter 3: Eliminativism About Phenomenal Consciousness	19
5.3. Chapter 4: Why Do People Believe in Phenomenality?	19
6. Pain	20
6.1. Chapter 5: The Bodily Theory of Pain	20
6.2. Chapter 6: The Bodily Theory and the Pain-in-Mouth Argument	20
7. The Extension of Psychological Predicates and Conclusion	21
7.1. Chapter 7: Does the Brain Think?	21
7.2. Chapter 8: Conclusion	22
<b>Chapter 2: What-it's-Like Talk is Technical Talk</b>	<b>23</b>
1. Introduction	23
2. The Technical View and the Non-Technical View	25
3. The Ability Sense	29
3.1. Why the Ability Sense is an Ordinary Sense	30
3.2. The Ability Sense and Knowing What It's Like	34
4. How Should We Interpret Lay People's Use of the 'What-It's-Like' Phrase?	36
5. The Arguments for the Technical View	42
5.1. First Argument	42
5.2. Second Argument	44
6. The Arguments for the Non-Technical View and Why They are Unconvincing	46
6.1. First Argument	47
6.2. Second Argument	47

6.3. Third Argument	48
7. Why the Technical View Matters	50
7.1. First Reason	50
7.2. Second Reason	53
8. Conclusion	53
<b>Chapter 3: Eliminativism About Phenomenal Consciousness</b>	<b>54</b>
1. Introduction	54
2. What Is Eliminativism About Phenomenality?	55
2.1. Defining Eliminativism	55
2.2. Clarifications	59
3. The Motivation for Eliminativism	61
3.1. Eliminativism vs. Non-Reductive Realism	61
3.2. Eliminativism vs. Reductive Realism	65
4. How Widespread is Belief in Phenomenality?	66
5. Conclusion	72
<b>Chapter 4: Why Do People Believe in Phenomenality?</b>	<b>73</b>
1. Introduction	73
2. Darwinism	74
2.1. The Theory	74
2.2. Challenges with Darwinism	75
3. Illusionism	75
3.1. The Theory	76
3.2. Challenges with Illusionism	77
4. Inferentialism	79
4.1. The Theory	79
4.2. Challenges with Inferentialism	81
5. Educationism	82
5.1. Narrow Educationism	82
5.2. Wide Educationism	86
5.3. Challenges with Educationism	86
6. Conclusion	87

<b>Chapter 5: The Bodily Theory of Pain</b>	<b>89</b>
1. Introduction	89
2. Different Senses of ‘Pain’ and Why the Debate is Non-Verbal	91
2.1. Different Senses of ‘Pain’	91
2.2. Why the Debate is Non-Verbal	93
3. The Bodily Theory	95
3.1. Pains as Proximal Activations of Nociceptors	96
3.2. The Experience-Dependence of Pains	99
3.3. Objections/Clarifications	100
4. Arguments for the Bodily Theory	101
4.1. The Common Sense Argument for the Bodily Theory	101
4.2. The Semantic Argument for the Bodily Theory	103
5. Arguments for the Experiential Theory and Why They are Unconvincing	104
5.1. The Common Sense Argument for the Experiential Theory	105
5.2. The Phantom Limb Argument for the Experiential Theory	108
5.3. The Scientific Argument for the Experiential Theory	108
6. Conclusion	109
<b>Chapter 6: The Bodily Theory and the Pain-in-Mouth Argument</b>	<b>110</b>
1. Introduction	110
2. Existing Views of the Pain-in-Mouth Argument	112
2.1. The Representational View	113
2.2. The Predicative View	114
2.3. The Implicature View	117
3. The Mereological View	118
4. The Mereological View vs. The Predicative View	119
5. The Mereological View vs. The Implicature View	121
6. Conclusion	125
<b>Chapter 7: Does the Brain Think?</b>	<b>126</b>
1. Introduction	126
2. The Nonsense View	130
3. The Literal View	132
4. Five Assumptions	133

5. Why Do Cognitive Scientists Say ‘The Brain Thinks’?	136
6. The Synecdoche View	139
7. The Synecdoche View vs. the Nonsense View and the Literal View	142
8. Conclusion	146
<b>Chapter 8: Conclusion</b>	<b>148</b>
1. Introduction	148
2. The Speculative Error Theory	149
3. Two Objections	153
4. Conclusion	155
<b>Bibliography</b>	<b>156</b>



# Chapter 1

## Introduction

The aim of this thesis is to make contributions to the philosophy of mind. But even though the aim is to contribute to the philosophy of mind, I often talk about words and how they are used. When I talk about words, I mention them rather than use them, with quotation marks indicating mention rather than use.<sup>2</sup> When I use words, I talk about what the words refer to, not the words themselves. For example, I distinguish different senses of the word ‘pain’ and say that I am only concerned with the phenomenon pain in one sense of the word and not pain in other senses of the word. The reason I so often focus on words and how they are used is that I have certain metaphilosophical views.

The first part of this introductory chapter provides an overview of my metaphilosophical views, which I call the linguistically sensitive approach. This approach consists of committing to (a) semantic externalism, conceptual freedom and conceptual pluralism, and (b) trying to avoid conceptual confusions (verbal disputes and conceptual conflations). §1-4 of this chapter explains these ideas and how they are connected. But some caveats are appropriate. Firstly, I do not commit to any particular theory of concepts or meaning. Rather, I mostly leave these notions intuitive when explaining (a) and (b). Given my conceptual pluralism (see below), I think there are several senses of ‘concept’ and ‘meaning’, so on some occasions where it is particularly useful, I specify further what I have in mind when using these terms. Secondly, I do not go into all the details surrounding (a) and (b) but introduce the ideas as they are relevant for my purposes. Further details can be found in the literature to which I refer. Thirdly, I do not give a defence of (a) and (b) but take them for granted. The point of introducing them here is to prepare the reader for other ideas in the philosophy of mind for which I do give a defence in the main body of the thesis (chapters 2-7).

The second part of this introductory chapter gives a roadmap to the rest of the thesis and explains how my metaphilosophical views determine how I navigate in philosophical debates. To this end, §5 covers the chapters that concern phenomenal consciousness, §6 covers the chapters that concern pain, and §7 covers the chapter that concerns the extension of psychological predicates and the conclusion of the thesis.

---

<sup>2</sup> I also use quotation marks as scare quotes, but when I do so it is clear from the context.

## 1. Semantic Externalism, Conceptual Freedom and Conceptual Pluralism

My metaphilosophy commits me to three claims: semantic externalism, conceptual freedom and conceptual pluralism.

The idea of *semantic externalism* is the idea that things external to a speaker's mind, such as the use of words or the nature of reality, determine facts about the meaning of the words of the speaker.<sup>3</sup> This minimal characterization of externalism does not commit me to all claims semantic externalists have defended. For example, I am not committed to denying that we have access to or control over meaning (e.g. Williamson 1994). The above idea of semantic externalism is consistent with the claim that if I use a word a certain way, then my using it that way is what ensures it has the semantic meaning it has, and my using it that way is something to which I have epistemic access. Let's say that I stipulate that what I mean by 'banana' is what other people call 'apple' and I consistently call apples 'bananas'. Semantic externalism is consistent with describing this as me introducing a new meaning of 'banana'. So externalism does not entail that we have no control over meaning or that I cannot know what the meanings of the words I utter are. But if I start claiming that 'banana' in my mouth refers to (what other people call) bananas and not apples, then semantic externalism predicts that I am wrong, since there are things external to my mind, such as my use of 'banana', that determine what 'banana' means in my mouth.

The idea of *conceptual freedom* is the idea that language puts no constraints on philosophy, or that one's use of words in philosophy is not constrained by anything. Put simply: we can speak however we want! Even if you use 'banana' to pick out bananas, I am free to use 'banana' to pick out apples. Of course, if I want to talk about bananas in the sense in which you use the term 'banana', then my use of 'banana' is constrained by something, namely your use of 'banana'. In this case, language *does* put constraints on my philosophy about bananas. Setting special cases of semantic externalism aside, I cannot both talk about bananas in the sense in which you talk about bananas, and use 'banana' to refer to apples, when you use 'banana' to refer to bananas and not apples. But given conceptual freedom, nothing turns on what we call 'bananas', and the fact that you call bananas 'bananas' does not prevent me from theorizing about apples while calling them 'bananas'.

The idea of *conceptual pluralism* is the idea that for many important expressions in philosophy there corresponds different concepts, senses, and properties (Chalmers 2011: 539). Some of the different senses, concepts and properties may be associated with the term pre-

---

<sup>3</sup> See Putnam (1975), Burge (1979) and Kripke (1980) for more on externalism and its motivation.

theoretically, others may be added during theorizing. Whether the origin of the association is pre-theoretical or theoretical does not matter. In the example above, we might say that there are different concepts, senses and things corresponding to the single expression 'banana'. So pluralism, in this imagined case, is true for 'banana'.

Conceptual freedom and conceptual pluralism are not the same. Conceptual freedom is a normative thesis saying that you are allowed to associate whatever concepts, senses and properties you like with an expression in philosophy. Conceptual pluralism is an empirical thesis saying that there are in fact different concepts, senses or properties associated with many important expressions in philosophy.

## **2. Conceptual Confusions**

The combination of semantic externalism, conceptual freedom and conceptual pluralism has a downside, namely *conceptual confusions*. These are misunderstandings that crucially depend on meanings or the use of words. In particular, there are two kinds of conceptual confusions that are important for my metaphilosophical views: verbal disputes and conceptual conflations.

### *2.1. Verbal Disputes*

A dispute is verbal if the parties of a dispute associate different concepts with an expression in the sentence about which they disagree, and the dispute is a result of their concepts diverging. Put in terms of meaning: a dispute is verbal if the parties mean different things with an expression in the sentence about which they disagree, and the dispute is a result of their meaning different things with that expression.

This is how the idea of verbal disputes is related to the ideas of conceptual pluralism, conceptual freedom and semantic externalism. For there to be a verbal dispute with respect to an expression E, conceptual pluralism must be true with respect to E, i.e. it must be true that E has different meanings. So the idea of verbal disputes depends on the idea of conceptual pluralism. Even though the idea of verbal disputes does not depend on the ideas of conceptual freedom and semantic externalism, the reason pluralism is true with respect to an expression E may often be conceptual freedom, in combination with semantic externalism, regarding E. That is, the reason pluralism is true with respect to E may often be that theorists use E in a peculiar way, which, given conceptual freedom, is perfectly valid, and given externalism, creates a distinct sense of E (determined by the peculiar use of E, which is external to the speaker's mind). So even though one only needs the idea of pluralism for the idea of verbal disputes, in

practice, the reason pluralism is true for a given expression may be conceptual freedom, in combination with semantic externalism, regarding that expression.

Some verbal disputes have important practical consequences and are not naturally described as ‘conceptual confusions’. For example, disputes about ‘marriage’ and ‘murder’ may have consequences for whether same-sex partners can marry and what conviction someone might get (Chalmers 2011: 516). Other verbal disputes have no important consequences. Chalmers calls the latter ‘mere verbal disputes’ (2011: 517). Verbal disputes of the first kind are important, but mere verbal disputes are pointless (2011: 520), since there is no point engaging in a dispute about the meaning of a word in a case where one agrees about the non-linguistic facts and the meaning of the relevant word has no important consequences. One does not achieve anything with mere verbal disputes, hence the pointlessness.

In addition to being pointless, mere verbal disputes typically involve defective communication. As Chalmers says: ‘[m]ere verbal disputes are usually impediments to understanding. In effect, they are obstacles that we do better to move beyond, in order that we can focus on the substantive issues regarding a domain’ (2011: 517). The point is that when disputants do not notice that their dispute is merely verbal, the result is typically not just pointlessness but also defective communication.

Consider an example. Let’s say that you claim that Mary is going to the bank, and I claim that Mary is not going to the bank, but what you mean by ‘bank’ is the financial institution and what I mean is the riverbank and we both believe that Mary *is* going to the financial institution and that she is *not* going to the riverbank. In this case, our dispute seems to be a waste of time, as we agree about all the non-linguistic facts and do not achieve anything with our dispute. In other words, our dispute is pointless. But our dispute is not just pointless, it also involves defective communication, as our dispute impedes us from realizing that we actually agree about where Mary is going. It is the pointlessness and typical defective communication that comes with mere verbal disputes that makes them apt for the label ‘conceptual confusions’.

How do we know whether a dispute is verbal? Chalmers (2011) introduces the method of elimination: eliminate use of the key expression in the sentence about which the parties (allegedly) disagree and then determine whether there remains a substantive (i.e. non-verbal) dispute (2011: 526). If no substantive dispute remains after eliminating the key expression, then this is evidence that the dispute is verbal. For example, when we disagree whether Mary is going to the bank, then we might try to state our respective theses without using the word ‘bank’. We will then see that your claim that Mary is going to the financial institution is perfectly

compatible with my claim that she is not going to the place next to the river, and since we actually agree about where Mary is going, this is evidence that our dispute was verbal.

A more philosophical example Chalmers mentions is that of free will. When a compatibilist says that free will is compatible with determinism and an incompatibilist denies this, one might ask them to state their theses without using the term ‘free will’. It might turn out that the parties disagree about a substantive issue, for example, whether moral responsibility is compatible with determinism. (Though whether this is substantive requires either that a substantive dispute remains if one reapplies the method of elimination to the dispute about moral responsibility, or that the concept of moral responsibility is primitive and cannot be clarified in more basic terms.) But it might also turn out that no substantive dispute remains, for example because what the compatibilist means by ‘free will’ is the ability to do what one wants, and what the incompatibilist means is the ability to ultimately originate one’s choices. If so, this is evidence that the dispute is verbal (Chalmers 2011: 530-532). In both the non-philosophical and the philosophical example, we see how the method of elimination is a tool for both identifying and resolving verbal disputes.

One final point before I move to conceptual confluences. According to Chalmers, the danger of verbal disputes should make us wary of questions of the form ‘What is X?’. He writes:

On the picture I favor, instead of asking “What is X ?,” one should focus on the roles one wants X to play and see what can play that role. The roles in question here may in principle be properties of all sorts: so one focuses on the properties one wants X to have and figures out what has those properties. But very frequently, they will be causal roles, normative roles, and especially explanatory roles (2011: 538).

I agree that we should be wary of questions of the form ‘What is X?’. But asking ‘What plays role R?’ rather than asking ‘What is X?’ – where X is specified as that which plays role R – is a mere verbal difference, so I see no reason to prefer the former to the latter question, and I shall occasionally let my discussion be driven by questions of the latter type. For example, in chapter 5 I defend a theory of pain. However, I do not simply ask ‘What are pains?’ and then go on to give a determinate answer (which, in my opinion, occurs too often in the pain literature). Rather, I distinguish senses of ‘pain’ and narrow my discussion to pain in the locatable sense, i.e. the referent of ‘pain’ in sentences like ‘There is a pain in my foot’. This specifies a role for pain (in the locatable sense) – namely, being the referent of ‘pain’ – such

that one can discuss what pains (in the locatable sense) are without engaging in verbal dispute (at least with respect to ‘pain’).

## 2.2. *Conceptual Conflations*

To engage in mere verbal disputes is only one way of being conceptually confused. A different kind of conceptual confusion is to conflate one meaning or concept associated with an expression with a different meaning or concept associated with the same expression. In the imagined example above where I am wrong about my own peculiar use of ‘banana’, my claim that ‘banana’ in my mouth refers to (what other people call) bananas and not apples is based on a conceptual conflation, namely that of my novel concept and the ordinary concept of banana. In other words, my claim is based on conflating the novel stipulated meaning and the ordinary meaning of ‘banana’. (I discuss possible philosophical examples of conceptual conflations in chapter 8.)

The connection with conceptual pluralism, conceptual freedom and semantic externalism is the same for conceptual conflations as for verbal disputes. Only conceptual pluralism with respect to an expression E is necessary to have a conceptual conflation with respect to E. But even though only pluralism is strictly necessary, in practice, the reason pluralism holds for an expression E is often conceptual freedom and semantic externalism regarding E. For example, the reason pluralism holds with respect to ‘banana’ in the example above is that I use ‘banana’ in a peculiar way, which, given conceptual freedom, is perfectly valid, and given externalism, creates a distinct meaning of ‘banana’. So in this case, it is conceptual freedom and semantic externalism that give rise to the conceptual pluralism, which again gives rise to the conceptual conflation.

Whereas verbal disputes may lead to defective communication, conceptual conflations may lead to confused questions and false claims about a domain. If we use the banana example again, my conceptual conflation may make me wonder why people say that bananas grow together in continuous bunches while I have only seen bananas growing separated from other bananas. This would be a confused question based on conflating the ordinary sense, and my peculiar sense, of ‘banana’. Similarly, if I claim that ‘banana’ as it is ordinarily used (in e.g. ‘banana split’) refers to what I call ‘bananas’, namely apples, i.e. that bananas in the ordinary sense of the word are apples, then I make a false claim about what bananas (in the ordinary sense) are.

To conflate concepts in this way is to be wrong about the meaning of an expression. Importantly, the claim that people can be wrong about meaning in this way is not about *speaker-*

*meaning*. Speaker-meaning is what a speaker intends to communicate with a sentence, or refer to with a term, so speaker-meaning is determined by the speaker's mind, such that there seems to be no room for an appearance/reality distinction. What the speaker thinks he means *is* the speaker-meaning. But given semantic externalism, the references of terms and truth-conditions of sentences – which we may call *semantic meanings* – are not determined by a speaker's mind in the way speaker-meaning is. This is why people can be wrong about meaning.<sup>4</sup>

### 3. Monism

There is a complication that is worth mention. I said that conceptual confusions with respect to an expression E depends on conceptual pluralism with respect to E. But Chalmers points out that even in cases where what I call 'monism' holds for an expression E, i.e. when pluralism does not hold, one can still have disputes with very similar properties to those I call 'verbal' (2011: 520). In fact, Chalmers himself wants a notion of verbal disputes that accommodates cases of monism. He does this by replacing talk of meanings with talk of beliefs about meanings. Thus, Chalmers offers the following characterization of a verbal dispute: 'A dispute over [a sentence] S is (broadly) verbal when, for some expression T in S, the parties disagree about the meaning of T, and the dispute over S arises wholly in virtue of this disagreement regarding T' (2011: 522). What Chalmers means by 'disagree about the meaning of T' is that the parties have conflicting beliefs about what a key expression in the sentence means. His point is that even if we grant monism with respect to an expression E that is involved in a dispute, i.e. that the parties use E with the same sense, we can still say that the dispute is 'broadly verbal'. The reason is that (a) the dispute is explained by the disputants conflicting beliefs about the meaning of E, (b) the dispute is resolved by clarifying what E means, and (c) the dispute has the characteristic pointlessness of a dispute in which a key expression is used with different meanings (Chalmers 2011: 520).

Consider an example. Let's say we have a dispute over whether Socrates was bald or not, and we both agree that he had a few hairs on the back of his head. In this case, the pluralist will say that we are using 'bald' with different meanings. Perhaps you are using it to refer to a state where one has no hairs on one's head whatsoever, and I am using it to refer to a state where most of the scalp is not covered by hair, regardless of whether there are a few hairs left. But the monist will deny this. For example, Williamson (1994) holds that vague predicates like

---

<sup>4</sup> See Kripke (1977) for more on the speaker vs. semantic meaning distinction.

'bald' have a hidden meaning, so on this view, it is not that we are using 'bald' with different meanings, it is just that we do not know which of us is right. Chalmers point is that even if we grant monism here, we can still say that our dispute is 'broadly verbal', since (a)-(c) above are satisfied.

One might insist that if there is a case where monism holds for an expression E, the dispute involving E is not verbal, since in such a case the parties do not use E with different meanings. But as Chalmers notes, nothing turns on whether we call such disputes 'verbal disputes'. On Chalmers' terminology, disputes where the parties mean different things with key expressions are 'narrowly verbal' and disputes where the parties mean the same but have different beliefs about what they mean are 'broadly verbal'. If one still insists that meaning different things with a key expression is essential for a dispute being verbal, then one can call broadly verbal disputes 'schmerbal disputes' instead (Chalmers 2011: 519-522).

Something similar can be said for conceptual confluations. In a case where monism holds, there is no conflation of different meanings of the relevant expression, since there is only one meaning. But one can still say that there is a 'broadly' conceptual conflation, similar to how (on Chalmers' terminology) there was a 'broadly' verbal dispute when two parties mean the same but have different beliefs about what they mean. The 'broadly' conceptual conflation consists in one having both the belief that an expression E has meaning<sub>1</sub> in context *c* and the belief that E has meaning<sub>2</sub> in *c*, but without realizing that E does not have both meanings in *c*. If we stipulate that monism holds for 'banana', then my thinking that 'banana' in my mouth refers to both apples and bananas will be an example of this. Of course, one can insist that this is not an instance of a conceptual conflation, since that consists in conflating meanings. But then we can call this a 'schmonceptual conflation' instead.

The point of this translation of talk about meanings and senses to talk about beliefs about meanings or senses is that the idea of conceptual confusions need not be construed in terms of conceptual pluralism, as I did above. One can adopt a broader notion of conceptual confusions in order to accommodate cases of monism, as Chalmers does with verbal disputes. But in this thesis, I shall only be concerned with cases where pluralism arguably holds, so unlike Chalmers, I need not accommodate monism in the characterization of conceptual confusions. For the purposes of this thesis then, it is simpler to think of conceptual confusions as arising from pluralism with respect to a key expression.



#### **4. The Linguistically Sensitive Approach**

I have now introduced the metaphilosophical ideas on which my philosophical approach depends. Call an approach that (a) commits to semantic externalism, conceptual freedom and conceptual pluralism, and (b) tries to avoid conceptual confusions a *linguistically sensitive approach*. This will be my approach in this thesis.

Metaphilosophy is often severed from first-order philosophy, i.e. the philosophy that is about other things than philosophy. But in what follows, I try to take a linguistically sensitive approach to the philosophy of mind. In the main body of the thesis (chapters 2-7) that engages with first-order debates, I do not charge other theorists with being subject to conceptual confusions, however. Rather, my paranoia for conceptual confusions determines how *I* navigate in first-order debates, in the sense that I distinguish different senses of key expressions and try to avoid conceptual confusions when engaging with first-order questions. Nevertheless, it is an interesting question to consider whether there are any examples of conceptual confusions in the first-order debates with which I engage. So I address this question in the conclusion of the thesis (chapter 8).

The chapters that follow address directly or indirectly three topics in the philosophy of mind: phenomenal consciousness, pain, and the extension of psychological predicates. I picked these topics using three criteria, as I wanted to write something that was (a) original, (b) controversial, and (c) defensible. It has been my intention that the individual chapters can be read more or less as independent papers without much knowledge of background and context. I think their contributions are of interest independently of my metaphilosophical views, so you do not need to care about or agree with my metaphilosophical views in order to care about or agree with my (first-order) philosophical views defended in the main body of the thesis. Nevertheless, the ideas of my metaphilosophical views are important to understand how I navigate in philosophical debates, hence the above discussion of these ideas.

The following sections of this introductory chapter provide a roadmap for the rest of the thesis and explains how I take the linguistically sensitive approach when navigating in philosophical debates.

#### **5. Phenomenal Consciousness**

Chapters 2, 3 and 4 are all related to debates about phenomenal consciousness. In chapter 2, I argue that the ‘what-it’s-like’ phrase has a technical meaning when it is used in the way that is characteristic in the literature on phenomenal consciousness, i.e. a meaning peculiar to

philosophy and some parts of cognitive science. Chapter 3 argues that eliminativism is a view worth taking seriously. Chapter 4 addresses the main problem for eliminativism – the problem of explaining why people believe in phenomenal consciousness.

### 5.1. *Chapter 2: What-It's-Like Talk is Technical Talk*

Chapter 2 is about the ‘what-it’s-like’ phrase as it is used in the way that is characteristic in the literature on phenomenal consciousness. Some philosophers claim that the phrase in this context has a technical meaning (Lewis 1995, Carruthers 2000, Byrne 2004, Janzen 2011, Mandik 2016), i.e. a meaning peculiar to a theoretical community. The relevant theoretical community is philosophy and some parts of cognitive science, so on this view, only philosophers and cognitive scientists use the ‘what-it’s-like’ phrase in the way that is characteristic in the literature on phenomenal consciousness. I call this the *technical view*. Others claim that the phrase is used in the same way outside of this theoretical community, hence that the phrase is non-technical (Hellie 2004, Farrell 2016, Stoljar 2016, Chalmers 2020b, Mehta 2021). I call this the *non-technical view*.

Chapter 2 argues in favour of the technical view and against the non-technical view. The first part of the chapter clarifies the technical view by distinguishing different senses of the ‘what-it’s-like’ phrase, which is important in order to avoid verbal disputes and conceptual conflation. It also offers two arguments aiming to show that a defender of the technical view can explain certain things that the defender of the non-technical view cannot explain. In the first argument, I claim that only a defender of the technical view can explain why the technical vs. non-technical disagreement exists. In the second argument, I claim that only a defender of the technical view can explain why some philosophers misunderstand what the ‘what-it’s-like’ phrase means in the literature on phenomenal consciousness. In addition, I address the arguments for the non-technical view and argue that they are unconvincing.

In the second part of the chapter, I move to why it matters that the ‘what-it’s-like’ phrase is technical in the literature on phenomenal consciousness. Firstly, I argue that the phrase says nothing informative about phenomenal consciousness and that definitions of phenomenal consciousness in terms of what it is like to be in mental states are trivial. Secondly, I argue that the fact that lay people use the ‘what-it’s-like’ phrase is not compelling evidence that they believe in phenomenal consciousness. Both these claims have further consequences for debates about phenomenal consciousness, which I draw on in chapters 3 and 4.

### 5.2. *Chapter 3: Eliminativism About Phenomenal Consciousness*

Chapter 3 discusses eliminativism about phenomenal consciousness, with the aim of showing that, for all realists have shown, it is a view worth taking seriously. First, I clarify the position by considering different definitions of phenomenality that can serve in a formulation of eliminativism, which is important in order to avoid verbal dispute. Then I explain the motivation for eliminativism – that it is ontologically simpler than non-reductive realism, and that it avoids the ‘hard problem’ facing reductive realism. Finally, I argue that there is no compelling evidence that belief in phenomenal consciousness is widespread outside of philosophy and some parts of cognitive science. The argument draws on chapter 2, which, to repeat, argues that lay people’s use of the ‘what-it’s-like’ phrase provides no compelling evidence that they believe in phenomenal consciousness. Chapter 3 extends this argument by considering other possible sources of evidence for lay belief in phenomenal consciousness, in particular empirical studies. Some researchers claim that these studies indicate that lay people believe in phenomenal consciousness, but I argue that all the major studies on the topic fail to control for plausible alternative hypotheses, hence that they do not show that lay people believe in phenomenal consciousness. The consequence of this argument is that we are currently not in a position to tell whether common sense counts in favour of eliminativism or realism. Considering all this, I conclude that, for all realists have shown, eliminativism is a view worth taking seriously.

### 5.3. *Chapter 4: Why Do People Believe in Phenomenality?*

Chapter 4 addresses the main problem for eliminativism – the problem of explaining why, given that phenomenality does not exist, people believe that it exists. According to Darwinism (Humphrey 2011), illusionism (Frankish 2016a, 2016b, Graziano 2016, 2019, Kammerer 2018, 2021), and inferentialism (Rey 1995, Clark et al. 2019, Schwarz 2019, Shabasson 2021), the belief in phenomenality is a product of biological or innate mechanisms. According to educationism (Balmer 2020), the belief in phenomenality is a product of cultural mechanisms. Chapter 4 surveys these recently proposed theories and suggests several improvements as well as points out where more work needs to be done. It also distinguishes different senses of ‘illusion’, ‘illusionism’ and ‘illusion problem’, which is important in order to avoid verbal disputes and conceptual connotations in discussions of illusionism and eliminativism. Finally, I argue that while we are currently not in a position to tell which of these theories offers the most promising eliminativism-friendly explanation of belief in phenomenality, the potential of (at least some) of these theories confirms the conclusion of chapter 3, namely that eliminativism is a view worth taking seriously.

## 6. Pain

Chapters 5 and 6 discuss pain in a sense that is independent of phenomenal consciousness, namely pain in the locatable sense. Chapter 5 develops and defends the Bodily Theory of pain, and chapter 6 discusses what a defender of this theory should say about the so-called pain-in-mouth argument.

### 6.1. *Chapter 5: The Bodily Theory of Pain*

Chapter 5 starts by distinguishing pain in the locatable sense – i.e. whatever ‘pain’ refers to in sentences like ‘There is a pain in my foot’ – from pain in the phenomenal sense and pain in the experiential sense, which is important in order to avoid verbal disputes and conceptual confluations. It then introduces the debate between the Bodily Theory and the Experiential Theory. According to the former, pains in the locatable sense are bodily occurrences located in extra-cranial body parts, according to the latter, pains in this sense are experiences or mental states located in the mind or brain.

The rest of the chapter develops and defends the Bodily Theory. First, I argue that defenders of the Bodily Theory should identify pains with proximal activations of nociceptors that cause experiences of pain, as this view accommodates acute pain, chronic/neuropathic pain, phantom limb pain and referred pain. Then I offer two arguments in favour of the Bodily Theory. The first argument is that the Bodily Theory accords better with common sense than the Experiential Theory, since empirical studies indicate that the majority of people believe that pains are located in the body rather than the mind. The second argument is that the Bodily Theory offers a more plausible semantics or interpretation of ordinary pain reports, since on this view, what makes ‘There is a pain in my foot’ true is that there *is* a pain in my foot. On the Experiential Theory, by contrast, ordinary pain reports are either false, or made true by there being a pain in my brain that e.g. represents bodily disturbance in my foot. But the former option is uncharitable and the latter is a less simple, less intuitive and less straightforward semantics than that offered by the Bodily Theory. Finally, I argue that the arguments for the Experiential Theory are unconvincing, and therefore that one should prefer the Bodily Theory.

### 6.2. *Chapter 6: The Bodily Theory and the Pain-in-Mouth Argument*

Chapter 6 is about what a defender of the Bodily Theory should say about the pain-in-mouth argument, which is the step from (1) There is a pain in my finger, and (2) My finger is in my mouth, to (3) There is a pain in my mouth.

Clearly, there is something wrong or misleading with (1)-(3). According to Tye (1995a, 1995b) and Carruthers (2000), the Experiential Theory offers the best account of why the inference to (3) sounds wrong. It is therefore crucial for defenders of the Bodily Theory to find an alternative diagnosis of the pain-in-mouth argument that entails the Bodily Theory, since otherwise the Experiential Theory might have an advantage over the Bodily Theory.

In chapter 6, I offer a novel view of the pain-in-mouth argument that entails the Bodily Theory of pain: the mereological view. According to this view, the sense of ‘in’ that is naturally interpreted as being operative in (1) and (3) is different from the sense of ‘in’ that is naturally interpreted as being operative in (2). While (1) says that there is a pain where a part of my finger is, (2) says that my finger is enclosed by my oral cavity. From these premises it does not follow that there is a pain where a part of my mouth is, which explains why (1)-(3) sounds wrong. I then argue that this mereological view is preferable to two recent accounts that also entail the Bodily Theory – (one version of) the predicative view (Noordhof 2001, 2002, 2005, Hyman 2003, Bain 2007, Liu 2020) and the implicature view (Reuter et al. 2019, Casser and Schiller 2021). I therefore conclude that defenders of the Bodily Theory should adopt the mereological view as an explanation of why the pain-in-mouth argument sounds wrong.

## **7. The Extension of Psychological Predicates and Conclusion**

Chapter 7 is the final chapter that contributes to first-order debates and discusses cognitive scientists’ ascription of psychological predicates to the brain, which is related to questions about the extension of psychological predicates. Chapter 8 concludes the thesis and considers a speculative error theory about some of the views I discuss in the main body of the thesis, according to which they are based on conceptual confluences.

### *7.1. Chapter 7: Does the Brain Think?*

Chapter 7 is about cognitive scientists’ ascription of psychological predicates to the brain, i.e. when they say things like ‘The brain thinks’, which has prompted philosophical debate. According to the Nonsense View, the relevant locutions of cognitive scientists are nonsensical or false (Bennett and Hacker 2003, 2007). According to the Literal View, they are literal truths and report the psychological properties of brains (Dennett 2007, Crane 2015, Figdor 2018).

In chapter 7, I propose a novel view that accommodates aspects from both these views. According to this view, which I call *the Synecdoche View*, we can make sense of the scientific practice by understanding the word ‘brain’ figuratively, namely as a synecdoche (a word for a

part referring to the whole, or vice versa) referring to the human being. Just like I can use ‘my new set of wheels’ to refer to my new car, so cognitive scientists use ‘brain’ to refer to the human being. On this view, ‘The brain thinks’ reports the thinking of the human being, not the thinking of the brain.

The plausibility of the Synecdoche View is not just interesting in its own right but has interesting consequences. Firstly, I argue that there is no reason to take the controversial locutions literally. Secondly, I argue that there is no reason to believe that these locutions indicate empirical support for the claim that brains possess psychological properties. These are substantive and controversial consequences of the Synecdoche View.

### *7.2. Chapter 8: Conclusion*

Chapter 8 concludes the thesis by taking up the idea of conceptual conflation discussed above. In the other chapters, I only make use of this idea for myself when navigating in first-order debates. But in chapter 8, I address the elephant in the room and consider the question of whether there are any examples of conceptual conflations in the debates with which I engage. I offer a speculative error theory, which points out three possible examples and claims that conceptual conflation is a source of these three mistaken philosophical views.

## Chapter 2

### What-it's-like Talk is Technical Talk

**Abstract:** It is common to characterize phenomenal consciousness as *what it is like* to be in a mental state. This chapter argues that the ‘what-it’s-like’ phrase in this context has a technical meaning, i.e. a meaning peculiar to a theoretical community. The relevant theoretical community is philosophy and some parts of cognitive science, so on this view, only philosophers and cognitive scientists use the ‘what-it’s-like’ phrase in the way that is characteristic in the literature on phenomenal consciousness. This claim is not just interesting in its own right but has, I argue, important consequences. Firstly, I argue that the phrase says nothing informative about phenomenal consciousness. Secondly, I argue that non-philosophers’ use of the phrase is not compelling evidence that they believe in phenomenal consciousness. These claims again have further consequences for debates about phenomenal consciousness.

#### 1. Introduction

It is common to characterize phenomenal consciousness as *what it is like* to be in a mental state. Some philosophers claim that the ‘what-it’s-like’ phrase in this context has a technical meaning (Lewis 1995, Carruthers 2000, Byrne 2004, Janzen 2011, Mandik 2016), i.e. a meaning peculiar to a theoretical community. The relevant theoretical community is philosophy and some parts of cognitive science, so on this view, only philosophers and cognitive scientists use the ‘what-it’s-like’ phrase in the way that is characteristic in the literature on phenomenal consciousness. Call this the *technical view*. Others claim that the phrase is used in the same way both inside and outside of this theoretical community, hence that the phrase is non-technical (Hellie 2004, Farrell 2016, Stoljar 2016, Chalmers 2020b, Mehta 2021). Call this the *non-technical view*.

This chapter defends the technical view. The claim that the meaning of the ‘what-it’s-like’ phrase is technical is only about the use of the phrase that is *characteristic* in the literature on phenomenal consciousness, it is not about all uses of the phrase that may occur in that literature or elsewhere (see below).

While I argue that the meaning of the ‘what-it’s-like’ phrase is technical, I do *not* offer an account of what the phrase means. The question of whether the phrase is technical is distinct from the question of what the phrase means, but they are related in the sense that an answer to the first question puts constraints on an answer to the second. If the meaning is technical, then

the meaning cannot be one that is familiar outside of philosophy and cognitive science, and if it is non-technical, then it has to be a meaning that is familiar outside of philosophy and cognitive science. So everything I say is consistent with different accounts of the meaning of the phrase as long as those accounts restrict themselves to meanings that are not familiar outside of philosophy and cognitive science.<sup>5</sup>

The chapter has two parts. In the first part, I defend the technical view. Even though several philosophers claim that the ‘what-it’s-like’ phrase is technical, this is typically not supported with compelling argument. The first part of the chapter tries to remedy that and offers two arguments aiming to show that a defender of the technical view can explain something the defender of the non-technical view cannot explain. In the first argument, I claim that only a defender of the technical view can explain why the technical vs. non-technical disagreement exists. In the second argument, I claim that only a defender of the technical view can explain why some philosophers misunderstand what the ‘what-it’s-like’ phrase means in the literature on phenomenal consciousness. I also address the arguments for the non-technical view and argue that they are unconvincing.

In the second part of the chapter, I move to why it matters that the ‘what-it’s-like’ phrase is technical. Firstly, I argue that the phrase says nothing informative about phenomenal consciousness and that definitions of phenomenal consciousness in terms of what it is like to be in mental states are trivial. Secondly, I argue that the fact that lay people use the ‘what-it’s-like’ phrase is not compelling evidence that they believe in phenomenal consciousness. Both these claims have further consequences for debates about phenomenal consciousness.

Before I lay out the structure of the chapter, I must make one clarification about phenomenal consciousness. Even though it is not always made very clear in the literature, there are two distinct concepts of phenomenal consciousness, one functional and one non-functional (cf. Chalmers 2003: 109, 112, 2018: 50). The functional concept is that of (in Chalmers’ 2003 terminology) type-A (i.e. *a priori*) materialism, in other words, analytic functionalism (e.g. Lewis 1995). The non-functional concept is that of type-B (i.e. *a posteriori*) materialism, such as representationalism (e.g. Tye 1995b) and higher-order theory (e.g. Carruthers 2000), and non-reductive (i.e. anti-materialist) theories, such as dualism (e.g. Chalmers 1996) and

---

<sup>5</sup> Note that the fact that some philosophers both offer an account of the meaning of the ‘what-it’s-like’ phrase and claim that the phrase is non-technical does not necessarily make their account of the meaning inconsistent with my claim that the phrase is technical, since they may be wrong about whether the meaning they associate with the phrase is familiar outside of philosophy and cognitive science. I believe the accounts of e.g. Hellie (2004), Stoljar (2016) and Mehta (2021) are all examples of this, but proving that here would take me too far off topic.



panpsychism (e.g. Strawson 2006). These two concepts of phenomenal consciousness are distinct concepts, since what type-A materialists mean by ‘phenomenal consciousness’ is functional by definition, but what type-B materialists and non-reductivists mean by the term is *not* functional by definition.

In all the chapters that concern phenomenal consciousness, I presuppose the non-functional concept of phenomenal consciousness. A different way to put this point is to say that I only count non-reductive theories and type-B materialism as realist views – i.e. holding that the relevant concept of phenomenal consciousness refers, or that phenomenal consciousness (as I use the term) exists. Thus, I do not count type-A materialism as realist, and what I mean by ‘reductive realism’ in what follows is type-B materialism, not type A. In fact, type-A materialists typically claim that the non-functional concept of phenomenal consciousness does *not* refer, so the type-A materialist is typically an eliminativist about phenomenal consciousness in this non-functional sense (Chalmers 2003: 109, 2018: 50, Lewis 1995: 143).

The chapter is structured as follows. §2 clarifies the technical and non-technical views. §3 discusses the ability sense of ‘knowing what it’s like’. §4 discusses how a defender of the technical view should interpret lay people’s use of the ‘what-it’s-like’ phrase. §5 offers two arguments in favour of the technical view. §6 argues that the arguments for the non-technical view are unconvincing. §7 discusses consequences of the technical view for debates about phenomenal consciousness. §8 concludes.

## **2. The Technical View and the Non-Technical View**

The literature on the ‘what-it’s-like’ phrase contains little discussion about the notion of technical meaning. But a minimal characterization that captures the idea relevant to the debate is the following:

The meaning *M* of an expression is technical if *M* is peculiar to a defined theoretical community, such that *M* is not widely known outside the relevant theoretical community.

I do not claim that a defined theoretical community is necessary for a technical meaning. Perhaps young children’s inventions of new words and peculiar uses of existing words count as introducing technical meanings without involving a defined theoretical community. But an exact definition of technical meaning that captures all cases is not what concerns me here, as

the above characterization is sufficient for the idea relevant to the debate between the technical view and the non-technical view.

Given the above characterization of technical meaning, one formulation of the technical view is the following: the *technical view* is the view that the meaning of the ‘what-it’s-like’ phrase relevant to phenomenal consciousness is peculiar to philosophy and some parts of cognitive science. In other words, the technical view is the view that only philosophers and cognitive scientists use the ‘what-it’s-like’ phrase in the way that is characteristic in the literature on phenomenal consciousness.

By contrast, the *non-technical view* is the view that the ‘what-it’s-like’ phrase as it is characteristically used in the literature on phenomenal consciousness does *not* have a meaning peculiar to philosophy and cognitive science. In other words, the non-technical view is the view that the ‘what-it’s-like’ phrase is used in the way that is characteristic in the literature on phenomenal consciousness both within and outside of philosophy and cognitive science.

Some expressions with technical meanings are only used by members of a defined theoretical community, such as ‘qualia’ and ‘implicature’. Other expressions are used both with a technical meaning within a defined theoretical community and with an ordinary (i.e. non-technical) meaning outside of the theoretical community. For example, ‘phenomenal’ is used both with a technical meaning by philosophers and cognitive scientists in connection with phenomenal consciousness, and with an ordinary meaning outside of philosophy and cognitive science, where it means great or fantastic. The ‘what-it’s-like’ phrase is clearly used outside of philosophy and cognitive science, so if it has a technical meaning in the literature on phenomenal consciousness, then it is like ‘phenomenal’ and not like ‘qualia’ and ‘implicature’.

This raises the question of what the ordinary senses of the ‘what-it’s-like’ phrase are. Identifying the ordinary senses of the ‘what-it’s-like’ phrase will be helpful to understand both the technical view and the non-technical view. To this end, I shall distinguish three ordinary senses – the evaluative-descriptive sense, the non-evaluative-descriptive sense, and the resemblance sense – and give examples of how these senses are operative in talk about mental states. All these senses are discussed in the debate at some point, and under various terminology, but they are never put together systematically in the way I do here. The evidence that these senses are ordinary is the linguistic intuition that the examples I provide are familiar and natural sounding.

The *evaluative-descriptive sense* is operative when one says that what it was like to watch a film was exciting, what it was like to hear a concert was boring, and what it was like to taste decayed food was disgusting. Here the ‘what-it’s-like’ phrase picks out one’s evaluative

judgement of the mental state, which describes the mental state in some evaluative term: exciting, boring, disgusting etc. (cf. Hacker 2002).

The evaluative-descriptive sense of the ‘what-it’s-like’ phrase cannot be the sense of the phrase relevant to phenomenal consciousness – regardless of whether that sense is technical or not. The reason is that many mental states do not prompt evaluative judgements and so are not like anything in the evaluative-descriptive sense, but realists about phenomenal consciousness believe that many (if not all) these mental states are like something in the sense relevant to phenomenal consciousness (cf. Snowdon 2010: 16). For example, my experience of walking down the street may not prompt any evaluative judgement and so is not like anything in the evaluative-descriptive sense. If you ask me what it was like to walk down the street I may answer ‘Nothing in particular. I was just walking down the street as I do every day.’ But realists about phenomenal consciousness plausibly believe that the experience of walking down the street is like something in the sense relevant to phenomenal consciousness.

The *non-evaluative-descriptive sense* is operative when one says that ‘it starts to become clearer’ after the optician asks what it is like to see the letters on the wall, or when the text book says that ‘it consists of firing of neurons in the visual cortex’ after asking what vision is like. In general, an object X being like something in the non-evaluative-descriptive sense just means that a non-evaluative property can be specified in a description of X as an answer to ‘How is X?’, or that one can describe X as being some non-evaluative way. The ‘ways’ picked out by the ‘what-it’s-like’ phrase in this sense are not just any properties, but properties that describe *how* an object is, as opposed to where or when it is (cf. Snowdon 2010, Stoljar 2016, Mehta 2021). For example, if I ask you what Norway is like or how Norway is, then you can say e.g. that Norway has fjords and mountains, or just above five million people – which is your non-evaluative description of what Norway is like, or how Norway is. (If you say it is fantastic or amazing, then that is your evaluative description of what Norway is like, or how Norway is.)

As mentioned above, a mental state can be like something in the sense relevant to phenomenal consciousness without being like anything in the evaluative-descriptive sense. This is why the evaluative-descriptive sense cannot be the sense relevant to phenomenal consciousness – regardless of whether that sense is technical or not. But there is no analogous dissociation for what mental states are like in the sense relevant to phenomenal consciousness and what mental states are like in the non-evaluative-descriptive sense. Perhaps there are mental states that are not like anything in the sense relevant to phenomenal consciousness that are like

something in the non-evaluative-descriptive sense.<sup>6</sup> But all mental states that are like something in the sense relevant to phenomenal consciousness are like something in the non-evaluative-descriptive sense, since all phenomenal states can be described as being some non-evaluative ‘way’. For this reason, the non-evaluative-descriptive sense of the ‘what-it’s-like’ phrase *can* be the sense relevant to phenomenal consciousness, *if* the meaning of the phrase is non-technical and the phrase is meant to pick out the phenomenal ‘way’ mental states are. Of course, if the meaning of the phrase is technical, then it cannot be the non-evaluative-descriptive sense.

The *resemblance sense* is operative when one says that the taste of lemon is like the taste of lime. Here the ‘what-it’s-like’ phrase picks out a similarity or resemblance between two phenomena (cf. Gaskin 2019). Just as one might say that one person is like another person – meaning the former resembles the latter, so one might say that one experience is like another experience – meaning the first experience resembles the second. All states that are like something in the sense relevant to phenomenal consciousness are like something in the resemblance sense, since everything resembles something in some respect. Therefore, the resemblance sense – just like the non-evaluative-descriptive sense – *can* be the sense of the ‘what-it’s-like’ phrase relevant to phenomenal consciousness. But it can only be the sense relevant to phenomenal consciousness *if* that sense is non-technical and the phrase is meant to pick out phenomenal resemblance, i.e. resemblance restricted to phenomenal properties. If the sense of the phrase relevant to phenomenal consciousness is technical, then it cannot be the resemblance sense.

With these distinctions in mind, we are now in a position to provide a second formulation of both the non-technical view and the technical view. As the evaluative-descriptive sense cannot be the sense of the ‘what-it’s-like’ phrase relevant to phenomenal consciousness, the defender of the non-technical view is left with three options. The first is that the phrase is used in the non-evaluative-descriptive sense – picking out the phenomenal ‘way’ mental states are. The second option is that the phrase is used in the resemblance sense – picking out phenomenal resemblance. The third option is that the phrase is used in a fourth ordinary sense: the phenomenal sense – picking out phenomenal consciousness. I am not sure if all defenders of the non-technical view have the same understanding of which of these senses is the sense of the ‘what-it’s-like’ phrase relevant to phenomenal consciousness. But that does not matter for present purposes, since it suffices to think of the non-technical view as a disjunction

---

<sup>6</sup> I am agnostic whether there are any such mental states, and if they exist, which they are, since it is controversial which states are phenomenal.

with three disjuncts: *either* the ‘what-it’s-like’ phrase is used in the non-evaluative-descriptive sense, *or* in the resemblance sense, *or* in a distinct but ordinary phenomenal sense.

The defender of the technical view denies all disjuncts: the ‘what-it’s-like’ phrase is neither used in the non-evaluative-descriptive sense, nor in the resemblance sense, nor in a distinct but ordinary phenomenal sense. According to the technical view, the phrase *is* used in a distinct phenomenal sense. But this sense is not ordinary. Rather, it is a sense that is peculiar to philosophy and cognitive science, i.e. a technical sense. The claim is only that the sense with which philosophers and cognitive scientists characteristically use the phrase in connection with phenomenal consciousness is a technical sense. This is consistent with lay people sometimes using the phrase in the non-evaluative-descriptive sense or the resemblance sense in connection with phenomenal consciousness. But even if lay people sometimes use the phrase that way – which I am agnostic about here – this is not the way philosophers and cognitive scientists characteristically use the phrase in the literature on phenomenal consciousness. Or so I argue.

### 3. The Ability Sense

There is one more sense that needs to be introduced: the *ability sense*. The ability sense is not a sense of the ‘what-it’s-like’ phrase but a sense of the complex phrase consisting of a cognitive verb like ‘know’ or ‘learn’ followed by the ‘what-it’s-like’ phrase. It is a distinct sense that needs to be introduced because there are arguably examples of lay people using e.g. ‘knowing what it’s like’ that cannot easily be analysed in terms of a sense of ‘know’ plus either the evaluative-descriptive sense, the non-evaluative-descriptive sense or the resemblance sense. But as I argue in the following section, these examples can be analysed in terms of the ability sense.

The ability sense can be operative in sentences such as ‘You do not know what it is like to experience this unless you have the experience’. To see that not all examples of sentences like this can be analysed by appealing to a sense of ‘know’ plus the evaluative-descriptive sense, the non-evaluative-descriptive sense, or the resemblance sense, we can stipulate a case where you both know what the experience resembles, what your evaluative judgement will be, and some non-evaluative property the experience has. Let’s say you want to taste a rare fruit, which you know resembles another fruit you like, and since you like the latter fruit you will find it delicious to taste the former. Let’s also say that you know the experience of tasting the rare fruit will have the non-evaluative property of consisting in such-and-such brain activity. Then you know what tasting the rare fruit is like in the resemblance sense, the evaluative-descriptive

sense, and the non-evaluative-descriptive sense. One might still say that there is a different sense in which you do *not* know what the experience is like, which ‘only experience can teach you’. One thing the experience will teach you is the ability to remember, imagine and recognize the experience. Thus, one might say that one sense of ‘knowing what it’s like’ refers to the possession of these abilities (Lewis 1988).

My claim is only that the ability sense is an ordinary sense that sometimes, not always, is the sense of e.g. ‘knowing what it’s like’. 3.1. defends the claim that the ability sense is *ordinary*, and 3.2. clarifies the claim that the ability sense only *sometimes* is the sense of ‘knowing what it’s like’ by discussing examples that cannot be analysed in terms of the ability sense.

### 3.1. *Why the Ability Sense is an Ordinary Sense*

Some might doubt that the ability sense is really an *ordinary* sense of ‘knowing what it’s like’ (cf. Lycan 1996: 93). They might think that the ordinary senses of ‘knowing what it’s like’ only concern knowledge of evaluative properties, non-evaluative properties, resemblance, and perhaps phenomenal properties, but not the possession of abilities. On this view, the ability sense is a philosophers’ invention, stipulated by defenders of the ‘ability hypothesis’, according to which knowing what it is like to see red consists in possession of the abilities to imagine, recognize and remember seeing red (Lewis 1988).<sup>7</sup> Why believe the ability sense is an ordinary sense?

One reason is the data uncovered by Gregory et al. (submitted). Gregory and colleagues conducted a study where they asked lay people unfamiliar with philosophy whether the colour-blind super-scientist Mary learns what it is like to see red after having a colour sight operation and then seeing a red tomato – closely modelled on Jackson’s (1982) original story with Mary’s release from the black and white room. To this question, most subjects tended to give positive responses. That is, subjects did not think Mary knew what it is like to see red pre-operation, but they thought she learned it post-operation. Gregory and colleagues also asked subjects to

---

<sup>7</sup> Lycan (1996) holds that the ability sense is not ordinary and claims that ‘S knows what it’s like to see blue’ never concerns the possession of abilities. His argument is the following: ‘S knows what it’s like to see blue’ means roughly ‘S knows that it is like Q to see blue’, where ‘Q’ suitably names some inner phenomenal property or condition’ (1996: 93). I agree that ‘S knows what it’s like to see blue’ *sometimes* means ‘S knows that it is like Q to see blue’ and that in the contexts it does, it does not mean ‘S possesses the ability to imagine, recognize and remember seeing blue’. But this is consistent with ‘S knows what it’s like to see blue’ sometimes meaning ‘S possesses the ability to imagine, recognize and remember seeing blue’. Lycan gives no argument that ‘S knows what it’s like to see blue’ *always* means what he thinks it does. So his claim about what it means is no more an argument against the claim that the ability sense is ordinary than the claim that ‘bank’ refers to the financial institution is an argument against the claim that ‘bank’ is an ordinary word for the place next to the river.

comments upon their response to the Mary scenario. Many of these comments are difficult to interpret. But some comments can plausibly be interpreted in a way suggesting that subjects understood the ‘what-it’s-like’ phrase in either of the senses discussed above, including the ability sense.

Here are some comments subjects wrote to explain why they did not think Mary knows what it is like to see red pre-operation:

‘Although she knows the theory behind it, it’s not a concept she can truly visualize.’

‘It is impossible to imagine a color you have never seen.’

‘No. Mary is having someone purposely give her a red tomato when she wakes up so that her brain will lock that color in and it will be memorized forever.’

‘She never actually saw it so I do not think she can imagine it. I think it has to be experienced.’

‘She only knows the description of red and red is still unknown to her. You can’t imagine something you’ve never seen before.’

And here are some comments subjects wrote to explain why they thought that Mary learns what it is like to see red post-operation:

‘She will learn in her mind what red looks like so it can register when she sees something else red.’

‘The red tomato would give her proof about the true color of red which she can store into her long term memory to know that the tomato is a form of red.’

‘This will be her first time ever seeing the color red so after seeing the tomato and knowing that it’s red, she will be able to identify the color red.’<sup>8</sup>

The first set of comments explains Mary’s ignorance of what it is like to see red pre-operation by appealing to her lack of the ability to imagine and remember seeing red. The second set of

---

<sup>8</sup> I am grateful to Daniel Gregory and colleagues for sharing the data from which these comments were taken.

comments explains Mary's learning what it is like to see red post-operation by appealing to her gaining the ability to recognize and remember red.

My argument for the claim that the ability sense is an ordinary sense is that subjects' understanding 'knowing/learning what-it's-like' in the ability sense best accounts for why they wrote the above comments. I shall first argue that neither of the senses of the 'what-it's-like' phrase discussed in the previous section fit with the above comments before arguing that the ability sense makes a good fit.

Consider first the sense of the 'what-it's-like' phrase relevant to phenomenal consciousness. If the subjects who wrote these comments understood 'knowing/learning what-it's-like' as involving this sense, then one would not expect the above comments. Rather, one would expect comments explaining Mary's ignorance of and learning what it is like to see red by appeal to something which is *not* merely a lack, or possession, of the ability to imagine, remember and recognize seeing red. This is because knowledge of phenomenal properties does not consist in the possession of these abilities. Indeed, the 'ability hypothesis', according to which Mary's learning what it is like to see red consists in her gaining these abilities was not advanced as an account of what phenomenal knowledge is, but as an *alternative* to the hypothesis that she gains phenomenal knowledge (Lewis 1988). Thus, the sense of the 'what-it's-like' phrase relevant to phenomenal consciousness is not plausibly the sense in which the subjects who wrote the above comments understood the 'what-it's-like' phrase.

Consider next the evaluative-descriptive sense. If the subjects who wrote the above comments understood the 'what-it's-like' phrase in this sense, then the appeal to the abilities to imagine, recognize and remember seeing red would seem irrelevant to Mary's ignorance of and learning what it is like to see red. This is because one being ignorant of or learning what seeing red is like in the evaluative-descriptive sense concerns one being ignorant of or learning that seeing red is e.g. appealing, amazing, terrible etc., it does not concern one's possession of the abilities to imagine, recognize and remember seeing red. Therefore, the subjects who wrote the above comments did not plausibly understand the 'what-it's-like' phrase in the evaluative-descriptive sense.

Then consider the non-evaluative-descriptive sense. Ignorance of and learning what it is like to see red in the non-evaluative-descriptive sense concerns ignorance of and learning the non-evaluative 'ways' seeing red is, i.e. any non-evaluative property that can be given as an answer to 'How is seeing red?'. If possessing the ability to imagine, recognize and remember seeing red is *not* a non-evaluative 'way' seeing red is, then it is clear that the subjects who wrote the above comments did not understand the 'what-it's-like' phrase in the non-evaluative-



descriptive sense, given their appeal to these abilities. But even if one grants that possessing the ability to imagine, recognize and remember seeing red *is* a non-evaluative ‘way’ seeing red is, then it is still not plausible that subjects understood the ‘what-it’s-like’ phrase in the non-evaluative-descriptive sense. Let me elaborate.

If the ‘what-it’s-like’ phrase in the description of the Mary scenario is understood in terms of the non-evaluative-descriptive sense and the relevant non-evaluative ‘way’ it is to see red is possessing the abilities to imagine, recognize and remember seeing red, then there are two ways of understanding the claim that Mary does not know what it is like to see red pre-operation. Perhaps the most natural interpretation is that she does not know *that she has* these abilities. But on this interpretation, it would not be relevant to comment that she lacks the abilities – which subjects do in the above comments – since the claim that she lacks the abilities is different from the claim that she does not know that she has them. So on this interpretation, it is not plausible that subjects understood the ‘what-it’s-like’ phrase in the non-evaluative-descriptive sense.

An alternative, perhaps less natural, interpretation of the claim that she does not know what it is like to see red – when what it is like to see red is understood as possessing the relevant abilities, understood as a non-evaluative ‘way’ it is to see red – is that she does not know *how to* imagine, recognize and remember seeing red. This fits with the above comments. But possessing the ability to imagine, recognize and remember seeing red is only one of the many ‘ways’ it is to see red. Different ‘ways’ it is to see red, is consisting of such-and-such neural activity, and having such-and-such effects, which Mary *does* know, even pre-operation. But if Mary knows most of the ‘ways’ it is to see red is and is ignorant of only one of them pre-operation, then it is strange that subjects did not mention this in their comments.

Consider an analogy. If one were asked whether a person knows how Norway is, or what Norway is like (in the non-evaluative-descriptive sense), and the person is Norwegian, has lived in Norway his entire life, is a professor of Norwegian history, geography and culture, then it would be strange to answer that he does not know how Norway is, or what Norway is like, on the ground that he is ignorant of one arcane fact about Norway, such as the exact number of moose there is in in the country. At least it would be strange to answer this without making clear that he knows how Norway is, or what Norway is like, in most other respects when asked to elaborate on one’s answer. Given that Mary does know most of the things seeing red is like in the non-evaluative descriptive sense (i.e. how seeing red is), it would be strange if the subjects who wrote the above comments understood the ‘what-it’s-like’ phrase in the non-evaluative descriptive sense and denied that Mary knows what seeing red is like pre-operation on the

ground that she lacks the abilities to imagine and remember seeing red. So regardless of whether possession of the abilities to imagine, recognize and remember seeing red is a non-evaluative ‘way’ it is to see red, the non-evaluative descriptive sense is not plausibly the sense in terms of which the subjects who wrote the above comments understood the ‘what-it’s-like’ phrase.

Finally, consider the resemblance sense. Anything resembles everything in some respect, so clearly seeing red resembles imagining, recognizing and remembering red. If the subjects who wrote the above comments understood the ‘what-it’s-like’ phrase in the resemblance sense, then their idea is that Mary does not know that seeing red resembles imagining and remembering seeing red pre-operation, but she knows that seeing red resembles recognizing and remembering red post-operation. But this seems like a strange idea. Given that Mary knows all physical information, it is plausible that she knows that seeing red resembles imagining and remembering seeing red pre-operation. And it is not clear why the operation should make her learn that seeing red resembles recognizing and remembering red post-operation. Thus, it is not plausible that the subjects who wrote the above comments understood the ‘what-it’s-like’ phrase in the resemblance sense.

I have argued that neither the sense of the ‘what-it’s-like’ phrase relevant to phenomenal consciousness, nor the evaluative-descriptive sense, nor the non-evaluative-descriptive sense, nor the resemblance sense fit with the above comments. The only option left is the ability sense, which *does* explain the content of the above comments. For the above comments explain Mary’s ignorance of what it is like to see red pre-operation by appealing to her lack of the abilities to imagine and remember seeing red, and they explain her learning what it is like to see red post-operation by appealing to her acquire the abilities to recognize and remember red. And not knowing what seeing red is like in the ability sense *is* just to lack these abilities, and learning what it is like *is* just to acquire them. Therefore, the above comments suggest that the subjects who wrote them understood the ‘what-it’s-like’ phrase in the ability sense, and since these subjects are lay people unfamiliar with philosophy, the ability sense is plausibly an ordinary sense.

### 3.2. *The Ability Sense and Knowing What It’s Like*

My claim that the ability sense is an ordinary sense only commits me to the claim that lay people *sometimes* use e.g. ‘knowing what it’s like’ in this sense, it does not commit me to the claim that ‘knowing what it’s like’ should *always* be understood in the ability sense.

For example, I do not claim that the ability sense is operative in Jackson’s (1982) ‘knowledge argument’, even though Jackson uses ‘knowing what it’s like’.

We can represent Jackson's argument as follows:

(P1) Mary knows all the physical facts about seeing red before her release from the black-white room.

(P2) Mary does not know what it is like to see red before her release.

(P3) Knowing what it is like to see red is knowing a fact about seeing red.

(C) Therefore, not all facts about seeing red are physical facts.

Given that I defend the technical view, I am committed to the 'knowledge argument' involving the technical sense of the 'what-it's-like' phrase relevant to phenomenal consciousness. Thus, it does not involve the ability sense of 'knowing what it's like', which I claim is an ordinary sense.

At this point, one may wonder what the relation is between my hypothesis that the ability sense is ordinary, and the 'ability hypothesis', which, as mentioned above, is the claim that knowing what it is like to see red consists in possessing the abilities to imagine, recognize and remember seeing red (Lewis 1988). The relation between these hypotheses depends on how one understands the 'ability hypothesis'.

The 'ability hypothesis' is sometimes taken to be an objection to the 'knowledge argument' (Conee 1994, Lycan 1996). But the mere claim that Mary learns what it is like to see red in the ability sense (i.e. that she gains some abilities) is consistent with Mary learning what it is like to see red in the sense relevant to phenomenal consciousness. So if the 'ability hypothesis' is supposed to be an objection to the 'knowledge argument', then the claim is not merely that Mary gains the abilities to imagine, recognize and remember seeing red, or that *a* sense in which Mary learns what it is like to see red concerns the acquisition of these abilities. Rather, it is the claim that, once one knows all physical facts, *the only* sense in which one can learn what it is like to see red is by acquiring the abilities to imagine, recognize and remember seeing red, i.e. that *the only* sense in which Mary learns what it is like to see red is in the ability sense.<sup>9</sup> My claim that the ability sense is ordinary entails that *a* sense in which Mary learns what it is like to see red concerns the possession of the abilities to imagine, recognize and

---

<sup>9</sup> I find it difficult to interpret Lewis (1988) on this point, i.e. whether he intended the 'ability hypothesis' to be an objection to the 'knowledge argument' or not. If he did not intend the 'ability hypothesis' to be an objection to the 'knowledge argument', then his objection to that argument stems from his eliminativism about phenomenal consciousness (cf. 1995: 143) rather than claims about abilities. If this is the case, then his discussion about abilities is perhaps only meant to point out the sense in which he, as an eliminativist about phenomenal consciousness, nevertheless thought that there is a sense in which it is true to say that (a) experience is the best teacher and (b) Mary learns what it is like to see red.

remember seeing red. But this does not commit me to the claim that *the only* sense in which Mary learns what it is like to see red is that she gains these abilities. The claim that the ability sense is ordinary is consistent with Mary learning what it is like to see red in senses of ‘learning what-it’s-like’ other than the ability sense.

For this reason, everything I have said so far is consistent with many of the claims of authors who object to the ‘ability hypothesis’ – understood as an objection to the ‘knowledge argument’, i.e. the claim that *the only* sense in which Mary learns what it is like to see red is by acquiring the abilities to imagine, recognize and remember seeing red. For example, Conee (1994), claims that a person looking at a green object can know what it is like to see green, despite not having the ability to imagine green (Conee 1994: 139). This is inconsistent with the claim that *the only* sense in which one can know what it is like to see green is by possessing the abilities to imagine, recognize and remember seeing green (which I doubt that anyone claims). But it is consistent with my claim that the ability sense is ordinary. Depending on how one interprets the ‘what-it’s-like’ phrase, the person in Conee’s example can know what seeing green is like in the technical sense relevant to phenomenal consciousness, the evaluative-descriptive sense, or the non-evaluative-descriptive sense. The claim that a person can know what seeing green is like in these senses is perfectly consistent with her not knowing what it is like in the ability sense and with my claim that the ability sense is ordinary.

We can add many more examples where ‘knowing/learning what it is like’ should not be understood in the ability sense. But that is consistent with my claim that the ability sense is ordinary, since this claim only commits me to the claim that the ability sense is *sometimes* the sense of ‘knowing/learning what it’s like’, it does not commit me to the claim that it is *always* the sense of ‘knowing/learning what it’s like’.

#### **4. How Should We Interpret Lay People’s Use of the ‘What-It’s-Like’ Phrase?**

This section illustrates how the technical view works in practice and applies it to examples of lay people using the ‘what-it’s-like’ phrase by interpreting these uses of the phrase as different from the use that is characteristic in the literature on phenomenal consciousness. All the examples are taken from Hellie (2004), Farrell (2016) and Stoljar (2016), who take these examples to support the non-technical view, since they believe these examples involve the use of the phrase that is characteristic in the literature on phenomenal consciousness (Hellie 2004: 339, Farrell 2016: 59-60, Stoljar 2016: 1183). The fact that the interpretations I offer are available shows that these examples provide no support for the non-technical view, however.

My strategy is to offer interpretations that involve either the evaluative-descriptive sense, the non-evaluative-descriptive sense restricted to the non-phenomenal ‘ways’ things are, the resemblance sense restricted to non-phenomenal resemblance, or the ability sense, since neither of these senses is the sense relevant to phenomenal consciousness. Let me elaborate.

As noted in §2 and §3, regardless of whether the ‘what-it’s-like’ phrase is technical or not in the literature on phenomenal consciousness, the sense of the phrase relevant to phenomenal consciousness can neither be the evaluative-descriptive sense nor the ability sense. The reason it cannot be the evaluative-descriptive sense is that a mental state can be like something in the sense relevant to phenomenal consciousness without being like anything in the evaluative-descriptive sense. The reason it cannot be the ability sense is that the ability sense is not a sense of ‘what it’s like’ but a sense of ‘knowing what it’s like’. So if one can plausibly interpret any of the examples of lay people using the ‘what-it’s-like’ phrase as involving either the evaluative-descriptive sense or the ability sense, then the defender of the technical view can accommodate these examples no less than the defender of the non-technical view.

In contrast to the evaluative-descriptive sense and the ability sense, the non-evaluative-descriptive sense and the resemblance sense *can* both be the sense of the phrase relevant to phenomenal consciousness *if* the phrase is non-technical.

The non-evaluative-descriptive sense picks out the non-evaluative ‘ways’ things are, but there are many non-evaluative ‘ways’ mental states that are not concerned with phenomenal consciousness. To repeat, the ‘ways’ mental states are, are any properties that can be specified in an answer to ‘How is the mental state?’, and many such ‘ways’ are not concerned with phenomenal consciousness (see below). So if any of the examples of lay people’s use of the ‘what-it’s-like’ phrase can be interpreted as being about a non-phenomenal ‘way’ mental states are, then the defender of the technical view can accommodate these examples no less than the defender of the non-technical view.

Similarly for the resemblance sense, which picks out the resemblance between two phenomena. Mental states resemble each other in many respects other than in terms of phenomenal properties, e.g. in terms of the objects they represent and the effects they cause. What causes people to say that one mental state resembles or ‘is like’ another mental state is an empirical question – whether it is the object represented, phenomenal properties, or the effects of the mental state. So if any of the examples of lay people’s use of the ‘what-it’s-like’ phrase can be interpreted as being about non-phenomenal resemblance, then the defender of the technical view can accommodate these examples no less than the defender of the non-technical view.

Having laid out the strategy, I shall now offer alternative interpretations of all the examples offered by Hellie, Stoljar and Farrell.

The first example is from The Beatles' song 'She Said, She Said':

She said 'I know what it's like to be dead' (Lennon and McCartney 1966, quoted by Hellie 2004: 369).

This can be interpreted as involving the ability sense, since the 'she' of the song is arguably expressing her rich life experience with the exaggeration that she knows what it is like to be dead, and her life experience is something the protagonist might not be able to imagine, given his lack of life experience. It can also be interpreted in the non-evaluative-descriptive sense restricted to the non-phenomenal 'ways' it is to be 'dead'. For example, if being 'dead' consists of having bad fortune, then what it is like to be 'dead' might be making one realize the contingencies of life, which is a non-phenomenal 'way' being 'dead' is. On this interpretation, she is saying that she knows how bad fortune makes one realize the contingencies of life.

The second example is from Everlast's song 'What It's Like':

God forbid you ever had to walk a mile in his shoes, because then you really might know what it's like to sing the blues (Schrody 1998, quoted by Hellie 2004: 369).

This can be interpreted as involving the non-evaluative-descriptive sense restricted to the non-phenomenal 'ways' singing blues are. The 'he' in this context is a poor alcoholic, so 'singing blues' is plausibly a metaphor for having a poor life. If so, then what it is like to sing the blues might be making one realize that the world is unjust, which is a non-phenomenal 'way' 'singing blues' is. On this interpretation, the above says that you would not like to walk a mile in the alcoholic's shoes, since then you would know how a poor life makes one realize that the world is unjust, perhaps with the implicature<sup>10</sup> that you would give up the romantic idea of a just world.

The third example is from Bee Gees' song 'To Love Somebody':

You don't know what it is like to love somebody (Gibb and Gibb 1967, quoted by Stoljar 2016:1183).

---

<sup>10</sup> An implicature is a content that a speaker conveys with an utterance that is not part of the literal meaning of the utterance.

This can be interpreted as involving the ability sense, because one does not have the ability to imagine the experience of love unless one has had the experience of love, and the protagonist of the song plausibly addresses this to an antagonist to which the protagonist – due to frustration – might exaggerate about by ascribing her a lack of experience of love. It can also be interpreted in the non-evaluative-descriptive sense restricted to the non-phenomenal ‘ways’ loving somebody are. For example, a non-phenomenal ‘way’ to love somebody might be making one do irrational things, in which case what it is like to love somebody is making one do irrational things. On this interpretation, the above says that you do not know that love makes you do irrational things, perhaps with the implicature that you have not experienced<sup>11</sup> love.

The fourth example is this:

Neither does he know what it is like to be scorched by lightning, but he has experienced the shrivelling effects of unrequited longing’ (unknown 1891: 541, quoted by Farrell 2016: 59)

This can be interpreted as involving the evaluative-descriptive sense, as what it is like to be scorched by lightning is contrasted with the effects of unrequited longing, one of which might be one’s evaluative judgement, e.g. one finding it dreadful. It can also be interpreted in the non-evaluative-descriptive sense restricted to the non-phenomenal ‘ways’ being scorched by lightning are. For example, a ‘way’ to be scorched by lightning is making one’s heart stop, in which case what it is like to be scorched by lightning is making one’s heart stop. On this interpretation, the above says that he does not know that being scorched by lightning makes one’s heart stop.

The fifth example is this:

Nina: And I should like to change places with you.

Trigorin: Why?

Nina: To find out how a famous genius feels. What is it like to be famous? What sensations does it give you? (Chekhov 1912, quoted by Farrell 2016: 60)

---

<sup>11</sup> Here and below, I use the verb ‘experience’ in a non-phenomenal way, meaning one has an episode occurring in one’s life.

This can be interpreted as involving the non-evaluative-descriptive sense restricted to the non-phenomenal ‘ways’ being famous are. For example, a non-phenomenal ‘way’ being famous is, is feeling (in a functional sense) superior, in which case what it is like to be famous is feeling superior. On this interpretation, Nina wants to experience how it is to be famous, e.g. whether it makes one feel superior or something else.

The sixth example is this (describing a blind person):

His great regret is that his normal sight at birth was too early to allow him to remember what it is like to see (unknown 1938: 7, quoted by Farrell 2016: 60).

This can be interpreted as involving the ability sense, since remembering what it is like to see in this context can be interpreted as remembering having the ability to imagine seeing, which (on this interpretation) the blind person regrets not having. It can also be interpreted in the non-evaluative-descriptive sense restricted to the non-phenomenal ‘ways’ seeing are. For example, one non-phenomenal ‘way’ seeing is, is enabling one to distinguish people by their faces, in which case what it is like to see is enabling one to distinguish people by their faces, which the blind person regrets not remembering.

The seventh example is this:

Drug-inspired psychedelic art tried to portray what it’s like to ‘see sounds’ and ‘taste colors’ while on an LSD trip (Cain 1969: F17, quoted by Farrell 2016: 60)

This can be interpreted as involving the resemblance sense restricted to non-phenomenal resemblance, since artists tried to portray what it is like to see sounds and taste colours by creating an artwork that presumably resembles the seeing of sounds and taste of colours in some way, which may be non-phenomenal. For example, the artwork may resemble the seeing of sounds and taste of colours in the sense that it prompts a reaction similar to that prompted by seeing sounds and tasting colours.

The eighth example is this:

His description of what it is like to ‘see’ as a blind man is fascinating and inspiring (Kirsch 1987, quoted by Farrell 2016: 60).



This can be interpreted as involving the evaluative-descriptive sense, as the evaluative judgements of a blind person to objects in the perceptual environment may be different from those who can see, and thereby be fascinating and inspiring. It can also be interpreted in the non-evaluative-descriptive sense, restricted to the non-phenomenal ‘ways’ seeing are, since the ‘way’ seeing is for a blind person (e.g. involving the coordination of a dog and a white cane) may be different from the ‘way’ it is for those who can see, and thereby be fascinating and inspiring.

The ninth and final example is this:

When I think of the future, I feel fear. I feel fear and I am a 34 year old man. What is it like for the children who live here in Gaza? What is it like for their parents? (Damo 2012, quoted by Farrell 2016: 60).

This can be interpreted as involving the evaluative-descriptive sense, since one could answer the 34 year old man by saying e.g.: ‘It is terrible for the children and parents too; they know that the future in Gaza is uncertain’ – thereby specifying the evaluative judgement of the children and parents. It can also be interpreted in the non-evaluative-descriptive sense restricted to the non-phenomenal ‘ways’ thinking of the future are. For example, one non-phenomenal ‘way’ thinking of the future might be, is making one hopeful but worried – in functional senses of these terms – in which case what it is like to think of the future is making one hopeful but worried. On this interpretation, the 34 year old man is wondering how it is for the children and parents to think of the future, e.g. whether it makes them hopeful but worried, or something else.

The above shows that all the examples offered by Hellie, Stoljar and Farrell can plausibly be interpreted as involving either the evaluative-descriptive sense, the non-evaluative-descriptive sense restricted to the non-phenomenal ‘ways’ things are, the resemblance sense restricted to non-phenomenal resemblance, or the ability sense. This means that examples of lay people’s use of the ‘what-it’s-like’ phrase – even those examples proponents of the non-technical view appeal to in support of their view – can be accommodated by the technical view.

Of course, the fact that the above examples *can* be interpreted as involving senses of the ‘what-it’s-like’ phrase that are not concerned with phenomenal consciousness does not show that these examples do not involve the use of the phrase that is characteristic in the literature on phenomenal consciousness, since examples of lay people’s use of the ‘what-it’s-like’ phrase *can* also be interpreted in the sense relevant to phenomenal consciousness. More generally,

when there are two competing interpretations available for certain locutions, then the fact that one interpretation can accommodate the relevant locutions does not show that that interpretation is more plausible than the other. Thus, the fact that examples of lay people's use of the 'what-it's-like' phrase *can* be interpreted in senses that are not concerned with phenomenal consciousness is no reason to prefer the non-technical view. The reason to prefer the non-technical view is the arguments I offer in the next section, and *those arguments* support the claim that the above examples do not involve the use of the 'what-it's-like' phrase that is characteristic in the literature on phenomenal consciousness but rather the alternative senses proposed in this section.

## 5. The Arguments for the Technical View

I have now clarified the technical view and discussed how defenders of this view should interpret lay people's use of the 'what-it's-like' phrase. But why should one believe that the technical view is true? This section offers two abductive arguments.

### 5.1. *First Argument*

The first argument for the technical view is that a defender of this view can explain why some philosophers believe that the 'what-it's-like' phrase is technical while others believe that it is non-technical, but it is not clear that a defender of the non-technical view can explain both these things.

Here are some statements demonstrating that some philosophers believe that the 'what-it's-like' phrase has a technical meaning in the literature on phenomenal consciousness:

'What it's like' or 'how it seems' are ordinary enough – but when used as terms for qualia, they are used in a special technical sense (Lewis 1995: 140).

[T]he terminology of 'subjective feel' and 'what-it-is-like' are quasi-technical in nature, having been introduced by philosophers (Carruthers 2000: 14n11).

[I]t is doubtful that 'There is something it's like for so-and-so to  $\phi$ ' has some 'special use to describe subjectivity' (dialects of analytic philosophy aside) (Byrne 2004: 215).

[I]n philosophical discussions of consciousness – talk of what it is like for S to V is to be understood in a relatively technical or specialized sense (Janzen 2011: 279).

‘[W]hat-it’s-like’ [...] is yet another technical term shedding no light on the term ‘phenomenal’ (Mandik 2016: 142).

Most of these philosophers do not provide explicit arguments for these claims. But even though they use slightly different formulations, it seems clear that they take the ‘what-it’s-like’ phrase in the relevant philosophical context to have a meaning peculiar to philosophy, and thereby is technical. By contrast, other philosophers (Hellie 2004: 336-339, Farrell 2016, Stoljar 2016: 1183, Chalmers 2020b: 237-238, Mehta 2021: 6) claim that the ‘what-it’s-like’ phrase has a non-technical meaning in the literature on phenomenal consciousness.

What explains this disagreement? A defender of the technical view can explain *both* why some philosophers believe that the ‘what-it’s-like’ phrase is technical *and* why others believe that it is non-technical. The reason some believe that it is technical is that it *is* technical. And the reason others believe that it is non-technical might be *either* because they conflate the philosophical use of the phrase with other (ordinary) uses – which is not so strange, given that there are (as noted in §2) several ordinary uses of the phrase and that philosophers typically do not clarify what they mean by the phrase – *or* because they are too entrenched in philosophy to reliably distinguish technical from non-technical language.

But it is not clear how a defender of the non-technical view can explain both why some believe that the phrase is technical and why others believe that it is non-technical. Of course, defenders of this view can say that the reason some believe that it is non-technical is that it *is* non-technical, but it is not clear how they can explain that some believe that it is technical. In fact, if the ‘what-it’s-like’ phrase really were non-technical, then one would not expect so many philosophers to claim that it was technical, since one would rather expect everyone to agree that it is non-technical. There is no analogous technical vs. non-technical disagreement about the use of other ordinary expressions in philosophy, so if the ‘what-it’s-like’ phrase was used in a perfectly ordinary way in the literature on phenomenal consciousness, then why would so many philosophers claim that it is technical?

The lack of an answer to this question points to an asymmetry: while defenders of the technical view can explain both why some believe that the ‘what-it’s-like’ phrase is technical and why others believe that it is non-technical, defenders of the non-technical view only seem able to explain why some believe that it is non-technical. Of course, defenders of the non-technical view will claim that defenders of the technical view have misunderstood how the phrase is used in the literature on phenomenal consciousness. But the point is that, given the view that the ‘what-it’s-like’ phrase has a non-technical meaning in this literature, there is no

explanation for why defenders of the technical view have (according to defenders of the non-technical view) misunderstood how the phrase is really used. By contrast, there *is* an explanation for why, according to defenders of the technical view, defenders of the non-technical view have misunderstood how the phrase is used in the literature on phenomenal consciousness. So only a defender of the technical view can explain both why some believe that the ‘what-it’s-like’ phrase is technical and why others believe that it is non-technical. That is, only a defender of the technical view can explain why the technical vs. non-technical disagreement exists.

## 5.2. *Second Argument*

The second argument for the technical view is that a defender of this view can explain why some philosophers misunderstand what the ‘what-it’s-like’ phrase means in the literature on phenomenal consciousness, but it is not clear that a defender of the non-technical view can explain this.

One example of a philosopher who misunderstands what the phrase means in the literature on phenomenal consciousness is Gaskin (2019), who interprets the ‘what-it’s-like’ phrase in the general resemblance sense, without the resemblance being restricted to phenomenal properties. Since the mere fact that experiences resemble something entails nothing that is difficult to explain, Gaskin thinks there is no ‘hard problem’. He writes: ‘the ‘further phenomenon’ that is supposedly inexplicable by ordinary science and is said to be grounded in the what-it’s-likeness of experience presupposes an understanding of WIL talk that is based on a simple linguistic error’ (2019: 695). Gaskin also thinks that it is a mistake to describe zombies as not being like anything:

[W]hen Chalmers writes that ‘there is nothing it is like to be a zombie’ (2010, p. 107), we can reply that in the ordinary and correct sense of ‘is like’ this is just false. Zombies, as they are characterized in the philosophical literature, are very similar to us indeed—in fact, they are identical to us in all respects except, supposedly, one. Hence there are plenty of things it is like to be a zombie; and the one respect in which, allegedly, it is like nothing to be a zombie cannot, it would appear, be stated. At least, WIL locutions will not do the trick: for zombies are physically and functionally indistinguishable from normal human beings, which means that their experiences will boast the usual range of resemblances (properties), in fact, exactly the same ones as ours do (2019: 696).

But this gets things the wrong way around. Given the idea of conceptual freedom mentioned in chapter 1, there is no ‘correct sense of ‘is like’’, and the fact that words are ordinarily used a certain way does not put any constraints on how we can use them in philosophy. Thus, ordinary senses of the ‘what-it’s-like’ phrase, such as the resemblance sense, do not put constraints on there being a hard problem or how one should describe zombies. Rather, the use of the ‘what-it’s-like’ phrase in connection with the hard problem and zombies puts constraints on what the ‘what-it’s-like’ phrase can plausibly mean in these contexts. This claim is not just an unfounded opinion but is supported by the plausible ideal of a charitable interpretation. For if it is not the use of the ‘what-it’s-like’ phrase in the literature on phenomenal consciousness that determines what the phrase means in this literature, but rather uses of the phrase outside of this literature (which Gaskin suggests), then we might end up with a very uncharitable interpretation of what philosophers are saying (such as that of Gaskin).

A second example of a philosopher who misunderstands what the ‘what-it’s-like’ phrase means in the literature on phenomenal consciousness is Hacker (2002), who interprets the phrase as ‘a miscegenous crossing of the form of a judgment of similarity with the form of a request for an affective attitudinal characterization of an experience’ (2002: 166). In other words, Hacker thinks that the ‘what-it’s-like’ phrase in the context of phenomenal consciousness illicitly combines aspects from how a relevantly similar phrase is used in the context of picking out resemblance with how a relevantly similar phrase is used in the context of picking out an evaluative judgement. Perhaps an easier way to put this claim is that Hacker thinks there is a conflation of the resemblance sense and the evaluative-descriptive sense. Based on this, Hacker claims that ‘it is misconceived to suppose that one can circumscribe, let alone define, *conscious experience* in terms of there being something which it is like for a subject to have it’ (2002: 166) and that ‘The very expression ‘There is something it is like for a person to have it’ is malconstructed’ (2002: 167).

But again, this gets things the wrong way around. Ordinary meanings of the ‘what-it’s-like’ phrase, such as the resemblance sense and the evaluative-descriptive sense, do not put constraints on whether one can successfully use the phrase in connection with phenomenal consciousness. Rather, the use of the phrase in connection with phenomenal consciousness puts constraints on what the phrase can plausibly mean in this context. Why? Again, because the alternative is an uncharitable interpretation of what philosophers are saying.

A third example of a philosopher misunderstanding what the ‘what-it’s-like’ phrase means in the literature on phenomenal consciousness is Snowdon (2010), who interprets the phrase in the general non-evaluative-descriptive sense, not restricted to phenomenal properties

or the phenomenal ‘ways’ mental states are. Snowden therefore claims that the biconditional *a mental state m is phenomenal if and only if there is something it is like to be in m*<sup>12</sup> if one considers ‘going from left to right, then it is trivial’, since mental states, ‘like anything else, must be some way’ (2010: 25). But ‘[c]onsidered from right to left, [...] then it is not even true’, since from the claim that a mental state has a non-evaluative property or is some ‘way’, it does not follow that it is phenomenal (2010: 25).

Once again, this gets things the wrong way around. Ordinary meanings of the ‘what-it’s-like’ phrase, such as the non-evaluative-descriptive sense, do not put constraints on how one can truthfully characterize phenomenal consciousness with a biconditional. Rather, the fact that the ‘what-it’s-like’ phrase is used in biconditionals like the above puts constraints on how one can plausibly understand the ‘what-it’s-like’ phrase. Once again, the reason is that the alternative is an uncharitable interpretation of what philosophers are saying.

What explains these misunderstandings? Here is an explanation: the ‘what-it’s-like’ phrase has a technical meaning when it is used in the way that is characteristic in the literature on phenomenal consciousness, but philosophers who use the phrase this way typically do not make clear that it is technical, and this misleads philosophers like Gaskin, Hacker and Snowden to believe that it is used in an ordinary way in which it is not used. This explanation entails the technical view, so it is only open for a defender of that view and not for a defender of the non-technical view. In fact, it is not clear how a defender of the non-technical view can explain why philosophers like Gaskin, Hacker and Snowden misunderstand the phrase, given that these are all native English-speakers. For if the phrase had a completely ordinary or non-technical meaning in the literature on phenomenal consciousness, then one would not expect native English-speaking philosophers to misunderstand the phrase in this way. Rather, it should have been clear what the phrase means in the relevant philosophical literature. But if the phrase is technical and people who use it do not make clear that it is technical, then misunderstandings like the above are likely to occur.

## **6. The Arguments for the Non-Technical View and Why They are Unconvincing**

The previous section offered two arguments in favour of the technical view. But if there is equal support for the non-technical view, then there would be no reason to prefer the technical view.

---

<sup>12</sup> Snowden actually has a different terminology, so I have changed the biconditional to conform to the terminology I use in this chapter. But nothing substantive turns on this.

This section discusses the arguments for the non-technical view and argues that they are all unconvincing. Thus, we should prefer the technical view.

### 6.1. *First Argument*

The first argument for the non-technical view is that philosophers do not seem to be aware that the ‘what-it’s-like’ phrase is technical when used in relevant philosophical contexts, which one would have expected if it really was technical (Farrell 2016: 56-57, Stoljar 2016: 1184). But as noted above, there are several philosophers who explicitly claim that the ‘what-it’s-like’ phrase is technical. Several others use scare quotes (e.g. Block 1995: 230, Levine 2001: 4, Kammerer 2021: 846). That some are not aware that the phrase is technical could be because they conflate the philosophical use of the phrase with other uses, or because they are too entrenched in philosophy to reliably distinguish technical from non-technical language. So the fact that some philosophers are unaware that the ‘what-it’s-like’ phrase has a technical meaning does not show that it has a non-technical meaning.

### 6.2. *Second Argument*

The second argument for the non-technical view is that we should not multiply technical meanings unless it is necessary (Mehta 2021: 6). This argument can be understood in different ways, for as noted in §2, the defender of the non-technical view holds either that the ‘what-it’s-like’ phrase is to be understood in the non-evaluative-descriptive sense restricted to the phenomenal ‘ways’ mental states are, or in the resemblance sense restricted to phenomenal resemblance, or in a distinct ordinary sense – the phenomenal sense.

If the defender of the non-technical view holds that the ‘what-it’s-like’ phrase is used in the non-evaluative-descriptive sense or the resemblance sense, then it is true that a defender of the technical view ends up with accepting more meanings of the phrase. The reason is that the technical view posits a technical sense distinct from the three ordinary senses discussed in §2, while (this version of) the non-technical view says that the phrase is to be understood in either the non-evaluative-descriptive sense or the resemblance sense, which are both among the three ordinary senses discussed in §2. But the argument was that we should not multiply meanings *unless it is necessary*. And in this context, it *is* necessary to take the ‘what-it’s-like’ phrase in a technical sense that is distinct from the non-evaluative-descriptive sense and the resemblance sense, since otherwise we cannot explain why the technical vs. non-technical disagreement exists, or why some philosophers misunderstand what the ‘what-it’s-like’ phrase means in the literature on phenomenal consciousness.

If the defender of the non-technical view holds rather that the phrase is used in a distinct ordinary sense – the phenomenal sense – then defenders of both the technical and non-technical views end up accepting the same number of meanings. For in this case, both the technical view and the non-technical view posit a sense that is distinct from the three ordinary senses discussed in §2, it is just that the former claims that this sense is technical while the latter claims that it is non-technical. But then it is not clear what the argument for the non-technical view is, since it is not clear why it is better to multiply non-technical meanings rather than technical meanings. The mere claim that it is better to say that the meaning is non-technical is not a reason to prefer the non-technical view, it is merely an expression of one's preference for that view.

### 6.3. *Third Argument*

The third argument for the non-technical view is not stated in the literature, but the argument I have in mind is a natural one in this context, so I will discuss it nevertheless. The argument is that the non-technical view explains why philosophers started using the 'what-it's-like' phrase in the way that is characteristic in the literature on phenomenal consciousness, but it is not clear why they did this if it is technical.

However, I think there is an explanation. Those who started using the phrase in the way that is characteristic in the literature on phenomenal consciousness were familiar with the idea of phenomenal consciousness. One feature of this idea is that there is a gap between third-person knowledge and knowledge of phenomenal consciousness: one cannot know a phenomenal property unless one has had a mental state instantiating the property. In other words: experience is the best teacher. Another thing they knew was the expression 'you do not know what it is like before you have the experience', which, as discussed in §2, can be understood in the ability sense that is not concerned with phenomenal consciousness. In any case, the expression seems to describe the epistemic gap that is part of the concept of phenomenal consciousness – that experience is the best teacher. My hypothesis then, is that philosophers started using this expression to describe the epistemic gap that is part of the concept of phenomenal consciousness. And since 'what-it's-like' refers to the object of knowledge in this use of the above expression, they also started to use the 'what-it's-like' phrase in isolation – without being preceded by cognitive verbs like 'know' – in the way that is characteristic in the literature on phenomenal consciousness.

Were the philosophers who started using the 'what-it's-like' phrase this way aware that they were introducing a technical sense? It depends on what they thought about the expression 'you do not know what it is like before you have the experience'. Given the technical view, lay



people do not use this expression with the sense of the ‘what-it’s-like’ phrase that is relevant to phenomenal consciousness. Rather, they use the expression with a different sense of the phrase, plausibly the ability sense. I am neutral about whether the philosophers who started using the ‘what-it’s-like’ phrase in the way that is characteristic in the literature on phenomenal consciousness were aware that the above expression ordinarily involves the ability sense of ‘knowing what it’s like’ and not the sense of the ‘what-it’s-like’ phrase relevant to phenomenal consciousness. If they were aware of this, then they were likely aware that they were introducing a technical sense, but if they were not aware of this, then their linguistic innovation is based on misunderstanding the above expression.<sup>13</sup>

If I am right that the use of the ‘what-it’s-like’ phrase that is characteristic in the literature on phenomenal consciousness is modelled on the use of the phrase connected to the ability sense, then that explains an otherwise curious coincidence, namely that ‘knowing what it’s like’ is taken to capture two senses in which experience is the best teacher.

As mentioned above, Lewis (1988) argued that the sense in which experience is the best teacher is that it makes one acquire the abilities to imagine, recognize and remember the experience. His ‘ability hypothesis’ is that learning what it is like to see red consists in acquiring the abilities to imagine, recognize and remember seeing red, hence that there is a sense in which Mary learns what it is like to see red, even though (according to Lewis) phenomenal consciousness does not exist. However, realists about phenomenal consciousness will say that there is a different sense in which experience is the best teacher, and that there is a different sense in which Mary learns what it is like to see red, namely a sense concerning phenomenal consciousness.

This raises the following question: How come there are *two* distinct senses of ‘knowing/learning what it is like’ – the ability sense on the one hand, and the sense of ‘know’ in combination with the sense of the ‘what-it’s-like’ phrase relevant to phenomenal consciousness on the other – such that both are taken to capture a way in which experience is the best teacher? How come there are two distinct senses of ‘knowing what it’s like’, such that Mary is taken to be ignorant of what it is like to see red before the release but is taken to know what it is like after the release – in *both* senses? This might seem like a strange coincidence.

---

<sup>13</sup> Who were these philosophers? I do not know, but the earliest uses of the ‘what-it’s-like’ phrase exemplifying the use that is characteristic in the literature on phenomenal consciousness, which I have come across, are Farrell (1950: 181) and Sprigge (in Sprigge and Montefiore 1971: 167-168). Farrell (2016) claims that Wittgenstein (1980 §91) and Russell (1926) also used the phrase in this way, but I am not convinced, since I think their use of the phrase can be interpreted in the non-evaluative-descriptive sense restricted to the non-phenomenal ‘ways’ mental states are, or as connected to the ability sense.

But I claimed above that the ability sense captures a way in which experience is the best teacher, and that the phenomenal sense is modelled on the ability sense precisely because one wanted to capture a (different) way in which one took experience to be the best teacher. And if this is true, then it is no coincidence that both senses of ‘knowing/learning what it’s like’ are taken to capture a way in which experience is the best teacher and thus that Mary is taken to learn what it is like to see red in both senses. This makes the above explanation of why philosophers started to use the ‘what-it’s-like’ phrase in the characteristic way look very plausible.

To sum up then, neither argument for the non-technical view is convincing. Given the arguments for the technical view, we should believe that that ‘what-it’s-like’ phrase is technical.

## 7. Why the Technical View Matters

I have now argued that the ‘what-it’s-like’ phrase as it is characteristically used in the literature on phenomenal consciousness has a technical meaning. But why care whether it is technical or not? This section offers two reasons why the technical view matters.

### 7.1. *First Reason*

The first reason the technical view matters is that if it is true, then the ‘what-it’s-like’ phrase says nothing informative about phenomenal consciousness. However, the phrase is sometimes used in a way that makes it look as if it says something informative. For example, Kriegel says that ‘[p]henomenal consciousness is the property mental states, events, and processes have when, and only when, there is something it is like for their subject to undergo them, or be in them’ (2006: 58). The fact that Kriegel uses the ‘what-it’s-like’ phrase on the right-hand side of this biconditional suggests that he takes the phrase to say something informative about phenomenal consciousness, since it is uncommon to state biconditionals that are completely trivial or uninformative.

But if the ‘what-it’s-like’ phrase has a technical meaning here, then it says nothing informative about phenomenal consciousness. A biconditional in which token technical expressions occur on both sides *can* be informative *if* (a) the expressions are not co-extensional, or (b) the technical meaning of one of the expressions is more widely known than the technical meaning of the other expression. For example, if I say ‘S conveys an implicature with an utterance U if and only if S conveys a content with U that is not part of the literal meaning of U’, then I use a technical term on both the left (‘implicature’) and right (‘content’) side. This is an informative biconditional, since ‘implicature’ and ‘content’ are not co-extensional terms.

And if I coin a new technical term ‘schneomenal’ by saying that ‘a mental state *m* is schneomenally conscious if and only if *m* is phenomenally conscious’, then I use technical expressions on both sides. This is also an informative biconditional, since the technical meaning of ‘phenomenally conscious’ is more widely known than the technical meaning of ‘schneomenally conscious’.

But neither (a) nor (b) are satisfied in Kriegel’s biconditional above. States in which there is something it is like to be are all phenomenal states and all phenomenal states are states in which there is something it is like to be, so ‘having phenomenal consciousness’ and ‘being like something’ are co-extensional. And even though the *non-technical* meanings of the ‘what-it’s-like’ phrase are more widely known than the technical meaning of ‘phenomenal’, the (assuming the technical view) *technical* meaning of the ‘what-it’s-like’ phrase is not more widely known than the technical meaning of ‘phenomenal’. Rather, those within the relevant theoretical community who know the technical meaning of the ‘what-it’s-like’ phrase typically know the technical meaning of ‘phenomenal’ too. So formulations like that of Kriegel satisfy neither (a) nor (b), and thus they are not informative. In fact, Kriegel’s biconditional only says that mental states have the property phenomenal consciousness when and only when they are phenomenally conscious.

One might think the fact that people use the ‘what-it’s-like’ phrase in a way suggesting that it says something informative about phenomenal consciousness is evidence that the phrase is non-technical, since it is difficult to see why people would make statements like that of Kriegel if those statements are completely trivial. But a defender of the technical view can deny this and claim that it is rather evidence that those who use the phrase this way *believe* that it is non-technical. A similar point can be made about defenders of the technical view who do *not* use the phrase in a way suggesting that it says something informative about phenomenal consciousness. That is, the fact that these philosophers do not use the phrase in a way suggesting that it says something informative about phenomenal consciousness is not evidence that the phrase is technical. Rather, it is just evidence that these philosophers believe that it is technical. And as argued in §3.1., a defender of the technical view can explain both the belief that the ‘what-it’s-like’ phrase is technical and the belief that it is non-technical, but it is not clear that a defender of the non-technical view can explain both these beliefs. So philosophers (not) using the ‘what-it’s-like’ phrase in a way suggesting that it says something informative about

phenomenal consciousness does not add any data that count in favour of either view other than the data discussed in §3.1.<sup>14</sup>

In any case, the fact that the ‘what-it’s-like’ phrase says nothing informative about phenomenal consciousness is not just interesting in its own right, it also has a further consequence. Some worry that a lack of an informative description of phenomenal consciousness impedes serious theorizing about the alleged phenomenon. As Mandik (2016) comments with regard to realism vs. eliminativism about phenomenal consciousness: ‘the terms ‘qualia’, ‘phenomenal properties’, etc. lack sufficient content for anything informative to be said in either affirming or denying their existence. Affirming the existence of what? Denying the existence of what?’ (2016: 148). And as Rosenthal (2019) comments with regard to the hard problem: ‘Having no informative description matters. The hard problem is ‘why and how do physical processes in the brain give rise to conscious experience?’ Without a clear, tolerably accurate description of what conscious experience is, we cannot begin to address that question or even evaluate whether doing so would be difficult’ (2019: 202). If these authors are right, then the lack of an informative description of what phenomenal consciousness is impedes serious theorizing about phenomenal consciousness. This is a serious worry and raises the following challenge for philosophers who theorize about phenomenal consciousness: *either* explain what phenomenal consciousness is *or* explain why the above worry is misplaced. One may have thought that the ‘what-it’s-like’ phrase would suffice for an informative description, but since it is technical, it will not do the job and the challenge remains.

Of course, the claim that the ‘what-it’s-like’ phrase is technical does not entail that this challenge cannot be answered. One possible answer would be to take the semantics people suggest for the ‘what-it’s-like’ phrase (e.g. Hellie 2004, Stoljar 2016, Mehta 2021) as an explanation of what phenomenal consciousness is. A different possible answer would be to claim that we all have a prior understanding of what phenomenal consciousness is and that this

---

<sup>14</sup> One may think that this raises the question of why the discussion in 3.2. – about Gaskin (2019), Hacker (2002), and Snowdon (2010) who misunderstand what the ‘what-it’s-like’ phrase means – *does* add data that are not discussed in 3.1. The reason is that the misunderstandings of Gaskin, Hacker and Snowdon are misunderstandings about *what* the phrase means when it is used in the way that is characteristic in the literature on phenomenal consciousness. This makes whatever belief they have about technicality/non-technicality irrelevant for whether the meaning the phrase actually has in that literature is technical or not. Unlike the misunderstanding that the phrase says something informative about phenomenal consciousness, the misunderstandings of Gaskin, Hacker and Snowdon cannot be explained by a belief *de re* about the meaning the phrase actually has in the literature on phenomenal consciousness being technical or not. The reason is that whatever beliefs Gaskin, Hacker and Snowdon have about technicality/non-technicality are not *de re* about the meaning the phrase actually has in that literature. Rather, their beliefs about technicality/non-technicality are about *different* meanings, in terms of which Gaskin, Hacker and Snowdon falsely think one should understand the phrase in the relevant literature, but neither of which is the meaning of the phrase in that literature.

understanding is sufficient to engage in serious theorizing about the phenomenon (Block 1995, Chalmers 1996). I shall not evaluate these possible answers here, however. The present point is that merely appealing to the ‘what-it’s-like’ phrase is *not* an answer to the challenge.

### *7.2. Second Reason*

The second reason the technical view matters is that lay people’s use of the ‘what-it’s-like’ phrase provides no compelling evidence that they believe in phenomenal consciousness. The question of how widespread belief in phenomenal consciousness is, is not just an interesting sociological question but has consequences for the debate between realists and eliminativists about phenomenal consciousness, since realists typically claim that the idea of phenomenal consciousness is something ordinary and innocent. For example, Chalmers says that eliminativists ‘might say that we feel pain in a non-phenomenal way or non-experiential or non-conscious way. But this claim is of dubious coherence. In the ordinary sense of the word ‘feel’, to feel pain is to experience pain. And when one feels pain in this sense, there is something it is like to undergo the pain, almost by definition’ (2018: 53). But if the ‘what-it’s-like’ phrase is technical, then lay people’s use of the phrase is not compelling evidence that they believe in phenomenal consciousness. There may of course be other sources of evidence that lay people believe in phenomenal consciousness. But I argue in the next chapter that these other sources do not provide compelling evidence either. And if there is no compelling evidence that lay people believe in phenomenal consciousness, then one cannot dismiss eliminativism on the ground that it denies common sense. This might make eliminativism a more credible view than what is often assumed.

## **8. Conclusion**

I have argued that the ‘what-it’s-like’ phrase as it is used in the way that is characteristic in the literature on phenomenal consciousness has a technical meaning. This is not just interesting in its own right but has important consequences. Based on this claim, I have argued that the phrase says nothing informative about phenomenal consciousness. I have also argued that lay people’s use of the phrase provides no compelling evidence that they believe in phenomenal consciousness. These consequences have further consequences for debates about phenomenal consciousness. So the question of whether the ‘what-it’s-like’ phrase is technical is not just a linguistic curiosity but has significant consequences for first-order questions in the philosophy of mind.

## Chapter 3

### Eliminativism About Phenomenal Consciousness

**Abstract:** Eliminativism about phenomenal consciousness is the view that phenomenal consciousness does not exist. In this chapter, I clarify the eliminativist position by considering different definitions of phenomenality that can serve in a formulation of eliminativism and by arguing that it is consistent with several things often thought to be inconsistent with it. I also explain the motivation for eliminativism by arguing that it is ontologically simpler than non-reductive realism, and that it avoids the ‘hard problem’ facing reductive realism. Finally, I argue that there is no compelling evidence for lay belief in phenomenality, hence that we are currently not in a position to say whether common sense counts in favour of eliminativism or realism. I conclude that, for all realists have shown, eliminativism is a view worth taking seriously.

#### 1. Introduction

Eliminativism about phenomenal consciousness is the claim that phenomenal consciousness does not exist. This is often considered an incredible view, and one gets the impression that many philosophers do not even take it seriously. For example, Carruthers says that ‘it is beyond dispute that there are such things as qualia’ (2000: 15). Chalmers says that denying the existence of phenomenal consciousness ‘is the sort of thing that can only be done by a philosopher’ (1996: 188). Levine says that ‘the qualiaphobe’s denial of conscious experience seems ludicrous’ (2001: 131). Nida-Rümelin says that eliminativism ‘strikes many people as absurd and that reaction is perfectly adequate. There is so much to do in philosophy in search of the truth that one should not lose too much time with absurd theories’ (2016: 160). And Strawson (2017) says that ‘[t]here occurred in the twentieth century the most remarkable episode in the whole history of ideas – the whole history of human thought. A number of thinkers denied the existence of something we know with certainty to exist: consciousness, conscious experience’ (quoted in Chalmers 2018: 55). These claims seem to suggest that eliminativism is so implausible that we do not even need to take the view seriously.

The aim of this chapter is to show that, for all realists have shown, eliminativism *is* a view worth taking seriously. To this end, I clarify the eliminativist position and argue that it is consistent with several things often thought to be inconsistent with it, such as there being

something it is like to be in mental states (in the ordinary senses of the ‘what-it’s-like’ phrase), the existence of feelings and experiences (in functional senses of these terms), and what we know to be common sense. I also explain the motivation for eliminativism by showing that it has theoretical advantages vis-à-vis both non-reductive and (arguably) reductive realism.

The structure of the chapter is this. §2 clarifies the eliminativist position by considering different definitions of phenomenality that can serve in a formulation of eliminativism and by arguing that it is consistent with several things often taken to be inconsistent with it. §3 explains the motivation for eliminativism by arguing that it is ontologically simpler than non-reductive realism, and that it avoids the ‘hard problem’ facing reductive realism. §4 considers the objection that realism accords better with common sense and argues that there is no compelling evidence that belief in phenomenality is part of common sense. §5 concludes.

## **2. What is Eliminativism About Phenomenality?**

This section clarifies the eliminativist position. 2.1. discusses potential definitions of phenomenality, realism and eliminativism, and 2.2. argues that eliminativism is consistent with several things often thought to be inconsistent with it.

### *2.1. Defining Eliminativism*

By ‘eliminativism’ I mean ontological eliminativism, not discourse eliminativism. Ontological eliminativism is the claim that phenomenality does not exist, while discourse eliminativism is the claim that talk about phenomenality should be eliminated from science or philosophy.<sup>15</sup> Neither form of eliminativism entails the other. But given the idea of conceptual freedom outlined in chapter 1, according to which we can speak however we want to, discourse eliminativism seems very hard to motivate. So I shall focus on ontological eliminativism, which is consistent with the claim that we should continue to use the word ‘phenomenal’ and its cognates in scientific and philosophical practice.

In the previous chapter, I said that I am only concerned with phenomenality in the sense in which non-reductive realists and type-B materialists (which is what I refer to as ‘reductive realists’) use the word, which excludes phenomenality in the functional sense in which type-A materialists (analytic functionalists) use the word. This constraint only says that phenomenality is not defined functionally. But many things are not defined functionally, so the non-functional

---

<sup>15</sup> For a defence of discourse eliminativism, see Irvine (2012).

constraint does not say much about what phenomenality really is. Thus, it does not say much about what realism is, and since eliminativism is the denial of realism, it does not say much about what eliminativism is either.

However, when discussing eliminativism it would be helpful with an informative definition of the position. And in order to avoid verbal dispute, the eliminativist must mean the same with ‘phenomenality’ and its cognates as the realist. So if we knew more precisely what the realist means with ‘phenomenality’ and its cognates, then we could formulate eliminativism as the denial of precisely that which the realist associates with these terms. Is there an informative definition of phenomenality that can ground a formulation of eliminativism?

One proposal is due to Dennett (1988), who characterizes phenomenality (or ‘qualia’ in Dennett’s terminology) as ineffable (not describable), intrinsic (unanalysable), private (interpersonal comparisons are impossible), and directly accessible (for the subject). If this was what realists meant by ‘phenomenality’ and its cognates, then Dennett’s view – that nothing satisfies this characterization – would be a formulation of eliminativism. But even though some realists believe that phenomenality has the above-mentioned properties, several realists deny this (e.g. Levine 2001: 132 and Tye 2018). So on the assumption that realists share one concept of phenomenality and talk about the same thing, Dennett’s (1988) position does not serve as a formulation of eliminativism.

A second proposal involves the notion of zombifiability, since several philosophers have suggested that we can understand phenomenality in terms of zombifiability. On this view, a mental state is phenomenal if it has properties that we can conceive of physical/functional duplicates as lacking (Horgan 2011: 61, Kriegel 2015: 52-53, McClelland 2016: 540-541, Carruthers and Veillet 2017: 80). Eliminativism could then be the thesis that there are no zombifiable properties. But this is problematic, for two reasons.

Firstly, neither ideal nor *prima facie* conceivability<sup>16</sup> seems to work in this context. Ideal conceivability is conceivability that cannot be undermined on ideal reflection, i.e. the reflection of an omniscient and non-mistake-making reasoner. But if zombifiability involves *ideal* conceivability, then zombifiability is arguably too strong to be necessary for phenomenality, since most realists do not commit to the ideal conceivability of zombies. *Prima facie* conceivability involves only a disposition to judge that something is conceivable, which *can* be undermined on ideal reflection. But if zombifiability involves only *prima facie* conceivability, then realism would be trivially true, since several philosophers are disposed to judge that

---

<sup>16</sup> See Chalmers (2002) for more on this distinction.



zombies are conceivable. But surely the truth of realism and falsity of eliminativism is not proved by the mere fact that some philosophers have the disposition to judge that zombies are conceivable.

Secondly, regardless of whether we mean ideal or *prima facie* conceivability – it can be argued that we do not know whether zombies are conceivable before we know whether physicalism is true (Brown 2010). The point is a general one about what we can conceive of after we have discovered an identity to hold. So for example, it is arguable that we cannot conceive of water that is not H<sub>2</sub>O after it was discovered that they are identical. Similarly, the idea is that if physicalism turns out to be true, then it may be that we cannot conceive of zombies, since then we would know that it is impossible for physical/functional duplicates of us to lack any properties that we have. If so, the discovered truth of physicalism would rule out the existence of phenomenality, given the zombifiability conception of phenomenality. But most realists do not take the truth of realism to depend on whether we know that physicalism is true. So zombifiability does not seem to help in formulating eliminativism.

A third proposal is defining phenomenality as *what it is like* to be in a mental state, and then define eliminativism as the thesis that mental states are not like anything or that what-it's-like-ness does not exist. But I argued in the previous chapter that the 'what-it's-like' phrase as it is used in this context has a technical meaning and cannot serve in an informative definition of phenomenality. In fact, the claim that mental states are phenomenal if and only if there is something it is like to be in those states only says that mental states are phenomenal if and only if they are phenomenal.

A fourth proposal is defining phenomenality using the words 'feeling' or 'experience'. For example, one could say that a mental state is phenomenally conscious if and only if it is an experience or a feeling. This would be informative, *if* 'feeling' and 'experience' – in the ordinary senses of these terms – concerned phenomenality. But if these words ordinarily do not concern phenomenality, for example, if they are ordinarily understood functionally and are technical in connection with phenomenality, then they cannot serve in an informative definition of phenomenality. The reason is the same as for the 'what-it's-like' phrase: a biconditional in which token technical expressions occur on both sides *can* be informative *if* (a) the expressions are not co-extensional, or (b) the technical meaning of one of the expressions is more widely known than the technical meaning of the other expression. And if 'feeling' and 'experience' are not ordinarily understood as concerning phenomenality, then the above definition satisfies neither (a) nor (b), in which case the above definition is uninformative. That is, all mental states that are phenomenal are experiences or feelings, and all experiences and feelings are

phenomenal (otherwise the above definition would be false). And even though the non-technical meanings of ‘experience’ and ‘feeling’ are more widely known than ‘phenomenal’, the (assuming that ‘experience’ and ‘feeling’ are technical in connection with phenomenality) technical meanings of ‘experience’ and ‘feeling’ are not more widely known than the technical meaning of ‘phenomenal’. Rather, those within the relevant theoretical community who know the meaning of ‘experience’ and ‘feeling’ typically know the meaning of ‘phenomenal’ too. So if ‘feeling’ and ‘experience’ are technical in connection with phenomenal consciousness, then the above definition only says that mental states are phenomenal if and only if they are phenomenal, and so is not informative.

To determine whether the definition of phenomenality in terms of experience or feeling is an informative definition or not one must therefore first determine whether ‘experience’ and ‘feeling’ are ordinarily understood as concerning phenomenality, in other words, whether ‘experience’ and ‘feeling’ are ordinary or technical in connection with phenomenality. But philosophers disagree over this question. Some claim that ‘experience’ and ‘feeling’ ordinarily concern phenomenality (Chalmers 2018: 53), while others deny that the ordinary use of these expressions concern anything but functionally defined properties (Mandik 2016: 141-142, Sytsma and Machery 2010: 309fn11, Sytsma and Ozdemir 2019: 245). These claims are typically not made with much compelling argument, however. So what reason is there to believe one rather than the other?

There is a relevant debate the outcome of which may provide evidence for either view, namely the debate about whether lay people believe in phenomenality. For if lay people *do* believe in phenomenality, then it is plausible that ‘feeling’ and ‘experience’ are ordinarily concerned with phenomenality, given that these words are used in studies designed to determine whether subjects believe in phenomenality (see §4). But if lay people do *not* believe in phenomenality, then it is not plausible that ‘feeling’ and ‘experience’ are ordinarily concerned with phenomenality, since it is implausible that lay people ordinarily use these words in seemingly referential ways without believing that they refer.

The problem is that, as I argue in §4, we are currently not in a position to tell whether lay people believe in phenomenality or not, so this debate is by no means settled. Therefore, we are not in a position to tell whether ‘feeling’ and ‘experience’ are ordinarily concerned with phenomenality. This has the upshot that we do not know whether the definition of phenomenality in terms of feeling or experience is informative, it all seems to depend on the outcome of the debate about whether lay people believe in phenomenality. But it is not ideal to work with a definition of phenomenality that we do not know whether is informative and whose

informativeness depends on the outcome of a debate separate from that between realism and eliminativism. So even though the definition of phenomenality in terms of feeling or experience *could* be informative, I shall not rely on this definition.

In what follows, I shall assume that we do not need an informative definition of phenomenality and that it suffices to give illustrations or typical examples and characterizations in synonymous terms (Block 1995: 230, Chalmers 1996: 4). As Block (1978) says when imagining someone asking what is meant by ‘phenomenal’ and its cognates: ‘If you got to ask, you ain’t never gonna get to know!’ (1978: 241). The idea is that we all have a prior understanding of what is meant and that this understanding is sufficient to engage in theorizing about the alleged phenomenon without an informative definition. The following discussion of eliminativism assumes that this is true.<sup>17</sup>

## 2.2. Clarifications

Eliminativism is inconsistent with mental states being like something – in the technical sense relevant to phenomenality. But importantly, this is consistent with mental states being like something in the three ordinary senses discussed in chapter 2, namely the evaluative-descriptive sense, the non-evaluative-descriptive sense (restricted to the non-phenomenal ‘ways’ mental states are), and the resemblance sense (restricted to non-phenomenal resemblance). Even though the eliminativist claims that there is nothing it is like to see red – in the technical sense relevant to phenomenality – there may be something it is like in these three ordinary senses. What it is like to see red in the evaluative-descriptive sense might be appealing, what it is like in the non-evaluative-descriptive sense might be making one able to distinguish ripe from unripe tomatoes, and what it is like in the resemblance sense might be seeing orange.

The fact that eliminativism is consistent with mental states being like something in the ordinary senses of the ‘what-it’s-like’ phrase is often underappreciated by eliminativists who use the ‘what-it’s-like’ phrase in formulating their view. Frankish says that ‘experiences do not really have [...] ‘what-it’s-like’ properties’ (Frankish 2016a: 15), Pereboom says that ‘the what-it’s-like features of sensory states [...] are illusory in that they don’t exist’ (Pereboom 2016: 173), and Kammerer says that ‘there is *nothing it is like* to be in any of our mental states’ (Kammerer 2021: 847). Neither of these authors specify what sense of the ‘what-it’s-like’ phrase they have in mind. But distinguishing the (technical) sense of the phrase relevant to phenomenality from the ordinary senses is not just important for clarity about the eliminativist

---

<sup>17</sup> But see Mandik (2016) and Rosenthal (2019) for criticisms of this idea.

position. It may also help to avoid the (as I argue below) unjustified impression that the eliminativist is denying something obvious, given that mental states obviously are like something in the ordinary senses, but not so obviously are like anything in the technical sense relevant to phenomenality. That is, while it may sound fair enough to say that one denies qualia or raw feels, to say that mental states are not like anything might create the erroneous impression that this is something more controversial than denying qualia or raw feels. But once we realize that the ‘what-it’s-like’ phrase in this context is technical and just means the same as ‘qualia’ and ‘raw feels’, and that eliminativism is consistent with mental states being like something in the ordinary senses of the ‘what-it’s-like’ phrase, then we can avoid this impression.

Another potential misunderstanding about eliminativism is this. It is sometimes claimed that eliminativism is inconsistent with feeling pain – in the ordinary sense of these terms (Chalmers 2018: 53), perhaps suggesting that eliminativism is inconsistent with the existence of feelings and experiences generally – in ordinary senses of these terms. But as noted above, I argue in §4 that it is controversial whether belief in phenomenality is something ordinary, in which case the existence of feelings and experiences – in the ordinary senses of these terms – may not be inconsistent with eliminativism. For eliminativism is not inconsistent with feelings and experiences in functional senses of these terms, since eliminativists only deny that the non-functional concept of phenomenality refers, they do not deny that other functional concepts refer. So given that it is an open question whether belief in phenomenality is ordinary (which I argue in §4), there is no reason to think that the ordinary senses of ‘feeling’ and ‘experience’ concern phenomenality, in which case eliminativism may be consistent with the existence of feelings and experiences in the ordinary senses of these terms.

Similarly for pain. I argue in chapter 5 that we can distinguish pain in the locatable sense – i.e. whatever ‘pain’ refers to in sentences like ‘There is a pain in my leg’ – from pain in the phenomenal sense. Even though eliminativism is inconsistent with the existence of pain in the phenomenal sense, it is consistent with the existence of pain in the locatable sense. According to the eliminativist, you can have a pain in your leg, it is just that there is nothing it is like – in the technical sense relevant to phenomenality – to have a pain in your leg. And since it is controversial whether belief in phenomenality is something ordinary, there is no reason to think that the ordinary sense of ‘pain’ is the phenomenal sense rather than the locatable sense, in which case eliminativism may be consistent with the existence of pain in the ordinary sense.

### 3. The Motivation for Eliminativism

This section explains the motivation for eliminativism. I first consider the main advantage vis-à-vis non-reductive realism, then the main advantage vis-à-vis reductive realism.

#### 3.1. *Eliminativism vs. Non-Reductive Realism*

The main motivation to adopt eliminativism rather than non-reductive realism is ontological simplicity: eliminativism posits fewer types of properties or entities than non-reductive realism (cf. Frankish 2016: 24).<sup>18</sup> Put a bit more formally:

- (P1) If a theory is ontologically simpler than another theory, then we should prefer the simpler theory, unless there are independently good reasons to adopt the less simple theory.
- (P2) Eliminativism is ontologically simpler than non-reductive realism.
- (P3) There are no independently good reasons to adopt non-reductive realism.
- (C) Therefore, we should prefer eliminativism.

(P1) is widely assumed to be true.<sup>19</sup> The truth of (P2) is clear from the fact that eliminativists deny the existence of phenomenal properties. The most controversial premise seems to be (P3). Are there really no good reasons to adopt non-reductive realism rather than eliminativism?

A non-reductive realist might think that there is antecedent support for realism, i.e. that phenomenality is a datum (Chalmers 2018: 53-55). The idea is then presumably that the existence of phenomenality is not invoked to explain something else, such that its support depends on the success of its explanatory role. Rather, the existence of phenomenality is itself in need of explanation, and as such it not questionable but enjoys antecedent support, i.e. support that is determined to hold independent of the debate between realists and eliminativists. For the only phenomena in need of explanation are those already determined to have support, since there is no need to explain a phenomenon whose existence lacks support. So if phenomenality is a datum, then that it is because its existence enjoys antecedent support. An eliminativist defending (P3) must therefore deny that there is any antecedent support for realism and hence deny that the existence of phenomenality is a datum. The question is then: what is

---

<sup>18</sup> Frankish (2016) also offers a ‘debunking argument’, which has been discussed and criticized elsewhere (e.g. Chalmers 2018), so I shall not discuss that argument here.

<sup>19</sup> Though it has been challenged (e.g. Willard 2014).

the alleged antecedent support for realism that (according to the realist) makes the existence of phenomenality a datum?

Perhaps the ‘antecedent support’ the realist has in mind is that she is acquainted with phenomenality – in the sense that she has direct and unmediated access to it. If this is what the realist means by ‘antecedent support’, then that begs the question against the eliminativist, since whether one is acquainted with phenomenality is not settled independent of the debate between realists and eliminativists. For insofar as eliminativists deny the existence of phenomenality they also deny that one has direct and unmediated access to it. So if ‘antecedent support’ means acquaintance with phenomenality, then eliminativists will claim that the realist’s appeal to ‘antecedent support’ is question-begging.

If the realist does not think that acquaintance with phenomenality provides antecedent support for eliminativism, perhaps she thinks that some other phenomena – whose existence does not depend on phenomenality – provides antecedent support. For example, one might think the best explanation of people’s intuitions, judgements, and beliefs that phenomenality exists appeals to the existence of phenomenality. Unlike acquaintance with phenomenality, these phenomena are distinct from, and do not depend on, the existence of phenomenality. So if the best explanation of people’s intuitions, judgements and beliefs that phenomenality exists appeals to the existence of phenomenality, then this would provide antecedent support for realism and thereby make the existence of phenomenality a datum. But in the next chapter I discuss various eliminativism-friendly explanations of why people have intuitions, judgements and beliefs that phenomenality exists.

The final phenomenon I can think of that realists might claim provide antecedent support for the existence of phenomenality is introspective representations of phenomenality – in a sense of ‘introspective representation’ that is neither direct acquaintance nor a cognitive state like an intuition, judgement or belief. For the realist might think that she has indirect and mediated access to phenomenality rather than direct and unmediated, via some introspective representation of phenomenality, e.g. a non-cognitive or quasi-perceptual representation. Unlike acquaintance with phenomenality, such a representation is distinct from, and does not depend on the existence of, phenomenality. So if the best explanation of why such a representation exists appeals to the existence of phenomenality, then this would provide antecedent support for realism and thereby make the existence of phenomenality a datum. But there are two eliminativist responses available.

One is to adopt *illusionism*, according to which the belief in phenomenality is based on a non-cognitive representation of phenomenality, but one which is non-veridical, i.e. what

illusionists call an ‘illusion’ of phenomenality. Non-cognitive representations of phenomenality would only provide antecedent support for realism if the best explanation of them appealed to phenomenality. However, illusionists typically claim that there is an alternative explanation, and hence that the existence of non-cognitive or quasi-perceptual representations of phenomenality provides no support for realism. The illusionist response to the realist’s appeal to non-cognitive representations of phenomenality is then to grant their existence but deny that their existence provides any antecedent support for realism.

The other response to the realist’s appeal to non-cognitive representations of phenomenality is to deny that there is reason to believe in such representations in the first place, which can be done by, for example, a defender of *educationism*. According to educationism, the belief in phenomenality is based on exposure to theoretical ideas. An educationist might claim that the belief in non-cognitive representations of phenomenality is also based on exposure to theoretical ideas: it is confabulation about why people believe in phenomenality. Unlike the illusionist, the educationist can grant that if there were non-cognitive representations of phenomenality, then they would provide antecedent support for realism. But the educationist can claim that there is no reason to believe in such representations. The response to the realist is then not to grant the existence of non-cognitive representations of phenomenality and claim that there is an alternative eliminativist-friendly explanation of them. Rather, it is to claim that there is no reason to believe there are such representations in the first place. So there is still no antecedent support for realism that would make the existence of phenomenality a datum.<sup>20</sup>

Regardless of whether the realist appeals to acquaintance, intuitions, judgements, beliefs, or non-cognitive introspective representations as ‘antecedent support’ for realism then, there are eliminativist responses available. I am not sure what else realists can mean by ‘antecedent support’, such that the appeal to antecedent support both supports realism and is non-question-begging against eliminativism.

One analogy that might be instructive in this context is the disagreement between Moore and the external world sceptic when Moore gestures with his hands to give a proof of the external world (cf. Chalmers 2018: 53). One who doubts the existence of the external world will deny that the existence of Moore’s hands is a datum, i.e. that the existence of Moore’s hands enjoys antecedent support that holds independent of the debate between Moore and the external world sceptic. Analogous to the realist above, Moore might insist that he has direct

---

<sup>20</sup> I discuss illusionism and educationism in more detail in chapter 4. The point of introducing them here is just to reply to the realist’s claim that ‘introspective representations’ provide antecedent support for realism.

acquaintance with his hands. But this depends on the existence of his hands, so it is not independent of the debate between Moore and the external world sceptic, in which case acquaintance does not provide antecedent support for the existence of Moore's hands. Instead of appealing to acquaintance with his hands, Moore might appeal to perceptual representations of his hands and claim that the best explanation of these representations appeals to his hands. But the sceptic can reply like the illusionist and grant the existence of these representations but deny that they provide support for belief in the external world. The sceptic will say that one need not appeal to Moore's hands to explain the existence of perceptual representations of his hands, since the existence of these mental representations can be explained by a sceptical scenario instead. Moore's proof is therefore generally considered either question-begging or ineffective as a dialectical move when debating the external world sceptic.<sup>21</sup>

Of course, one might say that Moore succeeds *by his own lights*. That is, if one assumes that the external world exists, then it is true that there were two hands there when Moore was gesturing with his hands. But the external world sceptic doubts that the external world exists, so this assumption begs the question against external world scepticism. This charge of begging the question can be true even if Moore – by his own lights – may be right that he has two hands. Another way to put this point is to say that Moore's appeal to his two hands is not question-begging *simpliciter*, for there is no such thing as begging the question *simpliciter*, one can only beg the question with respect to someone or something. But insofar as the existence of his hands entails the existence of the external world, then the appeal to his hands *does* beg the question with respect to external world scepticism. The same point applies to realism about phenomenality. If the realist assumes that phenomenality exists, then she succeeds by her own lights in claiming that phenomenality exists. This assumption is not question-begging *simpliciter*. But it *does* beg the question against eliminativism.

To sum up then, the eliminativist seems to be able to defend all three premises of the argument above. (P1) and (P2) are not very controversial. Given that the phenomena realists can appeal to in support of their view either depend on the existence of phenomenality, or can be explained, or denied, by eliminativists, the eliminativist seems able to offer a reasonable defence of (P3) as well. If this is true, the eliminativist seems to have a solid argument against the non-reductive realist.

---

<sup>21</sup> See Carter (2012) for a review of different diagnoses as to exactly what goes wrong in Moore's proof.



### 3.2. *Eliminativism vs. Reductive Realism*

Eliminativism is consistent with the existence of the physical/functional properties with which reductive realists identify phenomenality, so ontological simplicity is not a reason to be an eliminativist rather than a reductive realist. The motivation to adopt eliminativism rather than reductive realism is that eliminativism avoids the problems associated with reductive realism. The major problem perhaps, is the ‘hard problem’ of explaining how and why physical processes give rise to phenomenality, which is commonly thought to arise for both reductive and non-reductive realists, on the ground that both hold that there is an epistemic gap between phenomenality and physical/functional properties, i.e. that one cannot derive facts about phenomenal properties from facts about physical/functional properties *a priori* (cf. Chalmers 2003: 112).<sup>22</sup> Some reductive realists are optimistic towards solving the problem (e.g. Carruthers 2000) but others are less optimistic (e.g. Levine 2001, Block 2002). In any case, eliminativism has an advantage over reductive realism, since it avoids the ‘hard problem’.

Related to this is the fact that eliminativism offers a solution to what Chalmers (2018) calls the ‘meta-problem’ of phenomenal consciousness, i.e. ‘the problem of explaining why we think [phenomenal] consciousness is hard to explain’ (2018: 6). It is ‘the problem of explaining problem intuitions: intuitions that reflect our sense that there is some sort of special problem involving [phenomenal] consciousness’ (2018: 12). The problem involving consciousness referred to here is the ‘hard problem’. The meta-problem is then to explain why people think the hard problem of consciousness is hard, which should be done in topic neutral terms, i.e. without mentioning phenomenal consciousness (2018: 16). Eliminativism offers the following solution to the meta-problem: people find the hard problem of phenomenal consciousness hard, because there is no solution to it. If one struggles with a problem for which, unbeknownst to one, there is no solution, one is likely to find it hard. For example, if you do not know that there is no present King of France you may find it very hard to find out whether he is bald or not. But there is no answer to that question and the reason is that there is no present King of France. Similarly, the eliminativist claims that there is no solution to the ‘hard problem’, since it is (according to the eliminativist) a pseudo-problem. The reason it is a pseudo-problem is that there is no such thing as phenomenal consciousness. Of course, some unsolvable ‘problems’ seem like a piece of nonsense or sophistry rather than hard problems, but if they seem this way

---

<sup>22</sup> Papineau (2011) denies that the epistemic gap entails the explanatory gap, but since this entailment is widely assumed to hold, I assume it in what follows.

one typically knows that they are unsolvable. The hard problem seems hard to realists, because it, *unbeknownst* to them, is unsolvable.<sup>23</sup>

This explanation of why the hard problem seems hard is only available for an eliminativist and not for a reductive realist who commits to there being a hard problem. This, like the fact that eliminativism avoids the hard problem, seems to point out an advantage for eliminativism vis-à-vis reductive realism.

#### 4. How Widespread is Belief in Phenomenality?

One potential objection to the above arguments is appealing to common sense and say that realism accords better with common sense (i.e. the intuitions or beliefs of the majority of the population) than eliminativism. For example, Chalmers says that ‘The intuition [that phenomenal consciousness exists] appears to be shared by the large majority of philosophers, scientists, and others; and it is so strong that to deny it, a type-A materialist [i.e. an eliminativist] needs exceptionally powerful arguments’ (Chalmers 2003: 112). As mentioned in the previous chapter, he also says that eliminativists:

might say they do not intend to deny that we feel pain. For example, they might say that we feel pain in a non-phenomenal way or non-experiential or non-conscious way. But this claim is of dubious coherence. In the ordinary sense of the word ‘feel’, to feel pain is to experience pain. And when one feels pain in this sense, there is something it is like to undergo the pain, almost by definition (2018: 53).

Chalmers’ idea is that belief in phenomenality belongs to common sense: in the ordinary sense of ‘feeling pain’, there is something it is like – in the sense relevant to phenomenal consciousness – to undergo the pain ‘almost by definition’. This, according to Chalmers, puts realism in a better position than eliminativism.

---

<sup>23</sup> Chalmers (2018) focuses mostly on the above formulation of the meta-problem, but he also offers other (non-equivalent) formulations, for instance, that the meta-problem is the problem of explaining why we think that there is a problem of [phenomenal] consciousness (2018: 6). He also says the meta-problem is the problem of explaining other intuitions, like metaphysical intuitions (about phenomenality being non-physical), knowledge intuitions (about first-person knowledge being special), modal intuitions (about zombies being possible or conceivable), etc. These problems are different from the problem of explaining why people think the problem of phenomenal consciousness is *hard*, and they require a different eliminativist explanation than the above. Perhaps the eliminativist can offer an explanation by appealing to the theories discussed in chapter 4. But if not, it is at least fair to say that eliminativism offers a solution to a *core aspect* of the meta-problem.

This argument can be made regardless of whether one has reductive or non-reductive realism in mind. Applied to the arguments of the previous section, the non-reductive realist can object to (P3) in the eliminativist argument and say that realism accords better with common sense, so even though eliminativism is ontologically simpler than non-reductive realism, we should still prefer non-reductive realism. And the reductive realist can say that even though eliminativism avoids the ‘hard problem’, we should still prefer reductive realism, since realism accords better with common sense.

But it is controversial how widespread intuitions or beliefs about phenomenality are. In fact, the eliminativist can deny that there is compelling evidence that they are widespread among lay people. What evidence is there that lay people have intuitions or beliefs about phenomenality?

I shall restrict my focus here to *empirical studies* which are taken to offer evidence for lay belief in phenomenality. By restricting my focus this way, I shall not examine certain things that some may have thought were potential evidence for lay belief in phenomenality.

Firstly, I shall not focus on the fact that lay people use the ‘what-it’s-like’ phrase, because, as argued in chapter 2, such talk can be understood as involving the evaluative-descriptive sense, the non-evaluative-descriptive sense restricted to the non-phenomenal ‘ways’ things are, the resemblance sense restricted to non-phenomenal resemblance, or the ability sense. None of these senses can be the sense of the ‘what-it’s-like’ phrase relevant to phenomenality, so if one can interpret lay people’s use of the phrase as involving these senses, then lay people’s use of the phrase is not compelling evidence that they believe in phenomenality.

Secondly, I shall not focus on the fact that lay people use the word ‘consciousness’, because – absent any reason to believe otherwise – talk involving ‘consciousness’ does not indicate belief in phenomenal consciousness rather than access-consciousness, self-consciousness or many of the other things called ‘consciousness’.<sup>24</sup>

Thirdly, I shall not focus on the fact that lay people use words like ‘feeling’ (of e.g. pain) and ‘experience’ (of e.g. red), because, as argued above, there is no reason to think that feelings and experiences in the ordinary senses of these terms concern phenomenality independent of the empirical studies to be discussed. Of course, a realist (such as Chalmers 2018: 53 quoted above) may think feelings and experiences in the ordinary senses of these terms concern phenomenality. But the point is that we need evidence for this claim, such as empirical

---

<sup>24</sup> See Block (1995) for more on different notions of consciousness.

studies, in which case it is this evidence and not the fact that lay people use words like ‘feeling’ and ‘experience’, that supports the claim that lay people believe in phenomenality. Consequently, I shall neither focus on the fact that lay people use the ‘what-it’s-like’ phrase, nor words like ‘consciousness’, ‘feeling’ and ‘experience’. I shall only focus on empirical studies.

The five most relevant studies were conducted by Knobe and Prinz (2008), Peressini (2014), Gottlieb and Lombrozo (2018), Diaz (2021), and Gregory et al. (submitted). Most of these authors believe that their studies offer compelling evidence for lay belief in phenomenality. But I shall argue that the data are ambiguous.

Knobe and Prinz (2008: 74-75) asked subjects to evaluate the following sentences:

- (1) Acme Corp. wants to change its corporate image
- (2) Acme Corp. believes that its profit margin will soon increase
- (3) Acme Corp. is now vividly imagining a purple square
- (4) Acme Corp. is feeling excruciating pain

Subjects generally found (1) and (2) to be natural sounding, but they found (3) and (4) to be weird sounding. Knobe and Prinz’ explanation for why subjects treat (1) and (2) differently from (3) and (4) is that they take feeling pain and imagining a purple square to be phenomenal states, but they take wanting and believing to be non-phenomenal states (2008: 75). If this were the only or the best explanation of the data, then there would be compelling evidence for lay belief in phenomenality. But Sytsma and Machery (2010) point out that an equally suitable explanation for why subjects treat (1) and (2) differently to (3) and (4) is that subjects believe that corporations lack the functional properties needed to instantiate imagining a purple square and feeling pain. Thus, Knobe and Prinz’ study does not show that lay people believe in phenomenality.

Peressini (2014) asked subjects the following question:

Can we ever be sure that you see red the way another person does?

Peressini found that subjects tended to give negative responses, which Peressini explains by subjects believing in phenomenality (2014: 877). Peressini’s reasoning seems to be that subjects know that we *can* be sure that the physical/functional properties involved in seeing red are interpersonally identical, so if they believe that we cannot be sure that you see red the way

another person does, that must be because they believe that the way one sees red does not supervene on physical/functional properties. Belief in properties that do not supervene on physical/functional properties is certainly very close to non-reductivist belief in phenomenal properties, so if Peressini's explanation of the data were the only or best explanation, then there would be compelling evidence that lay people believe in phenomenality. But an equally suitable explanation of why subjects gave negative responses to the question is that they believe that it is practically impossible to be sure that the physical/functional properties involved in seeing red are identical, and thus they believe that we cannot ever be sure that you see red the way another person does. Consequently, Peressini's study does not show that lay people believe in phenomenality.

Gottlieb and Lombrozo (2018) asked subjects whether science can fully explain various (46 in total) mental phenomena, such as having depression, logical reasoning, falling in love, having an addiction, making mathematical errors, and using one's imagination. First, they introduced two notions: 'privileged first-person access' and 'introspection'. This is how they describe privileged first-person access: 'Only an individual him- or herself can know that he or she is experiencing \_\_: an outside observer might be able to guess but can't truly know' (2018: 124). And this is how they describe introspection: 'the examination of one's own internal feelings or reflection' (2018: 124). When asking subjects whether science can fully explain mental phenomena, they found that the extent to which subjects agreed that a phenomenon involves privileged first-person access and introspection correlated with the extent to which subjects agreed that the phenomenon was difficult to explain. In other words, subjects thought that mental phenomena involving privileged first-person access and introspection are more difficult to explain than mental phenomena not involving privileged first-person access and introspection (2018: 124). Chalmers interprets these results as indicating that subjects think that phenomenal consciousness is hard to explain (2018: 14).<sup>25</sup> His reasoning seems to be that privileged first-person access entails that other people than the subject of a mental state *in principle* cannot know what mental state he or she is undergoing, which is reminiscent of some things realists say about phenomenality, such as inverts being possible. But Gottlieb and Lombrozo's formulation of privileged first-person access does not make clear whether they mean this, or rather just that other people than the subject of a mental state *in fact* cannot know what mental state he or she is undergoing, which is consistent with the state being non-phenomenal. For example, it is true that other people *in fact* cannot know what I intend to vote

---

<sup>25</sup> Gottlieb and Lombrozo do not claim this themselves, as they do not mention phenomenal consciousness but only say that their aim with the study was to document people's intuitions about the limits of science.

for in the next election, at least as long as I do not tell them, and I know what I intend to vote for by what Gottlieb and Lombrozo call ‘introspection’. But this does not entail that my intention is phenomenally conscious, at least Chalmers offers no reason to believe this. So one may just as well interpret the subjects of Gottlieb and Lombrozo’s study as thinking that mental phenomena of which other people than the subject *in fact* cannot know, and of which the subject knows through introspection, are difficult to explain.

Diaz (2021: 65) asked subjects whether the properties of pain are fully explained in terms of neural activity and whether the feeling of pain is identical with neural activity. He found that some lay people do not think that neural activity fully explains the properties of pain and that the feeling of pain is not identical with neural activity. Diaz thinks this shows that these lay people have ‘problem-intuitions’ about phenomenal consciousness (2021: 67). But this is not clear. For it is unclear that subjects understood Diaz’ question as being about pain in the phenomenal sense rather than pain in the locatable sense, i.e. whatever ‘pain’ refers to in sentences like ‘There is a pain in my leg’. If subjects understood Diaz’ question as being about pain in the locatable sense, then the belief that the properties of pain are not fully explained by neural activity can be a result of lay people’s knowledge that the explanation of pain (in the locatable sense) may involve other things than neural activity, such as the nature and location of noxious stimulus, the mood or emotional state of the patient that depend on events in daily life, the general well-being and life situation of the patient, the patient’s (lack of) attention to the pain, the patient’s expectations about the painful effects of a stimulus, and so on. It is very implausible that neural activity alone explains the properties of pain in this sense (Corns 2020), and it is plausible that lay people are aware of explanatory factors other than neural activity. So their belief that neural activity does not fully explain the properties of pain may have nothing to do with phenomenality. Regarding the belief that the feeling of pain is not identical with neural activity, it is not clear why that belief should commit lay people to belief in phenomenality. True, they arguably believe that the feeling of pain is non-physical, but this may be because they believe that all mental states – not just the states taken to be phenomenal – are non-physical.

As discussed in chapter 2, Gregory et al. (submitted) asked subjects whether the colour-blind super-scientist Mary learns what it is like to see red after having a colour sight operation and then seeing a red tomato – closely modelled on Jackson’s (1982) original story about Mary’s release from the black and white room. They found that the majority of subjects tended to give positive responses, which Gregory et al. explain by lay people having similar intuitions about phenomenal consciousness as those philosophers typically have in connection with

Jackson's story about Mary (submitted: 3-4). But even though Gregory and colleagues intended to use the 'what-it's-like' phrase in the sense relevant to phenomenal consciousness, it is not clear that subjects understood the phrase in this way, as none of the comments subjects gave to elaborate on the answers they had given rule out a non-phenomenal interpretation.

Many of these comments are difficult to interpret. But some comments can plausibly be interpreted in a way suggesting that subjects understood the 'what-it's-like' phrase in the evaluative-descriptive sense, e.g. 'She only knows the physical information, not the mental or emotional response to seeing a red tomato'. Other comments can plausibly be interpreted in a way suggesting that subjects understood the 'what-it's-like' phrase in the non-evaluative descriptive sense restricted to the non-phenomenal 'ways' it is to see red, e.g. 'Just because she has an understanding of colour doesn't mean that she could identify one colour over another'. Yet other comments can plausibly be interpreted in a way suggesting that the subjects understood the 'what-it's-like' phrase in the resemblance sense restricted to non-phenomenal resemblance, e.g. 'She has never seen a color like red or yellow or green. She cannot know what red looks like yet'. And finally, as argued in detail in chapter 2, some comments can plausibly be interpreted in a way suggesting that the subjects understood 'knowing what-it's-like' in the ability sense, e.g. 'It is impossible to imagine a color you have never seen'.<sup>26</sup> Given that none of the subjects' comments rule out a non-phenomenal interpretation, Gregory et al.'s study does not show that lay people believe in phenomenality.

In all these studies, the problem is that the authors did not control for plausible alternative hypotheses, with the result that the studies do not show that lay people believe in phenomenality.<sup>27</sup> Of course, this does not entail that lay people do *not* believe in phenomenality. In order to show that one would need to provide compelling evidence, which I have not provided.<sup>28</sup> My claim is not that lay people do not believe in phenomenality but that we are currently not in a position to say that common sense counts in favour of either realism or eliminativism, since whether lay people believe in phenomenality is an open empirical question.

---

<sup>26</sup> Again, I am grateful to Daniel Gregory and colleagues for sharing the data from which these comments were taken.

<sup>27</sup> There may of course be other problems as well, such as the nature and size of the group of subjects in the studies. For example, Peressini's study was conducted with a group of only 73 subjects, all taken from a university logic class (2014: 868), which is arguably a small number and demographically speaking may not represent the population as a whole. But for the sake of charity, I have assumed that we can grant the validity of the subjects of the studies. The point is that even if we grant this, the studies still do not show that lay people believe in phenomenality.

<sup>28</sup> Some experimental philosophers believe there is compelling evidence that lay people do not believe in phenomenality (Sytsma and Machery 2010, Sytsma and Ozdemir 2019). But their results are ambiguous, so I think there is neither compelling evidence that lay people believe in phenomenality nor compelling evidence that they do not believe in phenomenality.

One possible reaction to this is that even though there may be no *empirical studies* showing that lay people believe in phenomenality, there may be sufficient *anecdotal evidence* that philosophers have gathered from talking to e.g. friends and family members. But exactly the same problem arises here: unless one controls for alternative hypotheses, the utterances of friends and family members is no compelling evidence for belief in phenomenality. For example, if I ask my cousin whether we can ever be sure that he sees red the same way as I and he answers ‘no’, then this does not show that he believes in phenomenality unless I have managed to rule out equally plausible hypotheses, such as the hypothesis that he believes that it is practically impossible to be sure that the physical/functional properties involved in seeing red are identical for him and me. I have not come across sufficient anecdotal evidence that controls for alternative hypotheses, so I shall not discuss anecdotal evidence further.

Another possible reaction is this: if the eliminativist explains away all apparent evidence for lay belief in phenomenality, then claims about lay belief in phenomenality look unfalsifiable or unverifiable. But this reaction is unfair, since I only claimed that authors of the above-mentioned studies failed to control for alternative hypothesis, I did not claim that hypotheses about lay belief in phenomenality are unfalsifiable or unverifiable. Indeed, if one had controlled for the mentioned alternative hypotheses, then one would have compelling empirical evidence that lay people believe in phenomenality. So more empirical research is called for to determine whether lay people believe in phenomenality, and importantly, research that *does* control for the alternative hypotheses discussed above. My claim is that we *currently* do not have any compelling evidence, since all the relevant studies failed to control for plausible alternative hypotheses.

## **5. Conclusion**

Eliminativism is the view that phenomenal consciousness does not exist. This view is consistent with there being something it is like to be in mental states in the ordinary senses of the ‘what-it’s-like’ phrase, with the existence of feelings and experiences in what may be the ordinary (i.e. functional) senses of these words, and what we know to be common sense. It also has advantages vis-à-vis both non-reductive and reductive realism. So contrary to what is sometimes assumed, it is a view worth taking seriously. In the next chapter, I address a problem for eliminativism – the problem of explaining why people have intuitions, beliefs and make judgements about phenomenality.



## Chapter 4

### Why Do People Believe in Phenomenality?

**Abstract:** According to eliminativism about phenomenality, phenomenality does not exist. But why, according to the eliminativist, do people have intuitions, beliefs and make judgements about phenomenality? According to Darwinism (Humphrey 2011), illusionism (Frankish 2016a, 2016b, Graziano 2016, 2019, Kammerer 2018, 2021), and inferentialism (Rey 1995, Clark et al. 2019, Schwarz 2019, Shabasson 2021), the belief in phenomenality is a product of biological or innate mechanisms. According to educationism (Balmer 2020), the belief in phenomenality is a product of cultural mechanisms. This chapter surveys these recently proposed theories and suggests several improvements as well as points out where more work needs to be done. I conclude that we are currently not in a position to tell which of these theories is the most promising eliminativist-friendly explanation of belief in phenomenality, but that the potential of some of these theories confirms the conclusion of the previous chapter, namely that eliminativism is a view worth taking seriously.

#### 1. Introduction

In the previous chapter, I argued that eliminativism about phenomenality is a view worth taking seriously. But if this is to be convincing, eliminativists must explain why people have intuitions, beliefs, and make judgements about phenomenality, given that realists can explain this by appealing to the existence of phenomenality. This is arguably the major problem for eliminativism.<sup>29</sup>

---

<sup>29</sup> Some philosophers claim that a major problem for eliminativism is the ‘resistance problem’ (sometimes called the ‘illusion meta-problem’), which is the problem of explaining why people are resistant to eliminativism (Chalmers 2018, Pereboom 2019, Kammerer 2021). But it is not clear who the ‘people’ resistant to eliminativism are. If one means that *philosophers and scientists* who believe in phenomenality are resistant to eliminativism, then their resistance can be explained by the cognitive bias that explains all theorists’ resistance to their opponents’ view, including eliminativists’ resistance to realism. And if one means that lay people unfamiliar with philosophy are resistant to eliminativism, then their resistance can be explained by their not being informed about the theoretical virtues of eliminativism. But if lay people are resistant to eliminativism and *not* because they are uninformed about the theoretical virtues of eliminativism, then one would need an alternative explanation. To this end, Kammerer (2021) provides a psychological explanation, according to which people are predisposed to judge that their introspective appearances are veridical. According to Kammerer then, if it is true that people have introspective appearances of phenomenality and they are predisposed to judge that their introspective appearances are veridical, then they will believe in realism and thus be resistant to eliminativism. Regardless of who the people resistant to eliminativism are then, the resistance problem seems quite tractable, so I shall only focus on the problem of explaining why people believe in phenomenality in what follows.

I argued in the previous chapter that we are currently not in a position to tell how widespread belief in phenomenality is. It is clear enough that many philosophers and scientists believe in it, but it is not clear to what extent lay people believe in any such thing. Before we have more empirical knowledge about how widespread belief in phenomenality is then, it would be nice to consider both explanations that are consistent with belief being quite widespread and explanations consistent with belief being quite limited. Darwinism (Humphrey 2011), illusionism (Frankish 2016a, 2016b, Graziano 2016, 2019, Kammerer 2018, 2021), and inferentialism (Rey 1995, Clark et al. 2019, Schwarz 2019, Shabasson 2021) claim that the belief in phenomenality is a product of biological or innate mechanisms. So these theories seem to predict that belief in phenomenality is widespread in the population, given that the relevant biological/innate mechanisms are widespread. By contrast, educationism (Balmer 2020) claims that the belief in phenomenality is a product of cultural mechanisms. These mechanisms can both be limited to a theoretical community and be widespread in the population, so educationism can explain belief in phenomenality even if this is limited to philosophers and scientists.

This chapter surveys these theories and suggests several improvements as well as points out where more work needs to be done. I conclude that our lack of knowledge about how widespread belief in phenomenality is, prevents us from knowing which of these theories is the most promising, but that the potential of (at least some of) these theories confirms the conclusion of the previous chapter, namely that eliminativism is a view worth taking seriously. For simplicity I focus on the explanation of belief, but the discussion straightforwardly extends to the explanation of intuitions or judgements as well.

§2-5 discuss Darwinism, illusionism, inferentialism and educationism. §6 concludes.

## **2. Darwinism**

Darwinism is the first theory appealing biological/innate mechanisms, and therefore assumes that belief in phenomenality is widespread in the population (across culturally isolated groups).

2.1. outlines the theory, and 2.2. outlines the main challenge for the theory.

### *2.1. The Theory*

According Humphrey's Darwinism, people believe in phenomenality because this belief is conducive for survival. In Humphrey's (2011) words, belief in phenomenality 'makes life more worth living' (2011: 75), creates a 'will to live' (2011: 86), and gives one a sense of

‘metaphysical importance’ (2011: 75). This, according to Humphrey, increases one’s investment in one’s own survival (2011: 75). Humphrey does not develop these claims in much detail, but the general idea is that just like the fact that the human brain is conducive for survival explains why creatures have evolved with the human brain, so the idea that belief in phenomenality is conducive for survival is supposed to explain why creatures have evolved with belief in phenomenality.

## 2.2. *Challenges with Darwinism*

As they stand, Humphrey’s reasons for claiming that belief in phenomenality has evolutionary advantages are underwhelming. Assuming that finding life more worth living and having a sense of ‘metaphysical importance’ increases one’s investment in one’s own survival, why should one believe that the belief in phenomenality makes life more worth living and oneself feel more important? There seems to be no reason to believe that eliminativists take their lives to be less worth living and themselves to be less important than realists do, since eliminativists do not behave in a way suggesting they are less encouraged to invest in their own survival than realists. And if belief in phenomenality has no impact on survival today, then it is not clear why the survival of our ancestors should have been impacted in the way suggested by Humphrey. Perhaps it matters that they were living in a primitive society and not a modern one, but Humphrey does not explain what difference that makes, and I am not sure how to make this difference matter in a way that removes the impression that Darwinism is an empirically implausible theory. In its current form then, Humphrey’s Darwinism is too vague and speculative, and much more needs to be said for it to be a compelling theory about why people believe in phenomenality.

## 3. **Illusionism**

There is an ambiguity in the word ‘illusion’. One can understand ‘illusion’ as referring to an inaccurate *cognitive* state, e.g. an intuition, belief or judgement. But one can also understand ‘illusion’ as referring to an inaccurate *non-cognitive* state, e.g. a perceptual or quasi-perceptual representation, on which cognitive states (intuitions, beliefs, judgements) are based. For this reason, there is a similar ambiguity in ‘illusionism’. Frankish says that ‘Illusionism makes a very strong claim: it claims that phenomenal consciousness is illusory; experiences do not really have qualitative, ‘what-it’s-like’ properties, whether physical or non-physical’ (2016: 15). On

this use of the term, ‘illusionism’ is simply a name for eliminativism about phenomenality. But there is also a different use of the term. Here is Kammerer:

Eliminativists can be distinguished by the way in which they propose to account for the fact that most of us falsely believe they are phenomenally conscious. Some eliminativists interpret this false belief as the result of a theoretical error—a kind of *doxastic mistake* [...] Alternatively, some eliminativists insist that phenomenal consciousness is not merely a theoretical posit; it is something which persistently and robustly *seems* to exist. In this kind of view, the belief in phenomenal consciousness is not the result of a doxastic mistake; it rather stems from a kind of *introspective illusion*. We can call this view “illusionism” (2021: 847-848).

On Kammerer’s use of the term, ‘illusionism’ is associated with a view about why people believe in phenomenality. In what follows, I use ‘illusionism’ exclusively for a view about why people believe in phenomenality and not for eliminativism about phenomenality. 3.1. explains the theory, and 3.2. outlines two challenges.

### 3.1. *The Theory*

According to *illusionism*, the belief in phenomenality is based on a separate intermediate representation – an ‘illusion’ in the second sense of the term outlined above (Frankish 2016a, 2016b, 2019, Graziano 2016, Kammerer 2018, 2021). This ‘illusion’ cannot itself be a cognitive state like a belief, since otherwise illusionism would leave the problem of explaining the belief in phenomenality unexplained. Thus, Frankish says that the illusion of phenomenality refers to ‘*quasi-perceptual* introspective appearances’ and not just ‘the *cognitive* illusion involved in judging that we are acquainted with [...] phenomenal properties’ (2016b: 11, 16).<sup>30</sup>

Other illusionists are less explicit about the nature of the alleged ‘illusion’ of phenomenality but say things which suggest a view similar to that of Frankish. For example, Kammerer says that ‘the belief in phenomenal consciousness is not the result of a doxastic mistake: it rather stems from a kind of *introspective illusion*’ (2021: 848) and that ‘illusionism is not simply the thesis according to which we *judge* phenomenal states to exist while they don’t, it’s the thesis according to which phenomenal states *appear* to exist while they don’t’ (2018: 58). Here Kammerer makes a contrast between (a) making a doxastic mistake or judging

---

<sup>30</sup> This use of ‘illusion’ is different from the use of ‘illusion’ that grounds Frankish’s understanding of ‘illusionism’. This is confusing, which is why I have stressed the ambiguity.

wrongly that phenomenality exists and (b) being subject of an introspective illusion that phenomenality exists or having phenomenality appear to exist while it does not. One natural interpretation of ‘introspective illusion’ and ‘appear’ is as involving a cognitive state like belief, but on this interpretation, there would be no contrast between (a) and (b). Thus, what Kammerer means by ‘introspective illusion’ and ‘appear’ must be a non-cognitive state, such as Frankish’s quasi-perceptual representation. In what follows, I use this cognitive vs. non-cognitive distinction to separate beliefs, intuitions and judgements on the one hand, and representations on which these are based, such as perceptual or quasi-perceptual representations, on the other. Illusionists do not use these labels themselves, but it clarifies their position, which is often formulated in unnecessarily ambiguous terms (such as Kammerer’s ‘introspective illusion’ and ‘appear’).

### 3.2. *Challenges with Illusionism*

The first challenge for illusionism is articulating what a non-cognitive ‘illusion’ of phenomenality is supposed to be. We are familiar with non-cognitive or quasi-perceptual representations of things with familiar properties of external physical objects like size, shape, colour, texture, sound, smell, etc., i.e. hallucinations. This also applies to quasi-perceptual representations (or hallucinations) of things that do not exist, like unicorns or ghosts. But the non-cognitive or quasi-perceptual representation of phenomenality is not supposed to represent phenomenality as having any such properties of external physical objects, since phenomenality is supposed to be a property of mental states and mental states do not have size, shape, colour, texture, sound, smell etc. Given this, it is not clear what a non-cognitive or quasi-perceptual representation of phenomenality is supposed to be. To put the point somewhat tendentiously: we know what it is to ‘see’ unicorns or ghosts – it is having a visual representation of a horse-like creature with a horn, or of a human-like creature that is transparent and walks through walls. But it is not clear what it is to ‘see’ phenomenal properties (what do they look like?).

Even though illusionists are not explicit about this, perhaps they think that the non-cognitive ‘illusion’ of phenomenality should be understood in terms of an inner sense theory of introspection, according to which introspection is a ‘self-scanning process in the brain’ (Armstrong 1968: 324). On this view, introspective representations are quasi-perceptual not in the sense that they represent objects with familiar properties of physical objects that we know through the ‘external’ senses. Rather, introspective representations are quasi-perceptual in the sense that the connection between the object of introspection and the introspective representation is causal, and that introspection involves attention and is in some sense non-

inferential. Given that the connection between the object of introspection and the introspective representation is causal – which is a contingent connection – the inner sense theory allows for introspective misrepresentation or ‘illusion’, similar to perceptual misrepresentation or illusion. This would explain what illusionists mean by a non-cognitive or quasi-perceptual ‘illusion’, which solves the first challenge for illusionism.<sup>31</sup>

The second challenge for illusionism is the illusion problem: the problem of explaining why people have the non-cognitive or quasi-perceptual ‘illusion’ of phenomenality, regardless of how we are to understand this alleged ‘illusion’ (cf. Chalmers 2018: 56, Frankish 2016a: 12, 37). We have such explanations for other illusions. For example, people have the visual illusion of a bent stick when a straight stick is put into water because of various reflectance and optical properties. But it is not immediately clear why people have a non-cognitive or quasi-perceptual ‘illusion’ of phenomenality.

To my knowledge, there is only one attempt to solve the illusion problem in the literature, namely that of Graziano (2016, 2019). According to Graziano’s (2016, 2019) ‘attention schema’ theory, the brain constructs internal models of worldly objects and processes, one of which is the brain’s own process of attention – in the sense of the deep processing of select signals at the expense of others (2016: 101). In order not to waste processing resources on pragmatically unimportant aspects of the model target, internal models are not accurate descriptions, but incomplete and schematic. Being incomplete and schematic, the model of attention lacks details about the physical properties of attention, which, according to Graziano, leads to the belief that there is a thing inside us that lacks physical properties. This theory might explain why people have a non-cognitive or quasi-perceptual ‘illusion’ of phenomenality, since the internal model of attention may be ‘quasi-perceptual’ in the sense of the inner sense theory of introspection, and attention (according to Graziano) correlates with states that realists take to be phenomenal.

However, Graziano’s theory faces the objection that it seems to predict that we have many other ‘illusions’, which do not cause corresponding intuitions or beliefs. So even if this theory explains the ‘illusion’ of phenomenality, it is not clear that the ‘illusion’ of phenomenality that Graziano posits can explain why people believe in phenomenality, which was the reason to posit the ‘illusion’ in the first place. Consider Graziano’s example of the internal model of an apple, which (according to Graziano) lacks microscopic physical details

---

<sup>31</sup> Though it is worth noting that the inner sense theory is not uncontested. See Shoemaker (1994) for an argument against the theory.

about the apple (2016: 102). One such detail is presumably that the apple contains H<sub>2</sub>O molecules. But the internal ‘illusion’ of the apple does not cause the corresponding belief that the apple lacks H<sub>2</sub>O molecules, so it is not clear why the illusion of attention causes a belief in something non-physical (phenomenality), and the illusion of the apple does not cause the belief that the apple lacks H<sub>2</sub>O molecules. As Chalmers (2020a) points out, if the theory is supposed to explain belief in phenomenality, it requires that people make a fallacy about attention that they do not make about other model targets. They fallaciously move from the model failing to represent attention as physical to representing attention as non-physical, analogous to the fallacy of moving from not representing an object as heavy to representing the object as light. But Graziano’s theory does not say why one should believe that people make this fallacy in the case of attention (2020a: 212-213). So it merely replaces the illusion problem with the problem of explaining why people make this fallacy. Consequently, illusionists have not succeeded in solving the illusion problem.

#### **4. Inferentialism**

Inferentialism is the final theory that appeals to biological/innate mechanisms. 4.1. outlines the theory, and 4.2. outlines some challenges.

##### *4.1. The Theory*

According to inferentialism, people believe in phenomenality because of mistaken unconscious inferences, i.e. mistaken inferences of which the subject lacks a higher-order awareness (Rey 1995, Clark et al. 2019, Schwarz 2019, Shabasson 2021). The unconscious inferences again are based on innate features of the human mind or brain. Shabasson (2021) seems to offer the most developed version, so I shall focus on that. The theory is very detailed and complex, so I focus on what I understand are the essential parts.

According to Shabasson, an introspecting subject makes an introspective judgement like the following:

- (1) I am having a mental state belonging to a certain type, such as seeing red, feeling pain, or drinking coffee.

Introspection then activates unconscious inferences like the following:

- (2) There is a property that justifies (1).
- (3) The property that justifies (1) can neither be a physical/neural property of the mental state I am having, nor an external property of a red object, of a body part, or of a sip of coffee, so it must be an internal mental property.
- (4) The internal mental property is a phenomenal property.

The reason to believe that people unconsciously infer (2) when introspecting is what Shabasson calls the ‘justification constraint’: ‘A subject judging that  $p$  necessarily believes, either explicitly or tacitly, that she is justified in judging that  $p$  at the time she judges that  $p$ ’ (2021: 7). Shabasson cites two reasons to believe that people unconsciously infer (3). The first is the idea that we do not have introspective access to the physical/neural properties of mental states, i.e. what Shabasson calls ‘introspective opacity’ (2021: 10). The second reason to believe that people unconsciously infer (3) is that upon philosophical reflection, one will deny that the property justifying (1) is an external property, since (according to Shabasson) considerations like the inverted spectrum shows that what justifies an introspective judgement like (1) is not an external property but an internal mental property (2021: 11). Shabasson thinks that the justificatory mental properties posited by (2)-(3) just are what people conceive of as phenomenal properties, hence (4).

Before I move on to the challenges with Shabasson’s inferentialism, there is a point that is worth comment. Shabasson portrays his inferentialism as offering a solution to the ‘illusion problem’. But we do not know how inferences alone can give rise to non-cognitive representations, we only know how inferences can give rise to cognitive representations. So Shabasson’s theory is arguably no solution to the illusion problem discussed above, since that is the problem of explaining why people have non-cognitive representations of phenomenality. The reason Shabasson calls inferentialism a solution to the illusion problem is, again, the ambiguity in ‘illusion’. On Shabasson’s use of the term, ‘illusion’ refers to false belief, ‘illusionism’ refers to eliminativism about phenomenality, and ‘illusion problem’ refers to the problem of explaining why people (falsely) believe in phenomenality. Even though Shabasson seems to think that inferentialism is concerned with the illusion problem that Frankish and Kammerer are concerned with, it is not. The reason is that what the latter authors mean by ‘illusion’ is not a false belief but an inaccurate non-cognitive representation, and what they mean by ‘illusion problem’ is not the problem of explaining why people (falsely) believe in phenomenality but the problem of explaining why people have non-cognitive representations of phenomenality on which (according to Frankish and Kammerer) the (false) belief in



phenomenality is based. This is why I treat inferentialism as a separate theory about why people believe in phenomenality, unrelated to illusionism and the illusion problem in my use of these words.

#### 4.2. *Challenges with Inferentialism*

Even if we grant that Shabasson's inference theory explains why people believe in phenomenality, the reasons he offers to believe in unconscious inferences need further work to be compelling.

As noted, in support of (2), Shabasson cites the 'justification constraint': 'A subject judging that  $p$  necessarily believes, either explicitly or tacitly, that she is justified in judging that  $p$  at the time she judges that  $p$ ' (2021: 7). Shabasson calls this a 'conceptual truth' (2021: 8), but it is not clear why one should believe that the justification constraint is true, since it seems possible to have beliefs for which one thinks that one lacks justification. Religious belief may be one example, and the belief that one is not a brain-in-vat may be another. It seems possible to have these beliefs without also believing that they are justified. Of course, one could just weaken the justification constraint and say that even though people do not always think that their judgements or beliefs are justified, they innately believe that their introspective judgements of the sort similar to (1) are justified. This is an interesting hypothesis, but more needs to be said about why one should believe that this weakened justification constraint is true.

In support of (3), Shabasson cites the fact that people do not have introspective access to the neural properties of mental states (i.e. what he calls 'introspective opacity'). But the fact that people do not have introspective access to the neural properties of mental states is not enough to predict that people do not think the property justifying (1) is a neural property. The reason is that one does not need to have epistemic access to what one thinks justifies one's belief, unless one imposes some internalist requirement (which Shabasson does not mention) on justification saying that one must have epistemic access. The hypothesis that we have an internalist requirement as an innate principle of our mind or brain is an interesting one, but just like with the weakened justification constraint, more needs to be said about why one should believe it is true. Thus, inferentialism needs more work to be a compelling eliminativism-friendly theory about why people believe in phenomenality.

## 5. Educationism

According to educationism, the belief in phenomenality is based on theoretical ‘education’, i.e. exposure to theoretical ideas. The theory was introduced by Balmer (2020), who claims that ‘an individual would only learn to make phenomenal judgements by picking up ideas from their surrounding culture’ (2020: 27).<sup>32</sup> However, Balmer does not say much more about the details of the theory, so I shall say a little more on how this kind of view could be developed.

The educationist idea is that people’s belief that phenomenality exists is the result of their picking up the idea in conversation, in written text, in a lecture, or whatever phenomenon that may provide one with theoretical ‘education’. This broad notion of education is not limited to university level teaching but includes all human communication about phenomenality. Unlike Darwinism, illusionism, and inferentialism then, educationism does not explain the belief in phenomenality by a biological or innate mechanism. Rather, the belief in phenomenality is explained by a cultural mechanism. So on this view, belief in phenomenality is not an integral part of being human, it may be limited to a fairly small theoretical community, such as analytic philosophy and some parts of cognitive science. But the view does not entail that belief is limited to a small theoretical community, as a cultural mechanism may be responsible for belief in lay people as well as in philosophers and cognitive scientists.

There are two versions of educationists, and which version one defends depends on whether one is targeting the beliefs of philosophers only or the beliefs of lay people as well as philosophers. We can call the version of educationism that only targets the beliefs of philosophers (and some cognitive scientists who engage with philosophy) *narrow* educationism, and the version of educationism that targets the beliefs of lay people as well as philosophers *wide* educationism. 5.1. discusses narrow educationism, 5.2. discusses wide educationism, and 5.3. discusses a challenge.

### 5.1. *Narrow Educationism*

One way of developing narrow educationism is by appealing to (a) students of philosophy’s desire to belong to those philosophers count as normal, and (b) philosophers’ rhetoric suggesting that only realists are normal. In combination, (a) and (b) explain why students of philosophy come to believe in phenomenality.

The idea that the desire to belong to a certain community is a source of cognitive bias has been discussed under various labels. For example, Kahan (2017) discusses the phenomenon

---

<sup>32</sup> Balmer (2020) calls the theory ‘soft-wired illusionism’, since he uses ‘illusionism’ as a label for eliminativism.

of ‘identity-protective cognition’, i.e. when one’s belief is motivated (often unconsciously) by a desire to protect one’s status as member of a certain group, rather than by evidence for its truth. Related to this, is what Funkhauser (2017) calls ‘signalling’, i.e. when one’s belief is motivated (often unconsciously) by a desire to ‘signal’ specific information about the believer to other people, which may be information about her group membership. Insofar as students of philosophy want to belong to those philosophers count as normal, this desire may be a source of belief in phenomenality *if* students of philosophy are also told that philosophers only count realists as normal.

The defender of narrow educationism must therefore provide evidence that students of philosophy are told things suggesting that philosophers only count realists as normal. Here a defender of narrow educationism can appeal to the rhetoric characteristic of philosophical ‘education’ about phenomenality. The rhetoric I have in mind is manifested in at least two ways.

One manifestation of rhetoric suggesting that philosophers only count realists as normal is bold dismissals of eliminativism. Take for example these claims quoted in the previous chapter:

‘[I]t is beyond dispute that there are such things as qualia’ (Carruthers 2000: 15).

‘To take the line that explaining our judgements about consciousness is enough [...] denies the evidence of our own experience. This is the sort of thing that can only be done by a philosopher’ (Chalmers 1996: 188).

‘[T]he qualiaphobe’s denial of conscious experience seems ludicrous’ (Levine 2001: 131).

‘According to illusionism about phenomenal consciousness, no living being is ever phenomenally conscious. This claim strikes many people as absurd and that reaction is perfectly adequate. There is so much to do in philosophy in search of the truth that one should not lose too much time with absurd theories’ (Nida-Rümelin 2016: 160).

‘There occurred in the twentieth century the most remarkable episode in the whole history of ideas – the whole history of human thought. A number of thinkers denied the existence of something we know with certainty to exist: consciousness, conscious experience’ (Strawson 2017, quoted in Chalmers 2018: 55).

I take these locutions to be representative of a confident rhetoric that seems quite unique within professional philosophy. Elsewhere in philosophy, philosophers typically express their views with hedging qualifications and thorough argument, but eliminativism is often dismissed

merely on the ground that the author is of the opinion that eliminativism is ridiculous or absurd. These bold dismissals of eliminativism suggest that philosophers only count realists as normal. Given that students of philosophy have the desire to be among those philosophers count as normal, bold dismissals of eliminativism may inculcate the belief in phenomenality.

A second manifestation of rhetoric suggesting that philosophers only count realists as normal is talk expressing intuitions. For example, claims about it being ‘intuitively plausible’ that Mary learns what it is like to see red (in the sense relevant to phenomenal consciousness), or that a zombie world ‘seems conceivable’ may have the same belief-inducing effect as bold dismissals of eliminativism. This is because claims about what is ‘intuitively plausible’ or what ‘seems conceivable’ is naturally interpreted as being about what *most people* – at least most of the people who have considered the relevant idea – find intuitively plausible and what seems conceivable to most people, rather than what a *single author* finds intuitively plausible or what seems conceivable to a single author. Thus, such claims expressing intuitions suggest that anyone denying them is abnormal and hence that philosophers only count realists as normal, given that the relevant intuitions are most naturally associated with, and accommodated by, realism. Given also that students of philosophy have the desire to be among those philosophers count as normal, claims expressing intuitions may inculcate the belief in phenomenality.<sup>33</sup>

We now have the two factors that together explain why students of philosophy come to believe in phenomenality: their desire to belong to those philosophers count as ‘normal’, and philosophical rhetoric about phenomenality suggesting that philosophers only count realists as normal. This explanation of philosophers’ belief in phenomenality is similar to the explanation of certain political beliefs that appeals to people’s desire to belong to those the political community considers as ‘normal’ and propaganda. Indeed, without the desire to belong to a certain community, a lot of propaganda – even that aimed at informed and competent people – would arguably never have worked. But there are plenty of historical examples where propaganda has led informed and competent people to believe in all sorts of things, despite it not being supported by evidence. So it is possible that the desire to belong to those philosophers count as ‘normal’, in combination with rhetoric suggesting that philosophers only count realists as normal, is the source of philosophers’ belief in phenomenality. Of course, it might be

---

<sup>33</sup> Perhaps talk expressing intuitions need not involve words specifically indicating intuition, such as ‘intuitively plausible’ and ‘seems’. One might express one’s intuition that Mary learns what it is like to see red (in the sense relevant to phenomenal consciousness) simply by saying that ‘Mary learns what it is like to see red’. If so, this claim might have the same belief-inducing effect as the same claim preceded by ‘It is intuitively plausible that...’. The defender of narrow educationism is not claiming that only talk involving locutions like ‘intuitively plausible’ inculcate the belief in phenomenality, but that any locution that (a) expresses intuition, and (b) suggests that the intuition is shared by philosophers, inculcates the belief in phenomenality.

surprising that the belief in phenomenality is the result of cognitive bias. But we know that cognitive biases are ubiquitous, so it would be no less surprising if philosophers escaped them.

The idea then is that a student of philosophy comes to the philosophy classroom without any concept of, or belief in, phenomenal consciousness, but with a desire to belong to those counted as ‘normal’. She then learns about phenomenality via standard thought experiments, such as the ‘knowledge argument’ and that of a ‘zombie’ world. This teaches her a new concept (of phenomenality), new senses of expressions she knows from before (such as the ‘what-it’s-like’-phrase, ‘consciousness’, ‘feeling’ and ‘experience’<sup>34</sup>), and expressions she does not know from before (such as ‘qualia’, ‘phenomenality’ and ‘raw feel’). However, her philosophical education about phenomenality does not only involve thought experiments and the use of technical language. It also involves rhetoric suggesting that philosophers only count realists as normal. Thus, it is not only the concept of phenomenality which is inculcated by her philosophical education, it is also (given her desire to be counted as normal) the belief that the concept refers.

According to this story, both the concept of, and belief in, phenomenality, are inculcated by philosophical education. But they are inculcated by different aspects of philosophical education, or mechanisms that are distinct, at least in theory. In practice, the mechanism that inculcates the concept of phenomenality (i.e. the use of thought experiments and technical language) likely converges with that which inculcates the belief (i.e. the use of rhetoric suggesting that philosophers only count realists as normal, given the desire to belong to those counted as normal), such that the inculcation of both the concept and the belief happen simultaneously. But in theory, these mechanisms may be isolated, such that one can have the concept without the belief. For example, if the philosophy teacher instead of saying that ‘It is intuitively plausible that Mary learns what it is like to see red upon release’ rather said that ‘Philosophers have stipulated a new, technical sense of the ‘what-it’s-like’ phrase, which is different to any of the ordinary senses, and some might claim that Mary learns what it is like to see red upon release in this technical sense’ – then this use of the thought experiment and technical language may inculcate the concept of phenomenality. But in lack of any rhetoric suggesting that non-realists are abnormal, narrow educationism predicts that the mere use of thought experiments and technical language will not inculcate belief in phenomenality. The same prediction would be made for individuals who do not have the desire to be counted as

---

<sup>34</sup> We are now assuming that only philosophers believe in phenomenality, in which case consciousness, feelings and experiences – in the *ordinary* senses of these terms – do not concern phenomenality.

‘normal’. So even though the mechanism by which the concept of phenomenality is inculcated in students of philosophy *in practice* may converge with the mechanism by which the belief in phenomenality is inculcated, the mechanisms are distinct mechanisms and need not converge.

### 5.2. *Wide Educationism*

Narrow educationism assumes that belief in phenomenality is limited to philosophers. If, by contrast, belief in phenomenality is widespread in the population and one is trying to explain the beliefs of lay people as well as philosophers, then peculiarities with philosophers and their education will not explain the target phenomenon. In a scenario where belief in phenomenality is widespread in the population, educationists will defend *wide* educationism and claim that realism is a false folk theory passed over from generation to generation. Defenders of wide educationism make a claim similar to that of eliminativists about folk psychology as a whole who claim that folk psychology as a whole is a false theory passed over from generation to generation, or atheists who claim that religions are false theories passed over from generation to generation.

Wide educationism can be developed from the idea that we teach our children that we have feelings and experiences, which – assuming that belief in phenomenality is widespread in the population – plausibly concern phenomenality in the *ordinary* senses of these terms. When children (presumably already in kinder garden age) learn the concepts of feeling and experience then, they learn the concept of phenomenality. Their learning the concepts of feeling and experience plausibly involves adults applying these concepts, which also makes the children think that the concepts refer and thereby also induces belief in phenomenality. In this case then, the mechanism is the same for inculcation of both the concept of, and the belief in, phenomenality, namely adults’ use of the concepts of feeling and experience. This is unlike narrow educationism described above, where distinct mechanisms inculcate the concept and the belief respectively. For in the case of narrow educationism, the use of thought experiments and technical language inculcate the concept, while rhetoric suggesting that philosophers only count realists as ‘normal’ (given the desire to belong to those counted as ‘normal’) inculcates the belief.

### 5.3. *Challenges with Educationism*

Like the other theories discussed above, educationism also leaves something unexplained. Regardless of which version of the theory one has in mind, educationism does not explain the *origin* of the belief in phenomenality. Educationism explains the beliefs of those exposed to

theoretical education, but this presupposes ‘educators’ – i.e. philosophy teachers in the case of narrow educationism, and people responsible for bringing up children in the case of wide educationism – from which people pick up the idea of phenomenality and thereby the educators’ belief in phenomenality. In order to explain the origin of the belief in phenomenality, the educationist would have to explain why someone started to believe in phenomenality in the first place without having picked up the idea from anyone else. In other words, educationism needs to be supplemented with a separate hypothesis to explain why someone invented phenomenality, whenever that was. By contrast, Darwinism, illusionism and inferentialism explain the origin of the belief in phenomenality.

I shall not attempt to track the origin of phenomenality here, however. But it is plausible that this may be done differently for narrow and wide educationism. If only philosophers believe in phenomenality, then realism may be a relatively recent innovation, given that it is far from obvious that historical philosophers were concerned about phenomenality, in which case one might be able to identify a known philosopher and his or her theoretical motivations for introducing phenomenality. But if belief in phenomenality is widespread in the population, then realism plausibly has an ancient origin, given that people have believed in feelings and experiences for a long time, in which case it might be more difficult to identify any *one* theorist whose theoretical motivations led to the idea of phenomenality.

## **6. Conclusion**

In the previous chapter, I argued that, for all realists have shown, eliminativism about phenomenality is a view worth taking seriously. But if this is to be convincing, eliminativists must arguably offer an explanation for why people believe in phenomenality. To this end, Darwinism, illusionism and inferentialism appeal to biological or innate mechanisms. By contrast, educationism appeals to a cultural mechanism. All theories are recent proposals for which I have suggested several improvements and pointed out where more work needs to be done.

Darwinism is the least developed theory and much more needs to be said for it to be convincing that belief in phenomenality is conducive for survival. Illusionism can be developed in combination with the inner sense theory of introspection, but it faces the illusion problem – the problem of explaining why people have non-cognitive ‘illusions’ of phenomenality. Inferentialism can be developed with a weak justification constraint saying that people innately believe that their introspective judgements are justified, and with an innate internalist

requirement on justification. But more needs to be said in support of these hypotheses. Finally, educationism appeals to cultural rather than biological or innate mechanisms, but it faces the challenge of explaining why theorists invented phenomenality. In short, all theories take something for granted that is arguably in need of explanation.

Which theory is the most plausible? Darwinism seems to predict that eliminativists invest less in their own survival than realists, which makes Darwinism (as it stands) a less plausible theory than the others. And while it may be difficult to distinguish illusionism and inferentialism (once developed) by empirical predictions, educationism makes an interesting prediction that neither of the other theories make. Educationism predicts that people not under the influence of theoretical education about phenomenality will not believe in phenomenality. This prediction should make it fairly straightforward to determine whether the cultural approach of educationism or the biological/innate approach of Darwinism, illusionism and inferentialism is the most plausible approach to explain belief in phenomenality. So, for example, if we gain empirical knowledge that only philosophers and cognitive scientists believe in phenomenality, then the cultural approach of educationism seems more plausible than that of the other theories. But if we gain empirical knowledge that belief in phenomenality is widespread across culturally isolated groups, then the biological/innate approach of Darwinism, illusionism and inferentialism seems more plausible.

In practice, however, things are complicated, since as argued in the previous chapter, it is difficult to tell whether belief in phenomenality is widespread in the population or whether it is peculiar to philosophers and some cognitive scientists. Therefore, we are currently not in a position to tell whether the cultural approach or the biological/innate approach for explaining belief in phenomenality is the most promising. Nevertheless, the potential of at least some of these theories confirms the conclusion of the previous chapter, namely that, for all realists have shown, eliminativism is a view worth taking seriously.

This ends the discussion related to phenomenal consciousness. The following two chapters move to a different topic – pain – and importantly, pain in a sense that is independent of phenomenal consciousness.



## Chapter 5

### The Bodily Theory of Pain

**Abstract:** One use of the word ‘pain’ is exemplified in sentences like ‘There is a pain in my foot’. According to the Experiential Theory, ‘pain’ in this context refers to an experience located in the mind or brain. According to the Bodily Theory, it refers to an extra-cranial bodily occurrence located in a body part. In this chapter, I defend the Bodily Theory. Specifically, I argue that pains are proximal activations of nociceptors that cause experiences of pain. This view is preferable to the Experiential Theory, because it accords better with common sense and offers a better interpretation or semantics of ordinary pain reports, and because the arguments for the Experiential Theory are unconvincing.

#### 1. Introduction

The noun ‘pain’ is used in many different ways. One use of the word is exemplified in sentences like ‘There is a pain in my foot’. I call the referent of ‘pain’ in this use of the word ‘pain in the locatable sense’, because this is a use of the word that is naturally followed by a phrase specifying a bodily location like ‘in my foot’. When speaking of pain in the locatable sense one *need* not specify a location, but one *can*, which is why ‘locatable’ is suitable.

Some philosophers believe that pains in the locatable sense are (extra-cranial) bodily occurrences located in body parts (Newton 1989, Byrne 2001, Hyman 2003, Hacker 2013, Hill 2014, 2017, Massin 2017, Reuter and Sytsma 2020, Bradley 2021).<sup>35</sup> Others believe that pains in the locatable sense are experiences or intentional mental states, which are typically taken to be located in the mind or brain (Armstrong 1968, Tye 1995a, 1995b, 2002, 2005a, 2005b, 2017,

---

<sup>35</sup> I shall mostly drop the qualification ‘in the locatable sense’ hereafter, so when I speak of pain in this and the following chapters, I mean pain in that sense unless otherwise specified. (I discuss pain in other senses of the word in §2.) I shall also drop the qualification ‘extra-cranial’, so whenever I speak of bodily occurrences, I mean occurrences in parts of the body other than the brain.

Hall 2008, Klein 2015). Call these views the Bodily Theory and the Experiential Theory respectively.<sup>36 37</sup>

This chapter is about what pains in the locatable sense are, i.e. what ‘pain’ refers to in sentences like ‘There is a pain in my foot’. Because experiences are taken to exist in the brain, and bodily occurrences exist almost anywhere in the body but the brain, what one takes to be the answer to this metaphysical question has consequences for where one takes pains to be located, and a theory of what pains are is simultaneously a theory of where they are. In what follows, I defend the Bodily Theory on the grounds that it accords better with common sense and offers a better interpretation or semantics of ordinary pain reports.

Specifically, my claim is that pains are proximal activations of nociceptors that cause experiences of pain. The claim is that actual human pains (in this world) are activations of nociceptors, and not that pains are necessarily (in all possible worlds) activations of nociceptors, as the latter claim would conflict with the possibility of multiple realization, i.e. the implementation of pain in something physically different from our bodies.

Even though there are defenders of the Bodily Theory in the literature, the majority do not attempt to develop this view in great detail.<sup>38</sup> This chapter aims to remedy that and develop the Bodily Theory as a compelling theory of pain. My contribution is threefold. Firstly, I develop the Bodily Theory systematically by showing how it accommodates acute pain, chronic/neuropathic pain, phantom limb pain, and referred pain. Secondly, I provide a novel argument in favour of the Bodily Theory – the Common Sense Argument. Thirdly, I provide objections to the common arguments for the Experiential Theory.

§2 distinguishes different senses of ‘pain’ and argues that the debate between the Bodily Theory and the Experiential Theory is non-verbal. §3 develops the Bodily Theory and shows how it accommodates acute pain, chronic/neuropathic pain, phantom limb pain, and referred pain. §4 offers two arguments in favour of the Bodily Theory – the Common Sense Argument

---

<sup>36</sup> This debate about *pain* is distinct from the debate about the *concept of pain*, where some argue that there is one bodily concept (Reuter and Sytsma 2020), others argue that there is one concept with both bodily and mental aspects (Borg et al. 2020), and yet others argue that there are two concepts – one mental and one bodily (Liu 2021).

<sup>37</sup> I use terms like ‘mental’, ‘mind’ and ‘experience’ in accordance with the tradition that takes these to be associated with the brain rather than the body, but nothing turns on this. If one holds a different view, for example, that the mental extends to the body and beyond, one may count as mental what I take to fall on the bodily side of the mind/body divide. In that case, one can make the contrast between the Experiential Theory and the Bodily Theory in physiological language, without terms like ‘mental’ and ‘experience’, e.g. contrasting the view that pains are occurrences in the central nervous system (i.e. the brain and spinal cord) with the view that pains are occurrences in the peripheral nervous system (i.e. nerves and ganglia outside the brain and spinal cord).

<sup>38</sup> The exception is Newton (1989), discussed in footnote 46.

and the Semantic Argument. §5 argues that the common arguments in favour of the Experiential Theory are unconvincing. §6 concludes.

## 2. Different Senses of ‘Pain’ and Why the Debate is Non-Verbal

This section distinguishes different senses of ‘pain’, which is important in order to identify the target phenomenon, avoid conceptual conflation, and show that the debate between the Experiential Theory and the Bodily Theory is non-verbal.

### 2.1. *Different Senses of ‘Pain’*

I said above that pain in the locatable sense is what ‘pain’ refers to in sentences like ‘There is a pain in my foot’. The Bodily Theory is only a claim about pain in this sense, not a claim about pain in other senses of the word.

Firstly, the Bodily Theory is not a theory about pain in the *phenomenal* sense, which is often referred to as ‘painfulness’ or ‘unpleasantness’ as well as ‘pain’.<sup>39</sup> Pain in the phenomenal sense is commonly taken to be a property of experiences and is not what the Experiential Theory holds is an experience itself and the Bodily Theory holds is a bodily object of an experience. Even though pain in both the locatable and the phenomenal senses of the word may be related, they are commonly taken to be dissociable. For example, several theorists believe that subjects with pain asymbolia have pain in the locatable sense, but no pain in the phenomenal sense (e.g. Grahek 2007, Bain 2014, Corns 2020). And eliminativists about phenomenal properties believe that we have pains in the locatable sense but that there is no pain in the phenomenal sense (e.g. Frankish 2019: 92). If we did not distinguish pain in the locatable sense from pain in the phenomenal sense, then we would beg the question against both these views, which would be unnecessarily strong for a view about pain in the locatable sense.<sup>40</sup>

Secondly, the Bodily Theory is not a theory about pain in the *experiential* sense. Many scientists and philosophers call the experience, feeling, or awareness associated with having a pain in the locatable sense ‘pain’. This is arguably not a common use of the word among lay

---

<sup>39</sup> Chalmers (1996) makes a similar distinction between pain in a functional sense and pain in a phenomenal sense (1996: 17).

<sup>40</sup> Aydede (2017, 2020) holds the unusual view that pain in the locatable sense *is* pain in the phenomenal sense and says that: ‘When I correctly report a pain in my elbow, I am introspectively reporting pain (a certain phenomenal quality) as literally being in my elbow’ (2017: 228). On this view, pain in the locatable sense is neither an experience nor a bodily occurrence, but a phenomenal property. This is an interesting view and for present purposes I have no objection to it. But since it is unusual and I focus on the contrast between the Bodily Theory and the Experiential Theory I shall not discuss this view further.

people. But since philosophers and scientists commonly use ‘pain’ to refer to what I shall refer to as ‘the experience of pain’, there is a sense in which ‘pain’ refers to an experience. On my view, the experience of pain is the experience of a bodily occurrence, so there is a distinction between pain (in the locatable sense) and the experience of pain. According to the Experiential Theory, by contrast, pain (in the locatable sense) is an experience, so on this view, there might not be any distinction between pain and the experience of pain. In any case, the present point is that the Bodily Theory is only a theory about the referent of ‘pain’ as it is used in contexts like ‘There is a pain in my foot’, it is not a theory about the referent of ‘pain’ as it is used in contexts where it clearly refers to an experience. There are several examples of the latter. For example, in Byrne’s (2001) talk about ‘the distinction between pains-as-experiences and pains-as-objects-of-experiences’ (2001: 229) the first occurrence of ‘pains’ refers to experiences. And when Aydede (2017) says that ‘there is no serious alternative to identifying pains with experiences’ (2017: 224) ‘pains’ refers to experiences, because Aydede distinguishes between pain in the experiential sense (which he believes is an experience) and pain in the locatable sense (which he does not believe is an experience, but a phenomenal property).<sup>41</sup> The Bodily Theory is not a theory about the referent of ‘pain’ in these examples where it refers to an experience.

This shows that the mere claim that ‘Pains are experiences’ is ambiguous, as this might *either* mean that pains in the experiential sense are experiences, *or* that pains in the experiential sense *and* pains in the locatable sense are experiences. Only the latter claim is in competition with the Bodily Theory. I suspect that many theorists who say that ‘Pains are experiences’ are not clear about which claim they are making, which makes it difficult to judge whether they defend the Experiential Theory, understood as a theory of pain in the locatable sense, or whether they are just concerned with pain in the experiential sense.

Of course, it may be confusing that the term ‘pain’ is used in so many different ways. But this is the linguistic reality we face, so the best one can do when discussing pain is to be clear which use of the word one has in mind. In any case, theorists like Aydede (2017, 2020) and Byrne (2001) are not experientialists as I use that word, because they do not claim that pains in the locatable sense are experiences, they only claim that pains in a different (experiential) sense of the word are experiences, and that is perfectly consistent with the Bodily Theory. My target is not *everyone* who claims that ‘Pains are experiences’, it is only those who claim that pains in the *locatable* sense are experiences. The Bodily Theory, which is only a

---

<sup>41</sup> Aydede (2017, 2020) is not explicit about this ambiguity, but has confirmed this in personal communication.

claim about pains in the locatable sense, is consistent with the claim that pains in a *different* sense of the word are experiences.<sup>42</sup>

## 2.2. *Why the Debate is Non-Verbal*

Given that ‘pain’ is used in so many different ways, one might think that the contrast I drew between the Experiential Theory and the Bodily Theory is superficial. Perhaps experientialists are only concerned with pain in the experiential sense? No. Experientialists are not just concerned with the referent of ‘pain’ in contexts where it obviously refers to an experience. Rather, they really believe that ‘pain’ refers to an experience in sentences like ‘There is a pain in my foot’. In order to see this, I must first elaborate a little on the Experiential Theory.

The Experiential Theory comes in two versions: representationalism and imperativism. Representationalists believe that pains represent something in the world, such as bodily disturbances or tissue damage (Armstrong 1968, Tye 1995a, 1995b, 2002, 2005a, 2005b, 2017). Imperativists hold either that pains have imperative content – conveying a command to protect a body part (Klein 2015) – rather than representational content, or that pains have *both* representational *and* imperative content (Hall 2008). With the distinction between representationalism and imperativism in mind, consider the following passages where experientialists identify pain in the locatable sense with an experience.

[T]he location of the pain in the hand is an intentional location, that is, that is simply the place where a disturbance feels to be [...] The pain itself will be the perception of these disturbances [...] and will be contingently identified, not with a physical happening at the ‘place of the pain’, but with an event in the central nervous system (Armstrong 1968: 316, 319-320).

[A] pain in the leg, I suggest, is a token sensory experience which represents that something in the leg is damaged (Tye 1995a: 228).<sup>43</sup>

---

<sup>42</sup> There are many other senses of ‘pain’, which are distinct from the three senses discussed above. ‘Pain’ is used to refer to emotional states like grief, the journal *Pain*, and bread in French, among other things. My claim is not that the word is used only in the locatable, phenomenal and experiential senses, but that we must be particularly careful to distinguish these three senses of the word. Since they concern importantly related phenomena the danger of conceptual conflation and verbal dispute is much more likely to arise from failing to distinguish these three senses than from failing to distinguish other senses of the word not mentioned in the main text.

<sup>43</sup> In Tye (2005a), Tye says that ‘a pain in a leg (viewed as an experience) represents that a certain quality is tokened in the leg’ (Tye 2005a: 101). He thinks the quality represented is tissue damage and confusingly calls tissue damage ‘pain’ (2005a: 101). But note that this cannot be pain in the locatable sense, given his claim that

The difference between a mild burning pain in your foot and a mild burning pain in your hand, then, is that the former commands you to protect your foot, whereas the latter commands you to protect your hand (Klein 2015: 90).

The sharp pain in your lower back when you're lifting something heavy tells you to stop doing that. The pain in your hand when you grasp something hot or sharp tells you to stop doing that – withdraw (Hall 2008: 534).

It is clear that these philosophers are concerned with pain in the locatable sense.<sup>44</sup> They really believe that pains in the locatable sense are not located in the body but in the mind or brain. As Klein puts it: '[P]ains *themselves* are located, if they're located anywhere, somewhere in the head. This is a general feature of intentionalist accounts [...] Representational accounts of pain have a similar structure: pains are in the head, but they represent some damage or disturbance at a bodily location, and that location corresponds to the felt location of pain' (2015: 88).<sup>45</sup>

This theory is my target: the theory that 'pain' refers to an experience in the mind or brain in sentences like 'There is a pain in my foot'. But again, my target is not everyone who says that 'Pains are experiences'. Nor is my target necessarily everyone who identifies as a representationalist or imperativist. Some philosophers might only be representationalists or imperativists about the experience of pain, i.e. pain in the experiential sense, and not about pain in the locatable sense. Other philosophers might not be clear what sense of 'pain' we should associate with their representationalism or imperativism. But the philosophers cited above are clearly representationalists or imperativists about pain in the locatable sense.

The fact that experientialists are concerned with pain in the locatable sense shows that the dispute between the Bodily Theory and the Experiential Theory is non-verbal. If experientialists were talking about pain in a different sense of the word, then this dispute would be merely verbal, and the Experiential Theory would not seem immediately relevant to answer the question of what pains in this familiar sense are. But since both theories are about the

---

pain in that sense is an experience. In Tye (2017: 483n6) he regrets calling tissue damage 'pain' and calls this use of the word 'confused and wrong-headed'.

<sup>44</sup> Though some experientialists also try to explain the painfulness of pains in terms of the content of the experiences they identify with pains, so their versions of the Experiential Theory are about *both* pain in the locatable sense *and* pain in the phenomenal sense.

<sup>45</sup> See Armstrong (1968: 316, 319-20) and Tye (1995a: 229) for similar claims about the location of pain.

referent of ‘pain’ in the same sense of the word, there is a genuine disagreement about the metaphysics of pain.

### 3. The Bodily Theory

According to the Bodily Theory, pains are bodily occurrences. But which bodily occurrences? Many things can be called ‘bodily occurrences’, so we should like something more specific for a precise theory of what pains are. This question is rarely addressed by advocates of the Bodily Theory. Byrne says that pains are ‘objects-of-experiences’ in body parts (2001: 229). Hyman (2003) says that pains are ‘states of a body part’, located in a body part, and consist in the body part hurting (2003: 15). Hacker (2013) says that pains are ‘physical’, not mental, and located in the body (2013: 263, 268). Massin (2017) says that pains are ‘non-intentional bodily episodes’ that are not necessarily experienced but are necessarily bad or disagreeable (2017: 321, 323, 329-330). Reuter and Sytsma (2020) say that pains are ‘bodily states’ and ‘qualities of bodily disturbances’ (2020: 1778, 1783). Finally, Bradley says that pains are ‘constitutively mind-dependent properties instantiated in part of the subject’s body’ (2021: 3) and that ‘a pain is nothing other than a body part’s *hurting* or *having a pain in it*, which is its instantiating the property PAIN’ (2021: 9).

These formulations tell us where pains are, but they do not tell us exactly what they are. Given both that physiology offers the best account of our bodies and occurrences in them, and that the Bodily Theory says that pains *are* bodily occurrences (or ‘states’, ‘episodes’, or ‘properties’) located in our bodies, a defender of the Bodily Theory should be able to say which of the bodily occurrences accepted by physiology pains are. For I take it that defenders of the Bodily Theory do not want to claim that they – through philosophical reflection – have discovered some occurrences in our bodies of which physiologists – who use the best empirical methods – are ignorant. The question is then: is the pain in my foot tissue damage, C-fiber stimulation, or something else accepted by physiology?

I propose the following account of what pains are:

For any event *x*, *x* is a pain if and only if *x* (a) is a proximal activation of nociceptors, and (b) causes an experience of pain.

§3.1. comments upon (a), §3.2. upon (b), and §3.3. addresses some objections.<sup>46</sup>

### 3.1. *Pains as Proximal Activations of Nociceptors*

By ‘nociceptors’ I mean neurons located in the peripheral nervous system that respond to a variety of stimuli – including noxious stimuli like extreme heat and cut, and innocuous stimuli like light pressure and small temperature changes. These receptors involve both A $\delta$ -fibers and C-fibers and are active when one has a pain, but also in everyday contexts unaccompanied by pain (Baliki and Apkarian 2015: 474-476). This understanding of nociceptors seems to be the most common (see e.g. Baliki and Apkarian 2015 and references therein). But some scientists seem to hold that instead of there being one kind of receptors that respond to different stimuli, there are two kinds of receptors – one that responds to innocuous stimuli, and one that responds to noxious stimuli, and that only the latter are nociceptors (e.g. Lee et al. 2011: 2). I am not in a position to judge which understanding is better, but at some level the difference is merely verbal. In any case, as the former seems to be the most common, I follow this and understand nociceptors as receptors that respond to both noxious and innocuous stimuli, the activity of which may occur with and without pain.

The Bodily Theory is sometimes associated with the view that pain is tissue damage, since standard cases of acute pain involve noxious stimulus and tissue damage. But as several philosophers rightly point out, it would be implausible to identify pain with tissue damage, as neither chronic/neuropathic pain (pain resulting from disorder of the somatosensory system), nor phantom limb pain (pain reported in an amputated, non-existent limb), nor referred pain (when the pain is reported as being in a different place from where the injury is) need involve tissue damage (Aydede 2017: 223, 2020: 146, Borg et al 2020: 32n9, 36). Still, as I am about to explain, all these kinds of pain involve activations of nociceptors.

---

<sup>46</sup> This formulation is close to both that of Hill (2017) and that of Newton (1989). According to Hill (2017), pain is ‘a type of disturbance that generally involves actual or potential damage, and that is grounded principally in the activity of nociceptive neurons known as C-fibers and A $\delta$ -fibers’ (2017: 61). This is close to (a) but misses (b). As I argue in §3.2., however, both (a) and (b) are necessary. Another difference between Hill’s view and mine is that Hill’s view is disjunctive: acute pains are bodily occurrences, but chronic pains and phantom limb pains are experiences in the brain (2017: 67). On my view, by contrast, all pains are bodily occurrences (see below).

With regard to Newton’s (1989) view, I suspect that the differences are mainly verbal. At one place, Newton says that pains are ‘powers in parts of our bodies to produce sensations of pain in us’ (1989: 576). At a different place, however, she says that ‘a person has a pain if and only if there is nociceptor stimulation’ (1989: 590). But given what Newton says elsewhere, I believe the latter biconditional is a slip of tongue. It should have been a conditional, since (as I discuss in §3.2.) the experience (or ‘sensation’ in Newton’s terminology) of pain is necessary for pain, which is why (b) is needed in addition to (a) in my formulation of the Bodily Theory above.



In *chronic/neuropathic pain*, nociceptors transmit signals to the brain – sometimes after innocuous stimulus and sometimes without being preceded by any stimulus – which, due to neural disorders, causes an experience of pain (Alles and Smith 2018, Baliki and Apkarian 2015: 482). The relevant neural disorders may be central mechanisms, so central mechanisms are part of the *explanation* of chronic/neuropathic pain. But the *phenomenon* chronic/neuropathic pain always involves an activation of nociceptors that can be identified as the pain.<sup>47</sup> In cases that do not involve phantom limbs, the pain is where the patient reports it as being. For example, in chronic back pain the pain is really in the back and in chronic headache the pain is really in the extra-cranial head, because there are activations of nociceptors in the back and extra-cranial head that cause experiences of pain.

In *phantom limb pain* – which is a type of chronic/neuropathic pain – nociceptors are activated in a different place from where the patient reports the pain as being, namely in the residual limb or stump. The nerves in the residual limb were damaged during amputation, and pain scientists consider it likely that those nerves play an important role in phantom limb pain. More specifically, there is an upregulation of voltage-sensitive sodium channels in the nociceptors, which increases spontaneous afferent input to the spinal cord. In other words, the nociceptors of the damaged nerves become hypersensitive and transmit signals to the central nervous system, even in the absence of any stimulus, which causes an experience of pain. There can also grow neuromas from the damaged nerves, i.e. enlarged endings of C-fibers and A-fibers that send signals to the central nervous system and cause an experience of pain (Coppes and Sang 2018: 2, Flor 2002: 184).

Pain scientists also appeal to central mechanisms in the explanation of phantom limb pain. For example, scientists hypothesize that amputation leads to a cortical reorganization in the brain such that the cortical areas of the amputated body part are taken over by adjacent zones of the body, which could explain why stimulation of those body parts results in phantom limb sensations (Coppes and Sang 2018: 2). But even though these central mechanisms are part of the *explanation* of phantom limb pain, the *phenomenon* phantom limb pain always involves activations of nociceptors that can be identified as the pain, namely those of the damaged nerves in the residual limb. These nerves were previously connected with the nerves in the limb where

---

<sup>47</sup> The question of what pains are is independent from the question of what explains the existence of token pains, and the claim that pains are activations of nociceptors is consistent with explanations of pains appealing to other mechanisms, such as central mechanisms, cognitive states, and much more. See Corns (2020) for more on the explanation of pain.

the patient reports feeling pain, so activity in these nerves can thus both be the pain the patient feels and contribute to cause the patient to misidentify the location of the pain.

In *referred pain* too, nociceptors are activated in a different place from where the patient reports the pain as being. Somatic referred pain, for example, is produced by noxious stimulation of spinal structures, while the pain is referred to and reported as being in lower limbs. Scientists believe that nociceptors connected to both the lower limbs and the injured tissue around the spine converge in the spinal cord and that this causes the referral (Bogduk 2009: 17). Explanations like this are the norm. Although the details of the mechanisms of referred pain remain controversial, it is generally believed that neurons in the central nervous system receive convergent inputs from various tissues at different locations in the body and that this leads higher centres to misidentify the actual input source (Arendt-Nielsen and Svensson 2001: 11, 15-17). It is also possible to induce referred pain to limbs with complete sensory loss due to anaesthetic block, i.e. limbs where nociceptive signals are blocked from travelling to the brain (Arendt-Nielsen and Svensson 2001: 14). In that case, there is no causal connection between the activation of nociceptors in the referred location and the experience of pain, which I claim is necessary for pain. Thus, my version of the Bodily Theory is committed to the pain having the location of the injury, not the referred location where the patient reports the pain as being. Just as the location is illusory for phantom limb pain, so it is illusory for referred pain. Interestingly, this prediction seems to accord with common sense, as indicated by Kim et al.'s (2016) studies, where the majority of subjects held that referred pains are located where the injury is (2016: 147, 154).

The upshot is that there are no pains without activations of nociceptors. Acute pain involving tissue damage obviously involves the activation of nociceptors, but this is also true of chronic/neuropathic pain, phantom limb pain, and referred pain. In lack of other bodily occurrences with which the Bodily Theory can plausibly identify pains, a defender of the Bodily Theory should hold that the activation of nociceptors is necessary for pain. The *argument* that it is necessary is the argument for the Bodily Theory, offered in §4.

The final point I want to make about the rationale for (a) in the formulation of the Bodily Theory concerns the qualification 'proximal'. This qualification is needed because there can be activations of nociceptors that cause experiences of pain without being pain. As explained above, nociceptors are not just active when one is exposed to a noxious stimulus or has a pain but are continuously active in everyday contexts unaccompanied by pain, for example, as the result of light pressure or small changes in temperature. However, an activation of nociceptors unaccompanied by pain may cause a different event, which may cause a different event (etc.),

which may eventually cause an experience of pain. By transitivity of causation, it follows that the original activation of nociceptors is a cause of the relevant experience. But it is not pain, because it is not the *proximal* activation of nociceptors, but rather, a distal one. For example, the activation of nociceptors in my back may cause me to move my hand to scratch my back, and my moving my hand may cause the book to fall down from the desk to my toe, which eventually causes an experience of pain in my toe. In that case, the original activation of nociceptors in my back is a cause of the experience of pain in my toe, but it is not pain, because it is distal, not proximal.

### 3.2. *The Experience-Dependence of Pains*

According to my formulation of the Bodily Theory, a proximal activation of nociceptors is not sufficient for pain. The reason is the fact just mentioned, namely that nociceptors can be activated without being accompanied by pain. Despite this, it does not follow, as some scientists suggest, that pains are not activations of nociceptors (Apkarian 2017: 74, Baliki and Apkarian 2015: 474).<sup>48</sup> Pains can still be identified with activations of nociceptors, it is just that they are not activations of nociceptors *simpliciter*. This is why (b) is needed in the above formulation of the Bodily Theory, saying that an activation of nociceptors is only a pain if it causes an experience of pain. Put differently: being a pain is not an *intrinsic* property of activations of nociceptors, but an *extrinsic* one, because only those activations of nociceptors that cause experiences of pain are pains.

This idea of pains is in several ways similar to Dretske's (2003: 5) idea of crocks – rocks that you experience. Not all rocks are crocks but crocks really *are* rocks and 'inherit' the properties of rocks, such as their location. And crocks depend on your experience: if your experience were absent there would be no crocks, only rocks. Similarly, not all proximal activations of nociceptors are pains, but pains *are* proximal activations of nociceptors and 'inherit' the properties of activations of nociceptors, such as their location. And pains depend on experience: no experience, no pain.

To depend on other things without being those other things is not peculiar to crocks and pains. A rock star depends on his fans in order to be a rock star. Without the fans he would just be a guy with a guitar, not a rock star, but it is the guy with the guitar, not the fans, who is the rock star. A widow depends on the death of her husband in order to be a widow, but it is the

---

<sup>48</sup> Though they seem to use 'pain' and 'pain perception' synonymously, so perhaps the claim is just that experiences of pain (i.e. pains in the experiential sense) are not activations of nociceptors, which is consistent with the Bodily Theory.

woman, not the death of the husband, who is the widow. A £5 note depends on the Bank of England in order to be £5, but it is the paper note, not the Bank of England, which is the £5. Similarly, the Bodily Theory claims that pains depend on experiences without being experiences.

This view differs substantively from the Experiential Theory. The claim that pains are bodily occurrences located in the body that depend on experiences in the brain (my version of the Bodily Theory) is not the same as the claim that pains are experiences located in the brain (the Experiential Theory).

### 3.3. *Objections/Clarifications*

There are three potential worries with this view that are worth addressing. The first is that if being a pain is an extrinsic property of activations of nociceptors, then it has no location, because properties are abstract objects and abstract objects have no location. But my claim about location is not about the *property* of being a pain, but about *instantiations* of that property, i.e. token pains. Rockstarhood, widowhood, and being £5 are extrinsic properties of guitar players, women, and paper notes respectively. If properties are abstract objects (which I am neutral about here), then these properties may have no location. But rock stars, widows, and £5 notes do have locations. Similarly, pains have bodily locations, even though the property of being a pain may have no location.

The second worry is that pains cannot be activations of nociceptors because activations of nociceptors do not figure in the content of experiences of pain, but pains *do* figure in that content. But the Bodily Theory is not meant to be an account of the *de dicto* content of experience. It is a theory of what pains are – the *de re* content of experience. The claim is not that experiences represent anything else than pains *de dicto*, but that the pains that experiences represent are identical with proximal activations of nociceptors.

The third worry is that the Bodily Theory is circular: in my claim that pains are proximal activations of nociceptors that cause experiences of pain, ‘pain’ occurs on both sides of the identification. I think two responses to this worry are available. The first is to deny that there is any circularity.<sup>49</sup> Even if one accounts for pain in terms of the experience of pain, one need not account for the experience of pain in terms of pain. Rather, one can account for the experience of pain in terms of ostensive definition. Thus, there is no circularity. The second response is to

---

<sup>49</sup> Menzies and Price (1993) make this point about dispositional accounts of colour and causation (1993: 194).

deny that there is any problem even if there is a circularity.<sup>50</sup> It is one thing to make a claim about the identity of pains and activations of nociceptors, and a different thing to explain what pains are to someone who is entirely ignorant of it or giving a synonym of pain terms. Explaining what pains are in terms of experiences of pain would not help someone who has absolutely no idea what pains are. And given reasonable assumptions about meanings, ‘pain’ cannot be synonymous with a complex phrase containing ‘the experience of pain’. But these worries are not relevant here, because the Bodily Theory is a view about the metaphysics of pain and only makes a claim about the identity of token activations of nociceptors and pains. Though circular in the sense that ‘pain’ occurs on both sides of the identification, this identity claim is neither viciously circular nor trivial, but legitimate and informative.

#### **4. Arguments for the Bodily Theory**

The previous section outlined what I take to be the best version of the Bodily Theory. But why should one believe that the Bodily Theory is true? This section offers two arguments: the Common Sense Argument and the Semantic Argument.

##### *4.1. The Common Sense Argument for the Bodily Theory*

The first argument in favour of the Bodily Theory is that it accords better with common sense than the Experiential Theory. The argument can be spelled out thus:

- (P1) If a theory accords better with common sense than another theory, then the former is more plausible than the latter, unless there are independently good reasons to adopt the latter.
- (P2) The Bodily Theory accords better with common sense than the Experiential Theory.
- (P3) There are no independently good reasons to adopt the Experiential Theory.
- (C) The Bodily Theory is more plausible than the Experiential Theory.

(P1) says that we should prefer a view that coheres better with our pre-theoretical beliefs than a view that demands more revision of our beliefs, unless the more revisionary view is

---

<sup>50</sup> Byrne and Hilbert (2011) make this point about the dispositional account of colour (2011: 342-343).

independently more plausible. This is at least supported by pragmatic considerations, perhaps also the historical success of less revisionary theories.

(P2) is supported by empirical studies. In Sytsma's (2010) study, the majority of subjects held that acute pains are located in body parts rather than the mind, and in Kim et al.'s (2016) study, the majority of subjects held that referred pains are located in the injured body part rather than the mind (2016: 147, 152). My version of the Bodily Theory agrees with these common sense beliefs about the location of pains, but according to the Experiential Theory, they are strictly speaking false.

It should be noted that Borg et al. (2020) and Salomons et al. (2021) claim that their experimental findings are in tension with the Bodily Theory, which may seem to threaten (P2). There are two findings that are relevant. The *first* finding is that the subjects in their studies judged that there is *no* pain in cases where a patient has bodily damage but denies pain, such as for athletes distracted from injuries and patients with congenital insensitivity to pain (Borg et al. 2020: 39-40, Salomons et al. 2021: 10-11). The *second* finding is that subjects judged that there *is* pain in cases with reported pain but no bodily damage, such as in chronic/neuropathic stomach ache and referred pain caused by electric shock to the brain (Salomons et al. 2021: 10-11).

But neither of these findings are in tension with my view. Setting aside the fact that the subjects in Reuter and Sytsma's (2020) studies judged the opposite to Borg et al.'s and Salomons et al.'s subjects about cases where patients have bodily damage but deny pain, it is true that the first finding arguably conflicts with a version of the Bodily Theory on which experience is not necessary for pain. This is because this version of the Bodily Theory would arguably count bodily damage as pain, even if the patient denies having pain.<sup>51</sup> But the first finding above does not conflict with my version of the Bodily Theory, on which the experience of pain is necessary for pain.

Regarding the second finding, it is true that this conflicts with a version of the Bodily Theory on which bodily damage is necessary for pain, because this version of the theory predicts that there is no pain in cases with no bodily damage, even if the patient reports pain.<sup>52</sup>

---

<sup>51</sup> Perhaps Massin's (2017) view is such a view. As mentioned above, Massin (2017) claims that pains are 'bodily episodes' that are not necessarily experienced but are necessarily bad or disagreeable (2017: 321, 323, 329-330). If Massin thinks that bodily damage is bad or disagreeable – in whatever non-experience-involving sense Massin understands these terms – even if the patient denies having pain, then bodily damage is (on Massin's view) sufficient for pain, which conflicts with the first finding above.

<sup>52</sup> Perhaps this is Reuter and Sytsma's (2020) view, as they claim that pains are 'qualities of bodily disturbances' (2020: 1783). If what Reuter and Sytsma mean by 'bodily disturbance' is bodily damage, then their view conflicts with the second finding above.

But the second finding above does not conflict with my version of the Bodily Theory, since on my view, it is the activation of nociceptors, not bodily damage, that is necessary for pain. In other words, none of the findings of Borg et al. and Salomons et al. are in tension with my view, so given the findings of Sytsma (2010) and Kim et al. (2016) mentioned above, (P2) stands.

(P3) is supported by the discussion in §5, where I argue that the common arguments for the Experiential Theory are unconvincing. Thus, there is reason to believe (P1), (P2) and (P3), in which case we can infer that the Bodily Theory is more plausible than the Experiential Theory.

#### 4.2. *The Semantic Argument for the Bodily Theory*

The second argument in favour of the Bodily Theory is that it provides a better interpretation or semantics of ordinary pain reports than the Experiential Theory.<sup>53</sup> According to the Bodily Theory, the semantics (i.e. truth-conditions) of ordinary pain reports are very simple, intuitive and straightforward. ‘There is a pain in my foot’ really means (i.e. has the truth condition) that *there is a pain in my foot*. Pain reports may sometimes be false, as in the case of phantom limb pain and referred pain, but at least the Bodily Theory requires us to take them at face value, such that when they are true, they are *literally* true. While it is right that the Bodily Theory requires the existence of experience for these pain reports to be true, what makes ‘There is a pain in my foot’ true is that there *is* a pain located in my foot.

With the Experiential Theory things are different, because there is no pain in my foot that makes ‘There is a pain in my foot’ true. One option is to say that pain reports are false. On this view, no one (except experientialists themselves) have ever been right about the location of their pains and people assert falsehoods all the time when they try to specify the location of their pains. But this interpretation of ordinary pain reports is very uncharitable, and therefore less plausible than the interpretation offered by the Bodily Theory. The other option for the experientialist is to say that pain reports are true, because ‘There is a pain in my foot’ does not really mean (i.e. does not have the truth-condition) that there is a pain in my foot. On an experiential-friendly semantics, ‘There is a pain in my foot’ means that I have a pain representing a bodily disturbance in my foot (representationalism), or that I have a pain informing me to protect my foot (imperativism). But this is a more complex, less intuitive, and less straightforward semantics than that offered by the Bodily Theory. Thus, we should prefer the Bodily Theory.

---

<sup>53</sup> A similar argument is made by Hyman (2003: 9) and Hill (2014: 168-169).

Both the Common Sense Argument and the Semantic Argument for the Bodily Theory presuppose the standard version of the Experiential Theory held in the literature, according to which pains are located in the mind or brain (e.g. Armstrong 1968: 316, Klein 2015: 88, Tye 1995a: 229). One might think that if those experientialists who believe that pains have representational content adopted a certain externalist theory of representational mental states and their content, then they could object to the arguments above. On the externalist theory I have in mind, representational mental states are partly constituted by the external objects that (on this view) determine the contents of those states (cf. McDowell 1986). So for example, the belief that the President is in the White House is partly in the White House, since the President, who determines the content of the belief, is there. On the externalist experiential view of pain then, the pain – representing a disturbance in a body part – is partly in the body, because there is a disturbance there determining the content of the pain. Thus, on this externalist experiential view, pains *are* located in body parts. This might make room for a semantics of ordinary pain reports that is equally plausible as that offered by the Bodily Theory. However, this externalist view would not accord as well with common sense as the Bodily Theory. Kim et al.'s (2016) and Sytsma's (2010) studies do not only suggest that lay people believe that pains are located in the body, but that they are located in the body *rather than* the mind. This accords with the Bodily Theory, but not the above externalist version of the Experiential Theory, since pains *are* located in the mind (i.e. brain and body) on the latter view. So even if defenders of the representationalist version of the Experiential Theory adopted the above externalist theory, there would still be reason to prefer the Bodily Theory.

### **5. Arguments for the Experiential Theory and Why They are Unconvincing**

The previous section offered two arguments for the Bodily Theory. Given these arguments, there is reason to believe that pains are bodily occurrences. Since the best candidates for bodily occurrences are proximal activations of nociceptors that cause experiences of pain, there is reason to believe that pains are proximal activations of nociceptors that cause experiences of pain. However, if there were equally strong arguments in support of the Experiential Theory, then we would have no reason to prefer the Bodily Theory, as there would be reasons to believe both that pains are bodily occurrences and reasons to believe that pains are experiences. This section evaluates the arguments for the Experiential Theory put forth by experientialists and argues that they are unconvincing. Thus, we should prefer the Bodily Theory.



### 5.1. *The Common Sense Argument for the Experiential Theory*

The first argument in favour of the Experiential Theory is due to Tye (1995a, 2002, 2005a, 2017) and draws on what (according to Tye) is common sense about pain. According to Tye, the following three claims are part of the common sense conception of pain:

PRIVACY: Only I can have my pains – they belong to one person only and cannot be had by other people (Tye 1995a: 228, 2002: 151, 2005a: 100, 2017: 478).

INCORRIGIBILITY: If I feel a pain, then I have a pain. Sometimes this idea is referred to as ‘epistemic authority’, ‘first person authority’, or ‘no appearance/reality distinction’, but the idea is the same: feeling a pain implies having a real pain, regardless of stimulus to and states of one’s body (Tye 2017: 478).

SUBJECTIVITY: If I have a pain, then I feel a pain. There are no unfelt pains, i.e. pains that escape one’s experience or awareness (Tye 2002: 151, 2017: 478).

There are two ways in which one can take these claims to support the Experiential Theory. On a *weak* version of the argument, the idea is that the Experiential Theory accords better than alternatives with PRIVACY, INCORRIGIBILITY and SUBJECTIVITY, which gives the Experiential Theory a theoretical advantage over the Bodily Theory, because we should prefer a view that demands less revision of our pre-theoretical beliefs (common sense) to a view that demands more revision. On a *strong* version of the argument, the idea is that the Experiential Theory offers the best explanation of the properties described by PRIVACY, INCORRIGIBILITY and SUBJECTIVITY. This version assumes not only that PRIVACY, INCORRIGIBILITY and SUBJECTIVITY are part of the common sense conception of pain, but that common sense is true and thus that PRIVACY, INCORRIGIBILITY and SUBJECTIVITY describe genuine *explananda*. This assumed link between common sense and metaphysics obviously makes the strong version of the argument more controversial than the weak version.

It seems to me that Tye has the strong version of the argument in mind (see Tye 1995a: 228, 2002: 151, 2005a: 100, 2017: 478), so one could object to his argument by denying the link between common sense and metaphysics. However, Tye (or other experientialists) could obviously adopt the weak version of the argument instead. I shall therefore argue that both the weak and the strong version of the argument fail independent of any link between common sense and metaphysics. They fail partly because empirical studies suggest that Tye’s account

of common sense is mistaken, and partly because my version of the Bodily Theory actually accords with what Tye takes to be common sense. Let's look closer at PRIVACY, INCORRIGIBILITY, and SUBJECTIVITY in respective order.

With regard to PRIVACY, data from experimental philosophy suggest that Tye is wrong about common sense. The majority of the subjects in Sytsma's (2010) studies allowed for shared pains when two patients share a body part, e.g. for Siamese twins joined at the torso who injure their shared foot. Thus, PRIVACY, which does not allow for shared pains, is plausibly not part of the common sense conception of pain.

As Sytsma's studies indicate what is common sense about privacy, one may think that experientialists can modify the appeal to privacy. For example, experientialists could argue that private pains are pains representing disturbance in, or conveying a command to protect, a body part belonging only to oneself, while shared pains are pains representing disturbance in, or conveying a command to protect, a shared body part. In other words, experientialists could argue that shared pains are pains with the *same content*. To say that two people 'share' or have the 'same' pain is not to say that there is one token state or experience belonging to two people. Rather, it is to say that two token states or experiences belonging to different people have the same content, analogous to two people having the same belief, i.e. there being two token states of believing that have the same content.

But there is no reason to believe that this experientialist account of private and shared pains is better than the one offered by the Bodily Theory. Only I can have the pain in my foot, because only I and no one else has that foot as a body part and only my brain and no one else's brain receives nociceptive input from that foot. The alleged cases of shared pains where privacy does not obtain are cases where people *share* body parts, which the Bodily Theory accommodates by the pain – i.e. the proximal activation of nociceptors that causes an experience of pain – being in that shared body part.

So appeal to (lack of) privacy indicates no advantage of the Experiential Theory over the Bodily Theory. Nevertheless, this brings out an interesting difference between the two theories: the Experiential Theory is compatible with PRIVACY, but the Bodily Theory is not. When two people (e.g. Siamese twins) report pain in a shared body part there is only one relevant activation of nociceptors, but, according to PRIVACY, two pains. Thus, the Bodily Theory contradicts PRIVACY, because it predicts that there is only one pain. By contrast, the Experiential Theory predicts that there are two pains, because there are two experiences.

With regard to INCORRIGIBILITY, the empirical data again suggest that Tye is wrong about common sense. The majority of the subjects in Reuter et al.'s (2014) studies allowed for

feeling of pain without real pain, for example, as the result of taking a drug (2014: 84-88). So INCORRIGIBILITY – the claim that if I feel a pain, then I have a pain – is plausibly not part of the common sense conception of pain. But regardless of this, neither the Experiential Theory nor the Bodily Theory entail anything about incorrigibility, and both theories are consistent with both INCORRIGIBILITY and what Reuter et al.’s studies indicate about common sense, so no appeal to (lack of) incorrigibility would indicate an advantage of the Experiential Theory over the Bodily Theory.

With regard to SUBJECTIVITY, the empirical data are more difficult to interpret than for PRIVACY and INCORRIGIBILITY. On the one hand, the majority of the subjects in Reuter and Sytsma’s (2020) studies judged that patients distracted from injuries (e.g. soldiers and athletes) have pains they do not feel (2020: 1787-1798), so these subjects believed that pains can occur without experience. But on the other hand, as discussed above, the majority of the subjects in Borg et al.’s (2020) and Salomons et al.’s (2021) studies judged that patients distracted from injuries have no pain (Borg et al. 2020: 39-40, Salomons et al. 2021: 10-11), suggesting that these subjects did not believe that pains can occur without experience. Moreover, Aydede (2020) found that the majority of pain scientists and clinicians take pains to be experience-dependent (2020: 154n27). Given these ambiguous data, it is difficult to judge what is common sense about subjectivity. But regardless of this, the Experiential Theory does not accord better with SUBJECTIVITY than the Bodily Theory. Gate control theory suggests that in cases where people distracted from injuries deny having pain, the peripheral signal does not travel to the central nervous system (Melzack and Wall 1965: 976). And my version of the Bodily Theory says that the activation of nociceptors located peripherally must cause an experience of pain (plausibly located centrally) in order to be pain, so on this view, people distracted from injuries who deny having pain have no pain. In fact, my version of the Bodily Theory entails SUBJECTIVITY and the Experiential Theory does not, so only if common sense *denied* SUBJECTIVITY would an appeal to common sense indicate an advantage of the Experiential Theory, contrary to what Tye suggests.

Summing up, while the evidence suggests that neither PRIVACY nor INCORRIGIBILITY are part of common sense, subjects are ambivalent about SUBJECTIVITY. But my version of the Bodily Theory is not only compatible with SUBJECTIVITY, it entails it. Therefore, the appeal to PRIVACY, INCORRIGIBILITY and SUBJECTIVITY indicates no advantage of the Experiential Theory over the Bodily Theory.

### 5.2. *The Phantom Limb Argument for the Experiential Theory*

The second argument in favour of the Experiential Theory is that it provides the best account of phantom limb pain (Hill 2017: 67, Tye 2017: 480).<sup>54</sup> The idea is that since there is no real limb, but given INCORRIGIBILITY, still a real pain, the pain cannot be in the limb. There cannot be a real pain in a non-real limb. Hence, the pain must be an experience in the mind or brain.

This argument is unconvincing, because there is no reason to think that the experientialist account of phantom limb pain is the best account. In §3.1. I argued that the pain can be located in the residual limb or stump, where signals from nociceptors are transmitted to the spinal cord. Contrary to what experientialists claim then, it is not true that the pain *must* be in the mind or brain. The Bodily Theory can accommodate phantom limb pain as pain with an illusory location. Thus, the Experiential Theory does not provide the best account of phantom limb pain.

### 5.3. *The Scientific Argument for the Experiential Theory*

The third argument in favour of the Experiential Theory is not explicitly stated as an argument but is something Tye (2005a) gestures towards – namely, that the Experiential Theory accords with pain science, on the ground that the International Association for the Study of Pain (IASP) defines pain as ‘An unpleasant sensory and emotional experience [a] associated with actual or potential tissue damage, or [b] described in terms of such damage’ (IASP 1979, cited by Tye 2005a: 100).

This is unconvincing taken as an argument, for two reasons. Firstly, the appeal to the IASP definition is an appeal to authority, not an appeal to evidence. It is an appeal to the opinion of scientists, not their reasons for having this opinion. It may of course be argued that a view that coheres with the opinion of scientists is more plausible than an alternative view that does not cohere with such opinion, unless there are independently good reasons to adopt the alternative view. But in this case, *there are* independently good reasons to adopt the alternative view (i.e. the Bodily Theory), as argued in §4. Secondly, the note immediately following the definition says that ‘[i]t [i.e. pain] is unquestionably a sensation in a part or parts of the body’ (IASP 1979), which seems to fit better with the Bodily Theory than with the Experiential Theory. It is therefore not clear that the Experiential Theory agrees better with the IASP

---

<sup>54</sup> As noted above, Hill’s (2017) view of pain is disjunctive: chronic pains and phantom limb pains are experiences in the brain, but acute pains are bodily occurrences.

definition than the Bodily Theory does, because it is not clear what pains are according to this definition. Perhaps a charitable interpretation requires one to replace ‘in’ with ‘of’ in the mentioned note, such that pain is an experience/sensation *of* a body part rather than *in* a body part. But with this change, it is not clear that the IASP definition is about pain in the locatable sense rather than pain in the experiential sense. Regardless how the note is interpreted, the appeal to the IASP definition provides no support for the Experiential Theory.<sup>55</sup>

## 6. Conclusion

All pains – acute pains, chronic/neuropathic pains, phantom limb pains, and referred pains – involve activations of nociceptors, which are bodily occurrences located in the body. Given that there are (I have argued) no convincing reasons to adopt the Experiential Theory, and that there are two reasons to adopt the Bodily Theory, I hold that pains in the locatable sense most plausibly *are* these bodily occurrences located in the body, and not experiences or mental states located in the mind or brain. This claim is perfectly consistent with the claim that pain in a different (experiential) sense of the word – i.e. the experience of pain – is an experience in the mind or brain.

The Bodily Theory does not attempt to answer all questions with regard to pain, it only aims to answer the metaphysical question ‘What are pains?’. There are several interesting questions that remain. For example, there are questions about the *quality of pain*, such as the question of what makes a pain burning rather than throbbing or aching. There are also questions about the *experience of pain* – i.e. pain in the experiential sense – including questions about phenomenal character (pain in the phenomenal sense), content (representational or imperative), and how the experience of pain motivates action. These questions about the quality of pain and the experience of pain are all interesting questions. But they are distinct from the metaphysical question to which the Bodily Theory is an answer. That said, the Bodily Theory is consistent with different accounts of the quality and experience of pain, and there is no direct inference available from claims about the metaphysics of pain to claims about the quality and experience of pain (nor is there a converse inference available). So defenders of the Bodily Theory are free to choose their favourite theory of the quality and experience of pain to complement the Bodily Theory.

---

<sup>55</sup> There is also a fourth argument for the Experiential Theory, having to do with the so-called pain-in-mouth argument. I discuss this at length in the next chapter and argue that it provides no reason to prefer the Experiential Theory.

## Chapter 6

### The Bodily Theory and the Pain-in-Mouth Argument<sup>56</sup>

**Abstract:** The so-called pain-in-mouth argument is the step from (1) There is a pain in my finger, and (2) My finger is in my mouth, to (3) There is a pain in my mouth. According to Tye (1995a, 1995b) and Carruthers (2000), the Experiential Theory of pain provides the best explanation of why the inference to (3) sounds wrong. This chapter is about what a defender of the *Bodily Theory* should say about (1)-(3). I argue that while both (one version of) the predicative view (Noordhof 2001, 2002, 2005, Hyman 2003, Bain 2007, Liu 2020) and the implicature view (Reuter et al. 2019, Casser and Schiller 2021) entail the Bodily Theory, they are not satisfactory accounts of the pain-in-mouth argument. To improve upon this, I offer a novel account of the pain-in-mouth argument – the mereological view – that both entails the Bodily Theory, and either solves or avoids the difficulties facing the predicative view and the implicature view.

#### 1. Introduction

In the previous chapter, I defended the Bodily Theory of pain, according to which pains (in the locatable sense) are bodily occurrences located in extra-cranial body parts. First, I developed what I think is the best version of the theory. On my view, pains are proximal activations of nociceptors that cause experiences of pain. Nociceptors are neurons located in the peripheral nervous system, so if pains are activations of nociceptors, then they are occurrences located in extra-cranial body parts. Then I argued that there are two reasons to adopt the Bodily Theory and no good reasons to adopt the Experiential Theory, according to which pains are experiences or intentional mental states located in the mind or brain. But there was one argument for the Experiential Theory I did not address, namely that of Tye (1995a, 1995b) and Carruthers (2000).

According to Tye and Carruthers, the Experiential Theory offers the best account of why the following ‘pain-in-mouth argument’ sounds wrong:

(1) There is a pain in my finger

---

<sup>56</sup> This chapter is based on an unpublished manuscript co-written by Arif Ahmed and me.

- (2) My finger is in my mouth
- (3) Therefore, there is a pain in my mouth

It might be natural to think that if A is in B and B is in C, then A is in C. But there is clearly something wrong or misleading in (1)-(3), which, according to Tye and Carruthers, is best explained by the Experiential Theory of pain (more on this below). Does Tye and Carruthers' argument show that one should adopt the Experiential Theory after all?

In this chapter, I shall not attempt to evaluate or criticize Tye and Carruthers' account of the pain-in-mouth argument. Rather, I shall seek an alternative account that entails the Bodily Theory. In the literature, there are two accounts that both entail the Bodily Theory: (one version of) the predicative view (Noordhof 2001, 2002, 2005, Hyman 2003, Bain 2007, Liu 2020) and the implicature view (Reuter et al. 2019, Casser and Schiller 2021). One might therefore think that defenders of the Bodily Theory should appeal to either of these views in explaining why the pain-in-mouth argument sounds wrong.

However, I argue that neither the predicative view nor the implicature view are satisfactory accounts of (1)-(3). So unless there is an alternative account that entails the Bodily Theory, the Experiential Theory seems to have an advantage over the Bodily Theory in that it accounts for the failure of (1)-(3). The aim of this chapter is therefore to offer a novel account of (1)-(3) that both entails the Bodily Theory and avoids or solves the challenges I identify for the predicative view and the implicature view.

Before I proceed, two clarifications are in order. Firstly, the only *explanandum* with which this chapter is concerned is the pain-in-mouth argument and why it sounds wrong. In the literature on the pain-in-mouth argument, however, there sometimes seems to be unclarity about this. The reason is that most of the above-mentioned philosophers compare (1)-(3) with other arguments that are also wrong or misleading in some way and believe that the fact that a particular account of (1)-(3) can or cannot be extended to analogous arguments is a *desideratum* in evaluating that particular account of (1)-(3). In this context, Tye compares (1)-(3) with: 'I want to be in City Hall' – 'City Hall is in the ghetto' – 'Therefore, I want to be in the ghetto' (1995a: 226-228, 1995b: 111). Carruthers compares (1)-(3) with: 'I want some nail-varnish on my finger' – 'My finger is in my mouth' – 'Therefore, I want some nail-varnish in my mouth' (2000: 120). Noordhof compares (1)-(3) with: 'There is a hole in my shoe' – 'The shoe is in the box' – 'Therefore, there is a hole in the box' (2001: 96-97). Hyman compares (1)-(3) with: 'I have a crease in my shirt' – 'My shirt is in my chest of drawers' – 'Therefore, I have a crease

in my chest of drawers' (2003: 21). Bain compares (1)-(3) with 'Amy's finger twitches' – 'Amy's finger is in her mouth' – 'Therefore, Amy's mouth twitches' (2007: 198). And Reuter et al. compare (1)-(3) with: 'There is tissue damage [or: inflammation] in my finger' – 'My finger is in my mouth' – 'Therefore, there is tissue damage [or: inflammation] in my mouth' (2019: 74, 76).

These other arguments are all interesting. But an account of (1)-(3) need not explain why *all* of these other arguments sound wrong. It is plausible that each argument requires a different explanation, as none of the authors mentioned above provide one explanation that can be extended to all of these, they only provide explanations for the pain-in-mouth argument and their preferred analogy. In other words, everyone can explain the pain-in-mouth argument and their preferred analogy, but no one can explain everyone else's preferred analogies. This means that the fact that an explanation for the pain-in-mouth argument can or cannot be extended to a particular analogous argument is not a decisive *desideratum* for or against that explanation of the pain-in-mouth argument. Thus, the *explanandum* with which this chapter is concerned is only the pain-in-mouth argument and why it sounds wrong.

Secondly, the reason I seek an account that *entails* the Bodily Theory is that I think this theory is true and am wondering whether the competing semantics (i.e. truth-conditions) of pain reports offered by the Experiential Theory and the Bodily Theory respectively shed equal light on why we accept and reject certain inferential connections, like that from (1) and (2) to (3). Tye and Carruthers' claim is that the best explanation of why the pain-in-mouth sounds wrong appeals to the semantics of pain reports offered by the Experiential Theory. Given that I defend the Bodily Theory, the question is therefore whether one can appeal to the semantics offered by the Bodily Theory instead, i.e. whether there is an account of why (1)-(3) fails that entails the Bodily Theory.

The structure of the chapter is this. §2 outlines the views of the authors mentioned above. §3 develops a novel view: the mereological view. §§4-5 compare the mereological view with the predicative and implicature views and argue that the mereological view either solves or avoids the problems facing these views. §6 concludes.

## **2. Existing Views of the Pain-in-Mouth Argument**

This section outlines the dominant views of the pain-in-mouth argument in the literature: the representational view, the predicative view, and the implicature view.



### 2.1. *The Representational View*

According to the Experiential Theory of pain, pains are experiences with intentional content. Tye (1995a, 1995b) and Carruthers (2000) adopt a common version of this theory, namely representationalism, on which pains have representational content and represent bodily disturbances. They suggest that, given representationalism about pain, ‘pain’ creates an intensional context. Thus, in Tye’s example, (1)-(3) is analogous to:

- (4) I want to be in City Hall
- (5) City Hall is in the ghetto
- (6) Therefore, I want to be in the ghetto

Tye and Carruthers hold that ‘want’ creates an intensional context for ‘in’, so that (6) does not follow from (4) and (5). Even if City Hall is in the ghetto, it does not follow that I want to be in the ghetto when I want to be in City Hall. Similarly (they claim), (3) does not follow from (1) and (2), because ‘pain’ creates an intensional context. From the premises that there is a pain representing some bodily disturbance in my finger, and that my finger is in my mouth, it does not follow that there is a pain representing some bodily disturbance in my mouth. This invalidity explains why the deduction of (3) from (1) and (2) sounds wrong (Tye 1995a: 226-228, 1995b: 111-112, Carruthers 2000: 120). A similar diagnosis can be generated from the other version of the Experiential Theory of pain, i.e. imperativism.<sup>57</sup>

If this representational view were the only explanation of why (1)-(3) sounds wrong, then the Experiential Theory would have an advantage over the Bodily Theory of pain, because the representational view is compatible with the Experiential Theory but incompatible with the Bodily Theory. In fact, the representational view entails the Experiential Theory by entailing that pains are representational states, i.e. mental states or experiences, which are typically taken to be located in the mind or brain. But if pains are mental states or experiences located in the mind or brain, then they are not bodily occurrences located in extra-cranial body parts, so the representational view is incompatible with the Bodily Theory of pain.

How should defenders of the Bodily Theory respond to this challenge? In what follows, I shall not attempt to evaluate or criticize the representational view. Rather, I shall seek an

---

<sup>57</sup> According to the standard version of imperativism, pains have imperative rather than representational content – they convey a command to protect a body part (Klein 2015). An imperativist could argue that from the premises that I should protect my finger, and that my finger is in my mouth, it does not follow that I should protect my mouth. Therefore, (1)-(3) is invalid, which explains why the inference to (3) sounds wrong.

alternative account of the pain-in-mouth argument that entails the Bodily Theory. First, I outline two accounts from the literature that both entail the Bodily Theory of pain: (one version of) the predicative view (Noordhof 2001, 2002, 2005, Hyman 2003, Bain 2007, Liu 2020) and the implicature view (Reuter et al. 2019, Casser and Schiller 2021). Then I outline my preferred account – the mereological view – and argue that it is better suited to explain why the pain-in-mouth argument sounds wrong than the predicative and implicature views.

## 2.2. *The Predicative View*

The predicative view comes in two versions: one non-semantic and one semantic. According to the *non-semantic* version, what explains our reaction to the pain-in-mouth argument is the thought connection between ‘There is a pain in X’ and ‘X hurts’. More specifically, the idea is that (3) is taken to entail ‘My mouth hurts’, but this is *not* something (1) and (2) are taken to entail. Given that entailment is taken to be transitive, (1) and (2) are therefore not taken to entail (3), which explains why the pain-in-mouth argument sounds wrong.

The non-semantic version of the predicative view makes no commitments about the semantics (truth-conditions) of the sentences (1)-(3), since it says nothing about what (1)-(3) mean or entail, it only says what (1)-(3) *are taken* to entail. This means that the non-semantic version is compatible with any view of pain and does not entail any of them, in which case it is independent of the question with which I am concerned. Given Tye and Carruthers’ claim that the best account of (1)-(3) appeals to the semantics offered by the Experiential Theory of pain, the question with which I am concerned is not just whether there is a *non-semantic* analysis of the pain-in-mouth argument that is *consistent* with the Bodily Theory. Rather, the question with which I am concerned is whether there is a *semantic* analysis of the pain-in-mouth argument that *entails* the Bodily Theory. In other words, I am seeking an analysis that appeals to the *semantics* of the sentences (1)-(3) offered by the Bodily Theory in order to understand why the corresponding inference sounds wrong, since only this would answer Tye and Carruthers’ challenge. Thus, the non-semantic version of the predicative view is not the sort of explanation we are looking for in the present context.

The *semantic* version of the predicative view says that ‘There is a pain in my finger’ really means (i.e. has the truth-condition) *My finger hurts*; and ‘There is a pain in my mouth’ really means *My mouth hurts*. Given that my finger can hurt while my mouth does not, even when my finger is in my mouth, (1)-(3) is invalid, which explains why the inference to (3) sounds wrong.

In principle, this semantic analysis of (1)-(3) can be taken in two ways. On one interpretation, the claim that ‘There is a pain in my finger’ has the truth-condition *My finger hurts* means that ‘pain’ in sentences like (1) and (3) does not refer. On this view, there are things that hurt in this world, but there are no *pains*, which means that pains (in the locatable sense) do not exist.<sup>58</sup> But if the Bodily Theory is true then pains *do* exist, because the Bodily Theory says that pains are bodily occurrences located in the body, and nothing can be a bodily occurrence located in the body (or anywhere else) while failing to exist. On the above interpretation then, the semantic version of the predicative view is incompatible with the Bodily Theory of pain (as well as any other theory of pain), because of course the Bodily Theory presupposes that pains exist. So given the aim of providing an analysis of (1)-(3) that entails the Bodily Theory, this interpretation of the semantic version of the predicative view makes the predicative view unsuited for present purposes. However, there is a second interpretation of the view.

On a second interpretation of the semantic version of the predicative view, the claim that ‘There is a pain in my finger’ has the truth-condition *My finger hurts* does *not* mean that ‘pain’ in sentences like (1) and (3) does not refer, since what it is for my finger to hurt is that there is a pain in it. Unlike the view that pains do not exist, this view is consistent with the Bodily Theory of pain. In fact, it entails it by entailing that the pain that (1) says is ‘in my finger’ *is* in my finger, which is what the Bodily Theory says. So unlike the versions of the predicative view discussed above, this view *is* the right sort of view we are looking for in the present context. Therefore, in the remainder of this chapter, what I mean by ‘the predicative view’ is this version of the view, understood as a semantic analysis that entails the Bodily Theory of pain.

Although defenders of the predicative view are not explicit about the semantic vs. non-semantic distinction, it is interesting to consider what kind of view the different predicativists are advocating.

Liu (2020) defends the non-semantic version. She writes:

A plausible explanation for the pain-in-mouth puzzle is the following: in English, pain reports using locative locutions are intuitively understood as entailing corresponding predicative locutions – the conclusion of the pain-in-mouth argument, i.e. ‘there is a pain in my mouth’, entails that ‘my mouth hurts’. This consequence should also be

---

<sup>58</sup> See Tye (2002: 151) for a similar point.

entailed by the premises. But it is not. So, the conclusion of the argument does not follow from the premises, and the pain-in-mouth argument is intuitively judged to be invalid (2020: 469).

This account of the pain-in-mouth argument is only about what the sentences (1)-(3) are ‘intuitively understood as entailing’, it is not about the semantics of these sentences. Thus, Liu only says that the pain-in-mouth argument is ‘intuitively judged to be invalid’, which is what the non-semantic version of the predicative view says, not that it *is* invalid, which is what the semantic version says. So Liu is offering a non-semantic explanation.

By contrast, Hyman (2003) defends the semantic version of the predicative view, namely the one that commits to the existence of pains and the Bodily Theory. He writes: ‘the more plausible explanation of why these inferences [of the pain-in-mouth argument and similar] fail [...] is not that ‘in’ does not have a spatial meaning when it is used to locate sensations. It is that sensations are modes, and these inferences are not valid in the case of modes at all’ (2003: 21). Hyman’s claim is that the inference to (3) is actually invalid, because he thinks that even though ‘in’ has a spatial meaning in (1)-(3) and literally locates a pain in my finger, pains are modes (or states) of a body part, and modes (or states) are not ‘inherited’ by my mouth from my finger, even if my finger is in my mouth. Thus, Hyman’s account is semantic and entails the Bodily Theory of pain.

Bain (2007) is more difficult to interpret. On the one hand, Bain claims that the inference to (3) is ‘obviously *invalid*’ (2007: 197), which suggests that he does not take the semantics or truth-conditions of (1) and (2) to validate the inference to (3), and hence that he defends the semantic version of the predicative view. But on the other hand, he claims that the predicative view he favours is completely independent of the semantics of (1)-(3) (2007: 196-198), which suggests that his account of (1)-(3) is supposed to be non-semantic.

Similarly with Noordhof (2001, 2002, 2005). On the one hand, Noordhof says that ‘When I say I have a pain in my mouth or a pain in my finger, I am describing a state of my mouth or the finger, that of being painful or hurting. Because my mouth and finger have spatial locations, the pain has a spatial location too’ (2005: 152, see also 2001: 197, 2002: 154). This suggests that Noordhof offers a semantic explanation, namely the one that entails the Bodily Theory. But on the other hand, he also denies that the predicative view has any ontological implications (2002: 153), which suggests that he takes the view to be non-semantic.

The version of the predicative view with which I am concerned then is Hyman’s view, maybe also Bain’s and Noordhof’s.

### 2.3. *The Implicature View*

According to the implicature view, the pain-in-mouth argument is valid, but the utterance of (3) carries a false conversational implicature, i.e. a false proposition that goes beyond the literal, semantic meaning of (3) (Reuter et al. 2019, Casser and Schiller 2021). On this view, (3) is true when (1) and (2) are true: there *is* a pain in my mouth when I put my injured finger in my mouth, namely the pain in my finger. But uttering (3) misleadingly conveys a false implicature, which explains why the pain-in-mouth argument sounds wrong.

Defenders of the implicature view offer different proposals as to what the false implicature is. According to Reuter et al., the false implicature is that there is something wrong with my mouth (2019: 74). According to Casser and Schiller, the false implicature is that my mouth is the ‘host’ of my pain, i.e. that my mouth is the body part without which my pain could not exist (2021: 11).

In any case, the defenders of the implicature view claim that utterance of (3) would convey their preferred implicature to any listener who interprets me on the assumption that I am being relevant and perspicuous. For it is hard to imagine circumstances in which it would be relevant to say that one has a pain in one’s mouth if one’s injured finger is in one’s non-injured mouth and saying so is hardly a perspicuous way to describe this situation. Utterance of (3) therefore carries a false implicature – either that there is something wrong with my mouth (Reuter et al.), or alternatively that my mouth is the ‘host’ of my pain (Casser and Schiller) – which is not carried by (1) and (2). This explains why the inference to (3) sounds wrong, even though it is (according to defenders of the implicature view) actually correct (Reuter et al. 2019: 74, Casser and Schiller 2021: 11).

Unlike the representational view but like (the relevant version of) the predicative view, the implicature view entails the Bodily Theory of pain. On this view, the semantics (truth-condition) of (1) is that there is a pain spatially located in my finger, the semantics of (2) is that my finger is spatially located in my mouth, and the semantics of (3) is that there is a pain spatially located in my mouth (namely the pain in my finger). Defenders of the implicature view claim that (1)-(3) does not involve different senses of ‘in’ (Reuter et al. 2019: 73-74, Casser and Schiller 2021: 9), which is why, on their view, (1)-(3) is valid and the diagnosis of the pain-in-mouth argument does not appeal to invalidity but rather a conversational implicature. But if the semantics of (1)-(3) are such that the corresponding inference is valid, namely those according to which (1) says that there is a pain spatially in my finger, and so on, then the

implicature view entails the Bodily Theory, since this is exactly what the Bodily Theory says, namely that the pain that (1) says is ‘in my finger’ really *is* in my finger.<sup>59</sup>

### 3. The Mereological View

The predicative view and the implicature view are the only existing accounts of the pain-in-mouth argument that entail the Bodily Theory of pain. But as I argue in §4-5, both these views face difficulties. I therefore think that defenders of the Bodily Theory should adopt an alternative analysis of (1)-(3), which I call the *mereological* view.

According to the mereological view, we can distinguish two spatial senses of ‘X is in Y’:

- X is  $in_1$  Y = X is where some part of Y is
- X is  $in_2$  Y = X is enclosed by a cavity of Y

$in_1$  is the sense in which the raisin is ‘in’ the cake, the brick is ‘in’ the wall, and the oxygen is ‘in’ the atmosphere. The mereological view says that the standard sense in which pains are ‘in’ body parts is  $in_1$ . And since my finger is not co-located with a part of my mouth, the sense in which my finger is ‘in’ my mouth is  $in_2$ . Thus, the pain-in-mouth argument trades on different spatial senses in which ‘X is in Y’: while  $in_1$  is operative in (1),  $in_2$  is operative in (2). This equivocation makes the pain-in-mouth argument invalid. From the premises that there is a pain where a part of my finger is, and that my finger is enclosed by my oral cavity, it does not follow that there is a pain where a part of my mouth is. This explains why the inference to (3) sounds wrong.

Recall that on my version of the Bodily Theory, pains are activations of nociceptors, so on this view, pains are where parts of our bodies are – they are located  $in_1$  body parts. The mereological view therefore entails the Bodily Theory, since it entails that pains are where parts of our bodies are, i.e. located  $in_1$  body parts. I shall now argue that the mereological theory is

---

<sup>59</sup> Even though defenders of the implicature view are not so explicit about why the implicature view entails the Bodily Theory, this is what Reuter et al. think. They say: ‘[...] if our proposal is correct, then this paper provides support for a bodily conception of pain. Not only can the pain-in-mouth argument no longer be used to support a representationalist view of pain, the results suggest that pain is a state of the body, not of the mind. The semantics of reports like ‘there is a pain in my fingertip’, taken at face value, tells us that pains are located in body parts’ (2019: 81). (Casser and Schiller are not so clear about the issue, see 2021: 15.)

better suited to explain why (1)-(3) sounds wrong than the predicative view and the implicature view.<sup>60</sup>

#### 4. The Mereological View vs. The Predicative View

The problem with the predicative view (i.e. the semantic version of the view that entails the Bodily Theory) is that it is at best unclear whether it actually explains why (1)-(3) sounds wrong. Recall that on this view, ‘There is a pain in my finger’ means (i.e. has the truth-condition) *My finger hurts* and that what it is for my finger to hurt is that there is a pain in it. This analysis still depends on the sentences (1) and (3). For insofar as this view is supposed to entail the Bodily Theory, we need to appeal to (1) and (3) in order to explain the meaning of ‘My finger hurts’ and ‘My mouth hurts’. That is, what the predicativist *means* by ‘My finger hurts’ is that *There is a pain in my finger*, what she *means* by ‘My finger is in my mouth’ is that *My finger is in my mouth*, and what she *means* by ‘My mouth hurts’ is that *There is a pain in my mouth*. Thus, we seem to be back where we started with the pain-in-mouth argument, in which case it is unclear whether (this version of) the predicative view explains the target phenomenon.

The predicativist might insist that all that is required of an explanation of the pain-in-mouth argument that appeals to invalidity is that it appeals to some semantics of (1)-(3) that makes it invalid. If the predicativist is right that (a) the semantics of (1) is that my finger hurts and the semantics of (3) is that my mouth hurts, and (b) it does not follow that my mouth hurts when my finger is in my mouth and my finger hurts, then this *does* explain why the inference to (3) sounds wrong.

But the point is that an analysis of (1)-(3) that appeals to invalidity is arguably supposed to *reduce* (1)-(3) to an invalid argument, i.e. render (1)-(3) to sentences whose semantic interpretation does *not* appeal to (1)-(3) and that make up an invalid argument. So if one needs to appeal to (1)-(3) to give a semantic interpretation of ‘My finger hurts’ – ‘My finger is in my mouth’ – ‘Therefore, my mouth hurts’ – which the predicativist needs at some point if she is to

---

<sup>60</sup> The mereological view has affinities with Tye’s (2002) view of (1)-(3), which also appeals to a distinction between senses of ‘in’. Tye writes: ‘Where there is a hollow physical object, *O*, the claim that something *X* is in *O* can be understood either to assert that *X* is within the cavity bounded by *O* or to assert that *X* is (at least partially) embedded within a portion of the cavity-surround (the top, bottom, and sides of *O*)’ (2002: 151-2). One difference between Tye’s account and the mereological view is that as far as I understand, Tye’s distinction only applies to *hollow* objects. A second, more important difference, is that Tye’s account entails the representational version of the Experiential Theory of pain. For instance, (1) is not analyzed as saying that there is anything in my finger, but rather as saying that my pain represents a bodily disturbance in my finger.

offer an analysis that entails the Bodily Theory of pain – then the predicativist has *not* reduced (1)-(3) to ‘My finger hurts’, and so on.

Of course, it may be *true* both that what it is to have a pain in my finger is for my finger to hurt and *vice versa*, and that it does not follow that my mouth hurts even if my finger hurts and my finger is in my mouth, which is what the predicativist under discussion claims. Similarly, it may also be *true* both that what it is to have a pain in my finger is for the world to be in state 1 and *vice versa*, and that it does not follow that the world is in state 2 from the claim that the world is in state 1 and that my finger is in my mouth. The predicative view and this ‘state-of-the-world’ view both appeal to a set of sentences that make up an invalid argument and which they say mean the same as (1)-(3). But insofar as they are supposed to entail the Bodily Theory of pain, (1) is needed to explain what they mean by ‘My finger hurts’ and ‘The world is in state 1’ respectively, and (3) is needed to explain what they mean by ‘My mouth hurts’ and ‘The world is in state 2’ respectively. Therefore, these analyses do not *reduce* (1)-(3) to an alternative set of sentences. Rather, once it is made clear what they mean by ‘My finger hurts’ etc. and ‘The world is in state 1’ etc. we seem to be back where we started with (1)-(3), in which case it is unclear whether the target phenomenon is explained.

Contrast this with the mereological view, which reduces (1)-(3) to ‘There is a pain where a part of my finger is’ – ‘My finger is enclosed by my oral cavity’ – ‘Therefore, there is a pain where a part of my mouth is’. One may ask what it is for a pain to be where a part of my finger is, but unlike the predicativist who wants to offer an account that entails the Bodily Theory, the defender of the mereological view never needs to appeal to (1) in order to explain what ‘There is a pain where a part of my finger is’ means. Rather, the defender of the mereological view can appeal to e.g. my version of the Bodily Theory and say that pains are activations of nociceptors and that nociceptors are located in the peripheral nervous system that is inside and part of my finger. This analysis does not appeal to (1) and (3) and therefore reduces (1)-(3) to ‘There is a pain where a part of my finger is’, and so on.<sup>61</sup> Thus, the mereological view is better suited to explain why (1)-(3) sounds wrong than the predicative view.

---

<sup>61</sup> Similar points apply to the other view that appeals to invalidity, namely the representational view, according to which (1)-(3) is reduced to ‘There is a pain representing some bodily disturbance in my finger’, and so on.



## 5. The Mereological View vs. The Implicature View

The previous section argued that a defender of the Bodily Theory of pain should not adopt the predicative view, since it is unclear to what extent this view explains what is wrong with (1)-(3). The other alternative in the literature that entails the Bodily Theory is the implicature view. The main problem with this view is that it is at best unclear whether the supposed implicature of (3) – either that something is wrong with my mouth (Reuter et al. 2019), or that my mouth is the ‘host’ of my pain (Casser and Schiller 2021) – passes the standard test for being an implicature, namely cancellability.

Conversational implicatures are cancellable in the sense that one can felicitously<sup>62</sup> both assert the sentence that carries the implicature *p* and cancel the implicature by adding something like ‘But not *p*’ or ‘But I do not mean to imply that *p*’ (Grice 1975 [1989]: 44). We can therefore use lack of cancellability as a useful criterion for disconfirming conversational implicatures: if a proposition conveyed is not cancellable, then it is not a conversational implicature (Zakkou 2018: 2). Consider Grice’s famous example, in which S is asked whether K is a good philosopher and S then answers: ‘He’s got good handwriting’. Does this answer carry the implicature that K is not a good philosopher? It depends on whether that proposition is cancellable. And given the felicity of ‘He’s got good handwriting, but I don’t mean that he is not a good philosopher’, the proposition that K is not a good philosopher *is* cancellable, in which case S’s answer may well carry the implicature that K is not a good philosopher.

The question for the implicature view then, is whether one can cancel the alleged implicature of (3), i.e. that there is something wrong with my mouth (Reuter et al), or alternatively that my mouth is the ‘host’ of my pain (Casser and Schiller). The question is not just whether one can *say* something like ‘There is a pain in my mouth, but nothing is wrong with my mouth’, or alternatively ‘There is a pain in my mouth, but my mouth is not the ‘host’ of my pain’. For of course, one can say whatever one wants. The question is whether one can say it *felicitously*. But intuitions may vary as to whether this is a felicitous utterance.

What would be helpful in this connection would therefore be more information about people’s intuitions with regard to the utterance in question. Fortunately, Reuter et al. conducted a study designed to find this out, where they collected the responses of 244 subjects who were exposed to *either* an argument of the form of (1)-(3) *or* an argument in which (3) had been replaced by something like:

---

<sup>62</sup> An utterance is felicitous if there is nothing wrong or strange with it.

- (7) Although it might sound misleading to say there is a pain in my mouth, technically speaking there is a pain located in my mouth

Subjects were then asked how much they agreed or disagreed with the conclusion, i.e. (3) or (7), answering on a seven-point Likert scale anchored at 1 with ‘Strongly Disagree’, at 4 with ‘Neither Agree nor Disagree’ and at 7 with ‘Strongly Agree’.

If (3) carries a false conversational implicature – regardless of whether that of Reuter et al. or that of Casser and Schiller – then (7) should cancel it, since (3) is contained as the final part of (7) and what comes before that final part – ‘Although it might sound misleading to say [...] technically speaking [...]’ – serves to cancel any misleading implicature of (3) by telling the listener to focus exclusively on the semantics (i.e. truth-condition) of (3). After all, according to the implicature view, (3) is straightforwardly true given (1) and (2), so cancellation of any misleading implicature ought to make (3) straightforwardly acceptable. In other words, if either the proposition that there is something wrong with my mouth, or the proposition that my mouth is the ‘host’ of my pain, is an *implicature* of (3), then one would expect the average subject to *agree* with (7).

As a comparison, suppose you said the following:

- (8) Last night Mary brushed her teeth at 11pm  
(9) Last night Mary went to bed at 11.15pm  
(10) Therefore, last night Mary went to bed and brushed her teeth

In conversation one may hesitate to accept (10), and it is plausible that the cause of this hesitation is a false implicature, namely that Mary went to bed and brushed her teeth in that order. In confirmation of this, suppose that instead of (10) you had said:

- (11) Although it might sound misleading to say that Mary went to bed and brushed her teeth last night, strictly speaking she did both go to bed and brush her teeth

Given (8) and (9), there should be no hesitation to accept (11), since (11) plausibly cancels the false implicature of (10) as predicted by an ‘implicature view’ about (8)-(10). Similarly, given the implicature view with which I am concerned, one should expect that (7) successfully cancels the alleged false implicature of (3) and thus that subjects agree with (7).

But this is not what Reuter et al. found. The subjects in their study were on average *neutral* about (7), its mean rating being 3.79. This score is closer to ‘disagree’ than to ‘agree’, and it is closest to ‘neither agree nor disagree’. In fact, the *modal* rating for (7) in the study was 1, the lowest possible (= ‘strongly disagree’).<sup>63</sup> This suggests that subjects on average did not find (7) felicitous, in which case (7) is not taken to cancel any misleading implicature. Given the cancellability test then, one may doubt that utterance of (3) conveys a false implicature, at least in lack of an explanation for why (7) did not succeed in cancelling this alleged implicature.

In response to this worry, Reuter et al. emphasize that the mean rating for (7), namely 3.79, was significantly higher than the rating for (3), namely 2.10 (2019: 80). It is true that the implicature view predicts a higher rating for (7) than for (3). But this is because it predicts *agreement* with (7), not just a higher rating, so pointing out that the rating for (7) was higher than for (3) does not dispel the above worry. Reuter et al. also suggest that the low rating for (7) ‘likely reflect[s] that our manipulation did not fully cancel the pragmatic implicature for all participants and/or that some participants did not recognize the technical sense of the terms suggested by the prompts’ (2019: 80-81). But I do not believe that this is a *likely* explanation of the low rating for (7), for two reasons.

Firstly, Reuter et al.’s subjects were all competent English-speakers, so it is not clear why they would misunderstand the ‘sense of the terms suggested by the prompts’ in (7). At least, this would require some explanation, which Reuter et al. do not provide.

Secondly, it is not clear what Reuter et al. mean by ‘the *technical* sense of the terms suggested by the prompts’ (my emphasis). On their own view, there *is* no technical sense of e.g. ‘located in my mouth’, if by ‘technical’ they mean something that is unusual or difficult to understand. On the contrary, the sense in which there is a pain in my mouth (on their view), is the same sense in which my finger is in my mouth, which is why, on the implicature view, (1)-(3) is valid. Given that there is nothing unusual or difficult to understand about the sense in which my finger is ‘located in my mouth’ there can be nothing unusual or difficult to understand about the sense in which there is a pain ‘located in my mouth’ either.

Besides, if Reuter et al. were to appeal to some other, genuinely ‘technical’ sense in which there is a pain located in my mouth given (1) and (2), then that must be a different sense from the ‘non-technical’ sense in which there is a pain located in my finger. But if the pain-in-mouth argument trades on different senses in which there can be a pain in a body part, then that equivocation would make (1)-(3) invalid, contrary to what Reuter et al. claim. If so, it would

---

<sup>63</sup> I am grateful to Kevin Reuter and colleagues for sharing the data from which this information is taken.

not be a misleading implicature that explains why the inference to (3) sounds wrong, since the activation of an implicature depends on my listener having grasped the semantics of (1)-(3). But if the semantics of (1)-(3) are such that the corresponding inference is invalid and my listener understands these semantics, then *that* is what explains why the inference sounds wrong – along the lines of the mereological view – and the appeal to implicature would be unnecessary and unmotivated. Consequently, Reuter et al.’s explanation of the low rating for (7) – that ‘some participants did not recognize the technical sense of the terms suggested by the prompts’ (2019: 81) – is not a plausible explanation, and thus it is at best unclear that utterance of (3) carries a false implicature.

The mereological view does not face this problem, because it makes no specific prediction about what response one would expect to (7). That said, there is a natural explanation of the data that is consistent with the mereological view. Recall that this view makes a distinction between two spatial senses of ‘X is in Y’:

- X is in<sub>1</sub> Y = X is where some part of Y is
- X is in<sub>2</sub> Y = X is enclosed by a cavity of Y

If one has sympathies with the Bodily Theory of pain, it is natural to interpret ‘in’ as in<sub>1</sub> for (7).<sup>64</sup> Therefore, subjects who adopt this natural reading of ‘in’ will continue to disagree with (7), regardless of the qualifying formulation ‘Although it might sound misleading to say [...] technically speaking [...]’. However, this qualifying formulation *might* trigger other subjects to adopt a more charitable but less natural interpretation of ‘in’ as in<sub>2</sub>. On this second reading, (7) is true if (1) and (2) are true, because, assuming the Bodily Theory of pain, the pain *is* in the oral cavity. Subjects who read it this way will therefore tend to agree with (7). On this explanation of the data, the overall result might well be an average Likert score close to ‘neutral’, but with a relatively high variance to reflect the prediction that many subjects will either strongly agree with (7) or strongly disagree with it. This is indeed what the data show.<sup>65</sup>

Note that this explanation of the low rating for (7) is only available for defenders of the mereological view, it is *not* available for defenders of the implicature view. The reason is that, as mentioned above, if the pain-in-mouth argument trades on different senses in which there

---

<sup>64</sup> For empirical evidence that people think that pains are located in body parts rather than the mind, see Sytsma (2010) and Kim et al. (2016) (discussed in the previous chapter).

<sup>65</sup> As mentioned above, the mean Likert rating for (7) is 3.79. The population variance is 5.08, which is much higher than for the original version of (1)-(3), where the variance was only 1.95. Again, I am grateful to Kevin Reuter and colleagues for sharing the data from which these statistics are taken.

can be a pain in a body part, then (1)-(3) is invalid, contrary to what the implicature view says, but in agreement with what the mereological view says. In any case, the low rating for (7) is a problem for the implicature view and not for the mereological view, so a defender of the Bodily Theory of pain should prefer the mereological view as a diagnosis of the pain-in-mouth argument.

To sum up then, the implicature view predicts that the alleged implicature of (3) – regardless of whether that of Reuter et al. or that of Casser and Schiller – is cancellable. But given the data uncovered in Reuter et al.’s study, it is at best unclear whether it is. The mereological view does not face this problem. Thus, defenders of the Bodily Theory of pain should prefer the mereological view as an explanation of why (1)-(3) sounds wrong.

## **6. Conclusion**

What and where are pains? The Bodily Theory and the Experiential Theory offer competing answers. In the previous chapter, I defended the Bodily Theory, but I did not consider Tye’s (1995a, 1995b) and Carruthers’ (2000) argument that the Experiential Theory offers the best explanation of why the pain-in-mouth argument sounds wrong. This chapter has addressed this argument by considering alternative accounts of the pain-in-mouth argument. I have argued that defenders of the Bodily Theory should adopt the mereological view, according to which (1)-(3) trades on different senses of ‘in’, which makes the pain-in-mouth argument invalid and thereby explains why the deduction of (3) from (1) and (2) sounds wrong. I therefore conclude that Tye and Carruthers’ analysis of (1)-(3) provides no reason to prefer the Experiential Theory, so given the arguments of the previous chapter, one should still believe in the Bodily Theory of pain.

This ends the discussion about pain. The following chapter concerns a debate related to a different topic, namely the extension of psychological predicates, which is the final first-order topic I address in this thesis.

## Chapter 7

### Does the Brain Think?

**Abstract:** It is common in cognitive science to ascribe psychological predicates to the brain, i.e. to assert that the brain sees, feels, thinks, etc. This has prompted philosophical debate. According to the Nonsense View, the relevant locutions of cognitive scientists are nonsensical or false (Bennett and Hacker 2003, 2007). According to the Literal View, they are literal truths and report the psychological properties of brains (Dennett 2007, Crane 2015, Figdor 2018). In this chapter, I propose the Synecdoche View, according to which cognitive scientists' locutions are figurative, with 'brain' referring to the human being, such that 'The brain thinks' reports the thinking of the human being, not the thinking of the brain. I compare this view to the dominant views in the literature and argue that it is a plausible alternative. I argue that this has two important consequences. The semantic consequence is that there is no reason to take the controversial locutions of cognitive scientists literally. The metaphysical consequence is that there is no reason to believe that the controversial locutions indicate empirical support for the claim that brains possess psychological properties, or that scientists have discovered that psychological predicates have a wider extension than what is often assumed.

#### 1. Introduction

Cognitive neuroscientists and psychologists often ascribe psychological predicates to the brain (or its parts). Here are some examples discussed by philosophers:

[...] the minor hemisphere is indeed a *conscious* system in its own right, *perceiving, thinking, remembering, reasoning, willing, and emoting* [...] both the left and the right hemisphere may be *conscious* simultaneously in different [...] mental experiences that run along in parallel. (Sperry 1974: 11, my emphasis, quoted by Bennett and Hacker 2003: 78, 389-390).

The left hemisphere [...] with its capacity for *making inferences and interpretations*, was more strongly influenced by the expectations for action common to a scene and

falsely *recognized* pictures consistent with the observed scene (Gazzaniga 1995: 225-226, my emphasis, quoted by Bennett and Hacker 2003: 390n16).

[T]he light of the Sun always comes from above. This means that concave objects will be light at the top and dark at the bottom, while convex objects will be light at the bottom and dark at the top. Our brain has this simple rule built into its wiring. It uses this rule to *decide* whether an object is concave or convex [...] (Frith 2007: 128, my emphasis, quoted by Crane 2015: 253).

If we are to take these locutions at face value, the brain (or its parts) is conscious, perceives, thinks, remembers, reasons, wills, emotes, recognizes, makes inferences and interpretations, recognizes, and decides.

Despite the widespread practice of cognitive scientists to ascribe psychological predicates to the brain, whether and how we can make sense of these locutions has in recent years been a matter of controversy among philosophers. In this debate, Bennett and Hacker defend what I (following Figdor 2018) call *the Nonsense View*, according to which cognitive scientists' ascriptions of psychological predicates to the brain are nonsensical or false. They believe that psychological predicates apply literally only to human beings or other animals as a whole and that brain behaviour – such as activity detected in an fMRI scan – is not a legitimate ground for ascribing psychological predicates to the brain. Against this, Dennett (2007), Crane (2015) and Figdor (2018) defend what I call *the Literal View*, according to which the locutions of cognitive scientists are literal truths and report the psychological properties of brains. On their view, brain behaviour *is* a legitimate ground for ascribing psychological predicates to the brain.<sup>66</sup>

In what follows, I shall propose a novel view that accommodates aspects from both these views. According to this view, which I call *the Synecdoche View*, we can make sense of the scientific practice by understanding the word 'brain' figuratively, namely as a synecdoche (a word for a part referring to the whole, or vice versa) referring to the human being. Just like I can use 'my new set of wheels' to refer to my new car, so cognitive scientists use 'brain' to refer to the human being. On this view, 'The brain thinks' reports the thinking of the human being, not the thinking of the brain. My main thesis is a modest one: that the Synecdoche View is a plausible alternative. I do not make the further claim that it is the best alternative. Still, the

---

<sup>66</sup> As I explain in footnote 74, Figdor's view is actually ambiguous and may be interpreted in a different way.

modest claim has interesting consequences. One consequence is that, contrary to what defenders of both the Nonsense View and the Literal View assume, there is no reason to take the controversial locutions literally. Another consequence is that, contrary to what e.g. Figdor claims (2018: 5-6, 61), there is no reason to believe that these locutions indicate empirical support for the claim that brains possess psychological properties, or that scientists have discovered that psychological predicates extend more widely than what is often assumed. Before laying out the structure of the chapter, three clarifications are needed.

Firstly, the debate with which this chapter is concerned is not about speaker-meaning. Speaker-meaning is (roughly) what a speaker intends to communicate with an utterance, and one can investigate what it is by asking the speaker of the utterance what he or she means. If the debate were about speaker-meaning, then the above-mentioned philosophical theories would be pointless, as one should rather ask the relevant scientists what they intend to mean when they say things like ‘The brain thinks’.<sup>67</sup> But asking scientists what they mean would not be relevant for the debate with which this chapter is concerned, since this debate is about the *semantic* meaning (i.e. truth-conditions) of the controversial locutions, not the speaker-meaning. The semantic meaning of the controversial locutions is less straightforward to investigate than the speaker-meaning, because, as will become clear, it involves considerations about e.g. what we are, the ideal of charitable interpretation, grounds for assertion, etc. Scientists are of course free to contribute to this debate, but that would involve engaging with considerations to be discussed in this chapter, it would be less helpful to unreflectively report speaker-meaning. Analogously, when philosophers discuss whether the language of physical theories refer to unobservable entities, they are not interested in what the scientists formulating those theories intend to communicate (speaker-meaning), but what the truth-conditions of those theories are. Just as it would be minimally interesting for the scientific realism debate to just ask scientists whether they intend to refer to unobservable entities, so it would be minimally interesting for the debate with which this chapter is concerned to just ask scientists what they intend to mean when they ascribe psychological predicates to the brain.

Secondly, the semantic question with which this chapter is concerned must be separated from the metaphysical question of whether the brain possesses psychological properties, i.e. whether it sees, feels, thinks etc. Still, these questions are related, since if the relevant locutions of cognitive scientists should be taken as literal truths, then these locutions indicate empirical

---

<sup>67</sup> However, it is not certain that one would get uniform answers. Cognitive scientists who discuss the relevant locutions disagree about how they should be interpreted (see Smythies 2009 and Bennett 2007), and those who do not discuss this may not have thought much about the issue.



support for the claim that brains possess psychological properties (Figdor 2018: 5-6, 61).<sup>68</sup> At least if we assume the commonly accepted epistemic authority of science. If these locutions should *not* be taken as literal truths, however, then they do not indicate such empirical support. In other words, an answer to the semantic question determines whether these locutions indicate empirical support for a positive answer to the metaphysical question.

Thirdly, even though I follow defenders of the Nonsense View and the Literal View in appealing to the literal/non-figurative vs. non-literal/figurative distinction when formulating the semantic views, these views do not rely on that distinction, and they could be formulated in terms of truth-conditions only. The Nonsense View is the view that the truth-conditions of ‘The brain thinks’ are never satisfied in the relevant cognitive scientific context. The Synecdoche View is the view that ‘The brain thinks’ inherits its truth-conditions from ‘The human being thinks’ (in the relevant context). And the Literal View is the view that the truth-conditions of ‘The brain thinks’ are satisfied independently of whether the truth-conditions for ‘The human being thinks’ are satisfied (in the relevant context). So even though the literal/non-figurative vs. non-literal/figurative distinction is controversial (Sperber and Wilson 2002) that need not affect the debate with which this chapter is concerned, and I shall continue to use terms like ‘literal’ and ‘figurative’.

§2 introduces the Nonsense View and reconstructs Bennett and Hacker’s (2003, 2007) argument for that view, which provides a useful background for the rest of the chapter. §3 introduces the Literal View. §4 identifies five assumptions made by defenders of both the Nonsense View and the Literal View, which is important both in order to understand the debate and to identify alternatives to the Nonsense View and the Literal View. §5 argues that one of these assumptions – that cognitive scientists ascribe psychological predicates to the brain on the basis of brain behaviour and not on the basis of human behaviour – is false. This opens for the Synecdoche View, which I develop in §6. §7 compares the Synecdoche View to the alternative views and argues that while it is preferable to the Nonsense View, the Literal View and the Synecdoche View are on par in the sense that there is no reason to prefer one rather than the other. §8 concludes that the Synecdoche View is a plausible alternative and draws out two consequences.

---

<sup>68</sup> Figdor writes: ‘[M]y metaphysical position – Anti-Exceptionalism – holds that all the relevant scientific evidence shows that psychological capacities are possessed by a far wider range of kinds of entities than often assumed. [...] Literalism claims that, in contexts standardly interpreted as fact-stating, uses of psychological predicates to ascribe capacities to entities in this wider range are best interpreted as literal with sameness of reference. Anti-Exceptionalism is the metaphysical position that underwrites the claim of sameness of reference (2018: 5-6).

## 2. The Nonsense View

According to (the standard formulation of) the Nonsense View, the locutions of cognitive scientists make no sense. This view is defended by Bennett and Hacker (2003, 2007) and their argument for the view is this.<sup>69</sup>

(P1) Only if an entity X can, under normal circumstances, behave in a way that satisfies criteria for the application of psychological predicates, do literal ascriptions of psychological predicates to X make sense.

(P2) It is only whole animals like human beings, not brains, that can, under normal circumstances, behave in a way that satisfies criteria for the application of psychological predicates.

(P3) Cognitive scientists make literal ascriptions of psychological predicates to the brain.

(C) Therefore, cognitive scientists' ascriptions of psychological predicates to the brain do not make sense.

By 'criteria' Bennett and Hacker (2003, 2007) mean the behavioural grounds on which we typically ascribe psychological predicates to animals and human beings. These include crying or limping for the ascription of pain, following moving objects with one's eyes or squinting in strong light for the ascription of seeing, and reaching out for things and trying to get them for the ascription of wanting. Similarly with other psychological predicates (2003: 81-83).

The motivation for (P1) is the claim that criteria are part of the meaning of psychological predicates. This claim is based on (a) a Wittgensteinian view of meaning, according to which the meaning of a word is determined by explanations of meaning, which again are determined by use<sup>70</sup>, and (b) the claim that to specify the criteria for the application of predicates such as 'being in pain', 'seeing', or 'wanting' is to give explanations of what these predicates mean (Hacker 1990: 552). Of course, Bennett and Hacker's claim is not that criteria have to be satisfied by an entity X at time *t* if the ascription of a psychological predicate to X at *t* is to make sense. It obviously makes sense to ascribe wanting to a person at *t* even though the person does not exhibit any behavioural characteristics of wanting at *t*. That is why (P1) says that literal ascriptions of psychological predicates to X only make sense if X *can* behave in a way that

---

<sup>69</sup> Bennett and Hacker do not present a recognizable deductive argument, leading Dennett to claim that their view is held 'without argument' (2007: 86). However, I believe that a charitable interpretation of what they say and some background from Hacker's other publications suggests the argument that I outline here.

<sup>70</sup> See Baker and Hacker (2005: 29-43, 129-158) and Hacker (2013: 123-127) for more on this view.

satisfies criteria for the application of psychological predicates (Bennett and Hacker 2003: 112, 82n35, 2007: 135). The qualification ‘under normal circumstances’ is necessary because there are some people who cannot behave in a way that satisfies criteria but to whom it still makes sense to ascribe psychological predicates, namely paralysed people. A paralysed person would be able to behave in a way that satisfies criteria ‘under normal circumstances’, i.e. if he or she were not paralysed.<sup>71</sup>

The motivation for (P2) is intuition. According to Bennett and Hacker, it is intuitively obvious that brains cannot – under any circumstances – cry, limp, follow moving objects with their eyes, squint in strong light, or reach out for and try to get things (2003: 83, 2007: 141, 149). On their view, it is only human beings (or animals), not brains, that can behave in a way that satisfies criteria (in their sense of the word) for the application of psychological predicates.

The motivation for (P3) is the existence of passages like the ones quoted in §1, where cognitive scientists apparently make literal ascriptions of psychological predicates to the brain (Bennett and Hacker 2003: 68-81). However, given Bennett and Hacker’s belief that cognitive scientists’ ascriptions of psychological predicates to the brain are nonsensical, one may think that they cannot accept (P3) because one may think the ascription of psychological predicates to the brain cannot be both literal and nonsensical. But Bennett and Hacker’s notion of nonsense is not that of unrecognizable gibberish. Rather, what they mean by ‘nonsense’ is a use of words that does not conform to explanations of meaning, which they take to be rules for the use of words. On this view of nonsense, it seems possible that one can ascribe psychological predicates to the brain – i.e. use ‘brain’ as the grammatical subject of psychological verbs – all in a literal and ordinary sense, even though such ascriptions are (according to Bennett and Hacker) nonsensical.

From (P1), (P2) and (P3), Bennett and Hacker conclude that cognitive scientists are committing a ‘mereological fallacy’: they ascribe to a part of an animal predicates that it only makes sense to ascribe to the whole animal (2003: 73, 2007: 132).

This is both how Bennett and Hacker themselves and their opponents (Dennett 2007, Crane 2015, Figdor 2018) understand the Nonsense View. But I think there is an alternative characterization of the position (at least a closely related one) that does not invoke Bennett and

---

<sup>71</sup> Bennett and Hacker do not themselves include the qualifier ‘under normal circumstances’, which made critics point to paralysed people as a counterexample to Bennett and Hacker’s claim that it only makes sense to ascribe psychological predicates to an entity that can behave in a way that satisfies criteria (Searle 2007: 105). But including the qualifier gets Bennett and Hacker around this worry.

Hacker's notion of nonsense. In more standard terminology, their view is really the view that the controversial locutions of cognitive scientists are *false*. And the argument is this:

(P1) Only if an entity X can, under normal circumstances, behave in a way that satisfies criteria for the application of psychological predicates, can the truth-conditions for literal ascriptions of psychological predicates to X be satisfied.

(P2) It is only whole animals like human beings, not brains, that can, under normal circumstances, behave in a way that satisfies criteria for the application of psychological predicates.

(P3) Cognitive scientists make literal ascriptions of psychological predicates to the brain.

(C) Therefore, cognitive scientists' ascriptions of psychological predicates to the brain do not satisfy the truth-conditions for literal ascriptions of psychological predicates to the brain.

For present purposes, we can take these alternative formulations of Bennett and Hacker's view as equivalent. But I shall refer to their position as the 'Nonsense View' and their argument as the 'Nonsense Argument'. In §3 I introduce the Literal View, which denies the conjunction of (P1) and (P2) but accepts (P3). In §6 I develop the Synecdoche View, which is agnostic about (P1) and (P2) but denies (P3).

### 3. The Literal View

According to the Literal View, the controversial locutions of cognitive scientists report literal truths about the psychological properties of brains. There are interesting differences in Dennett's, Crane's and Figdor's different formulations of this view, but the essential claim that matters in the present context is the same, namely that there are behavioural similarities between brains and humans that warrant the extension of psychological predicates to brains (Dennett 2007: 78, Crane 2015: 259-262, Figdor 2018: 96).

What behavioural similarities are there? Defenders of the Literal View are not explicit about this. But Crane (2015) and Figdor (2018) claim that the similarities in question can be captured by mathematical models (Crane 2015: 262, Figdor 2018: 96), which means that the similarities in question need not be what one would intuitively think of as behavioural similarities. Even though the brain may not cry, squint in strong light and try to get things, there may be other 'quantitative' similarities between the brain and the human being that warrant ascribing pain, seeing and wanting to the brain. Consequently, defenders of the Literal View

can maintain that the locutions of cognitive scientists make sense and express truths, which, after all, a charitable interpretation of the linguistic practice of a successful science should aim for.

Even though defenders of the Literal View disagree with the conclusion of the Nonsense Argument, they agree with Bennett and Hacker that cognitive scientists use the word ‘brain’ and psychological predicates with an ordinary literal sense, i.e. they accept (P3). Given this, they must deny either (P1) or (P2) or both.

If they take (P1) to be false, their objection to the Nonsense Argument is that one can ascribe psychological predicates to an entity (a brain) that cannot behave in a way that satisfies behavioural criteria (in Bennett and Hacker’s sense), because it behaves in a way sufficiently similar to something satisfying such criteria. If they take (P2) to be false, their objection is that brain behaviour is sufficiently similar to human behaviour to satisfy behavioural criteria. The difference between these alternatives is just a matter of what one calls ‘criteria’ – i.e. whether one accepts Bennett and Hacker’s restricted notion of criteria or whether one adopts a less restricted notion where behavioural similarity captured in mathematical models counts as satisfying criteria. This is merely a verbal difference and is not important for understanding the Literal View. What *is* important is that defenders of the view reject the conjunction of (P1) and (P2) because they think there is sufficient similarity between brains and human beings.

The fundamental disagreement between the Nonsense View and the Literal View then is about whether brain behaviour is a legitimate ground for ascribing psychological predicates to the brain. Defenders of the Nonsense View hold that brain behaviour is *not* a legitimate ground, while defenders of the Literal View hold that brain behaviour *is* a legitimate ground. In §5, I argue that this dispute about brain behaviour is based on a misunderstanding of the scientific practice, because the locutions of cognitive scientists are not based on the observation of (or a mathematical model of) brain behaviour. This makes the dispute about brain behaviour orthogonal to the semantic question with which the debate is concerned.

#### **4. Five Assumptions**

Before I can argue that the controversial locutions of cognitive scientists are not based on observation of (or a mathematical model of) brain behaviour, I must first make explicit five assumptions implicit in the debate between defenders of the Nonsense View and the Literal View. This is important both in order to understand the debate and to identify alternatives to the Nonsense View and the Literal View.

The first assumption is this:

- (1) We are human beings and not brains.

If defenders of the Nonsense View and the Literal View held that we are brains (like Parfit 2012), then the Nonsense View would be that it makes no sense (or is false) to say that beings like us think, while the Literal View would be that it is literally true to say that beings like us think. But it would have been surprising if the disagreement between these views had been whether beings like us think in lack of any statement that that is what the debate is about. Thus, it is plausible that defenders of both views assume (1).

The second assumption is this:

- (2) We possess psychological properties.

Given (1), it is plausible that defenders of the Nonsense View and the Literal View assume (2). If they denied (2) and held that we do not think, then defenders of the Nonsense View would be committed to the claim that neither we human beings, nor brains, think, and defenders of the Literal View would be committed to the claim that only brains think, while we human beings do not. But it would have been surprising if they had thought that these claims were true without making that explicit in the context of the debate about the controversial locutions of cognitive scientists. Thus, it is plausible that defenders of both views assume (2).

The third assumption is this:

- (3) Cognitive scientists use psychological predicates with an ordinary literal sense.

That defenders of the Nonsense View assume (3) is clear from the fact that this view is motivated by the claim that the brain does not satisfy criteria (in Bennett and Hacker's sense) for the ascription of thinking – in the sense in which human beings think – which is an ordinary literal sense. For it would not matter whether the brain does not satisfy criteria for the ascription of thinking in an ordinary literal sense if the sense of 'thinking' relevant to the locutions of cognitive scientists was not ordinary and literal.

A similar point applies to the Literal View, since this view is motivated by the claim that the brain behaves in a way that is sufficiently similar to the behaviour on the basis of which one ascribes thinking – in an ordinary literal sense – to the human being. For it would not matter

whether the brain behaves in a way that is sufficiently similar to behaviour that grounds the ascription of thinking in an ordinary literal sense if the sense of ‘thinking’ relevant to the locutions of cognitive scientists was not ordinary and literal.

In other words, if, contrary to what (3) says, cognitive scientists used psychological predicates with a figurative or technical sense unfamiliar to non-scientists, then the Nonsense View and the Literal View would have no relevance for the semantics of the controversial locutions of cognitive scientists. Thus, defenders of both views plausibly assume (3).

The fourth assumption is this:

(4) Cognitive scientists use the word ‘brain’ with the ordinary literal sense.

If cognitive scientists used the word ‘brain’ with a figurative or technical sense, then it would not matter that the brain – in the ordinary literal sense – does not satisfy criteria for the ascription of psychological predicates, which is the claim that motivates the Nonsense View. It would also not matter that the brain – in an ordinary literal sense – behaves similarly to the human being, which is the claim that motivates the Literal View. So if, contrary to what (4) says, the defenders of these views held that cognitive scientists do not use ‘brain’ with the ordinary literal sense, then again, neither of these views would have any relevance for the semantics of the locutions of cognitive scientists. Thus, defenders of both views plausibly assume (4).

The fifth assumption is this:

(5) Cognitive scientists ascribe psychological predicates to the brain on the basis of brain behaviour and not on the basis of human behaviour.

Given the dispute about whether brain behaviour is a legitimate ground for ascribing psychological predicates to the brain, it is clear that defenders of both the Nonsense View and the Literal View assume (5). It is also clear from specific passages. For example, Bennett and Hacker say that ‘brain activity detected by PET or fMRI [...] does not show that the brain is thinking’ (2003: 83) and that ‘[t]he constitutive grounds upon which competent speakers of our language apply such [psychological] expressions to animals and human beings, namely, what they say and do, cannot be satisfied by a brain or its parts’ (2007: 149). And Crane (2015) – commenting on ascribing inference to the brain – says that ‘[i]nferences relate propositions, so the claim is that what is going on in the brain resembles a relationship between propositions’

(2015: 262). It is difficult to make sense of these passages without appeal to (5), so defenders of both views plausibly assume (5).

In the remainder of this chapter, I shall follow defenders of the Nonsense View and the Literal View in assuming (1), (2), and (3), and I shall not provide any argument for these assumptions. But in §5 I argue that (5) is false. This makes (4) questionable, which I argue in §§6-7.

### **5. Why Do Cognitive Scientists Say ‘The Brain Thinks’?**

Contrary to the assumption of defenders of the Nonsense View and the Literal View, cognitive scientists *do* ascribe psychological predicates to the brain on the basis of the behaviour of the human being.

Take again the passage from Frith quoted in §1 (and by Crane 2015: 253):

[T]he light of the Sun always comes from above. This means that concave objects will be light at the top and dark at the bottom, while convex objects will be light at the bottom and dark at the top. Our brain has this simple rule built into its wiring. It uses this rule to decide whether an object is concave or convex [...]

The passage continues:

[...] which you can test by looking at the figure below. [...] We interpret the spots as concave and convex because the shading suggests there are shadows caused by light coming from above (Frith 2007: 128).

On the assumption of defenders of the Nonsense View and the Literal View, this should be taken as ascribing a decision to the brain on the basis of (a mathematical model of) brain behaviour and not on the basis of human behaviour. But the above passage suggests otherwise. Frith does not ascribe a decision to the brain on the basis of an fMRI scan or anything like that, but rather on the basis of the behaviour of the whole human being. When Frith says that you can test how the brain decides whether an object is concave or convex by looking at the figures in his book, ‘you’ refers to the reader who (given the assumption that we are human beings) is a human being. So it is the behaviour of human beings, how ‘we’ interpret shadows, which is the ground for asserting that the brain decides whether an object is concave or convex. The



studies Frith draws on are experiments where responses of human subjects serve as the evidence from which conclusions about ‘brain’ decisions are drawn.

The same applies to the quoted passages of Sperry (1974) and Gazzaniga (1995). They both discuss studies of patients who have undergone ‘cerebral commissurotomy’ or ‘split-brain surgery’, i.e. an operation that cuts the nerves between the hemispheres such that signals cannot be sent between them (Sperry 1974: 5). In such cases, subjects can identify objects of perception presented to one visual field, hand, foot or nostril, which correspond to one hemisphere, but they cannot identify the same objects if presented to the other visual field, hand, foot or nostril, which correspond to the other hemisphere. The asymmetries in the cognitive and perceptual abilities of these patients shows that the hemispheres perform differently on the same functional task (Sperry 1974: 7, Gazzaniga 1995: 225). This research is the context for the passage from both Sperry and Gazzaniga.

Here again is the passage from Sperry:

[...] the minor hemisphere is indeed a conscious system in its own right, perceiving, thinking, remembering, reasoning, willing, and emoting [...] both the left and the right hemisphere may be conscious simultaneously in different [...] mental experiences that run along in parallel. (Sperry 1974: 11, quoted in §1 and by Bennett and Hacker 2003: 78, 389-390).

These claims are not based on the observation of (or a mathematical model of) brain behaviour. Rather, they are based on the observation of the behaviour of human beings – the subjects of the studies that Sperry discusses. That is why Sperry also says that:

The presence of two rather separate streams of conscious awareness is manifested in many ways in different kinds of testing situations. For example, following surgery, these *people* are unable to recognize a visual stimulus that they have just looked at as if it is presented across the vertical midline in the opposite half visual field; that is, the normal perceptual transfer that one expects to find between the left and right halves of the field of vision is lacking [...] these *people* are unable to name or to describe verbally objects seen in the left half field of vision [...] (Sperry 1974: 7, my emphasis).

The ‘people’ Sperry discusses are plausibly human beings, and it is their behaviour (including verbal) which serves as the evidence from which conclusions about ‘hemisphere’ psychological properties are drawn.

Here again is the passage from Gazzaniga (1995), who (as noted) also discusses split-brain patients:

The left hemisphere [...] with its capacity for making inferences and interpretations, was more strongly influenced by the expectations for action common to a scene and falsely recognized pictures consistent with the observed scene (Gazzaniga 1995: 225-226, quoted in §1 and by Bennett and Hacker 2003: 390n16).

Again, these claims are not based on the observation of (or a mathematical model of) brain behaviour, but rather on the observation of human behaviour. For as Gazzaniga points out:

Recognition tasks merely require a *subject* to judge whether a stimulus such as a printed word has been seen before on a previously studied list (Gazzaniga 1995: 222, my emphasis).

The ‘subject’ is plausibly a human being, so the ground for which psychological predicates are ascribed to the ‘hemisphere’ in this context is human behaviour, contrary to the assumption of defenders of the Nonsense View and the Literal View. The same is true of other examples discussed in the debate (e.g. Crick 1995: 31, 57, 170, Damasio 1996: 172-173, LeDoux 1998: 69).

The fact that cognitive scientists *do* ascribe psychological predicates to the brain on the basis of the behaviour of the human being, contrary to what defenders of the Nonsense View and the Literal View assume, has two consequences. The first is that the dispute about whether brain behaviour is a legitimate ground for ascribing psychological predicates to the brain – i.e. the dispute about (P1) and (P2) of the Nonsense Argument – is orthogonal to the semantic question with which the debate is concerned. It does not matter for the semantics of the relevant locutions of cognitive scientists whether brain behaviour is a legitimate ground for ascribing psychological predicates to the brain, since these locutions are *not* based on the observation of (or a mathematical model of) brain behaviour. The second consequence is that an alternative semantics of the locutions of cognitive scientists becomes available: the Synecdoche View.

## 6. The Synecdoche View

The lesson of §5 was that cognitive scientists ascribe psychological predicates to the brain on the basis of human behaviour. But human behaviour is not uniquely a ground for ascribing psychological predicates to the brain, it is simultaneously a ground for ascribing such predicates to the human being. In the example of Frith (2007), how human beings interpret shadows is not just a ground for saying that the *brain* decides whether an object is concave or convex but simultaneously a ground for saying that the *human being* decides whether an object is concave or convex. And in the examples of Sperry (1974) and Gazzaniga (1995), how human beings identify objects of perception presented to one visual field, hand, foot or nostril, but cannot identify the same objects if presented to the other visual field, hand, foot or nostril, is not just a ground for saying that the *hemisphere* perceives, remembers, recognizes etc. It is simultaneously a ground for saying that the *human being* perceives, remembers, recognizes etc.

Given that the grounds for ascribing psychological predicates to the brain are simultaneously grounds for ascribing psychological predicates to the human being, and the three first assumptions identified in §4 – that (1) we are human beings and not brains, (2) we possess psychological properties, and (3) cognitive scientists use psychological predicates with an ordinary literal sense – I propose a novel view: *the Synecdoche View*. Synecdoche is a subclass of metonymy where a term for a part refers to a whole (or vice versa), like ‘my new set of wheels’ refers to my new car. On the Synecdoche View, we can take the controversial locutions of cognitive scientists as non-literal figures of speech, namely synecdoches in which a term for a part (‘brain’) refers to a whole (human being). In the example of Frith (2007), the ascription of deciding to the brain reports the deciding of the human being, and in the examples of Sperry (1974) and Gazzaniga (1995), the ascription of perceiving, remembering, recognizing etc. to the brain reports the perceiving, remembering, recognizing etc. of the human being.<sup>72</sup>

Of course, it may be that the scientists themselves implicitly adopt a semantic view like the Literal View and believe that ‘brain’ and ‘hemisphere’ refer to the brain and hemisphere respectively. Perhaps this is why Sperry does not simply say that the minor hemisphere perceives, remembers, recognizes, and so on, but also claims that ‘the minor hemisphere is indeed a conscious system *in its own right*’ (1974: 11, my emphasis), as this would seem like an odd thing to add if he held the Synecdoche View. But as noted in §1, the Synecdoche View

---

<sup>72</sup> The research on split-brain patients has prompted the ‘how many minds?’ debate (e.g. Schechter 2015). But whether one thinks of the asymmetries in the cognitive and perceptual abilities of these patients as manifesting two minds or one disunified mind does not matter for my claim that the subject of the psychological properties is the human being. The Synecdoche View is consistent both with the claim that these patients have two minds and with the claim that they have one disunified mind.

is not about speaker-meaning, so if these scientists implicitly adopt the Literal View, then the speaker-meaning of ‘The brain thinks’ diverges from the semantic meaning (i.e. truth-condition) of that sentence. But that is no objection to the Synecdoche View, because that view is only a claim about semantic meaning, not speaker-meaning.

One important fact about synecdoches is that they often exploit associations of parts and wholes (Lakoff and Johnson 1980: 35-40). Given that psychological properties are closely *associated* with the brain we can see what might motivate using ‘brain’ as a synecdoche for ‘human being’. Using ‘brain’ synecdochically in talk of psychological properties indicates the *functional relevance* of the brain for instantiating such properties. It would have been arbitrary to use e.g. ‘lung’ instead of ‘brain’ just as it would have been arbitrary to use ‘doors’ instead of ‘wheels’ in ‘my new set of wheels’, given that wheels are more closely associated with the function of a car than doors are. Of course, this does not prevent anyone from using ‘lung’ and ‘doors’ as synecdoches for human being and car respectively, since one is free to use words however one likes. The point, however, is that it would (in normal contexts) be arbitrary to use ‘lung’ and ‘doors’ that way, while the synecdochical use of ‘brain’ and ‘wheels’ might be motivated by the functional relevance of brains and wheels.

Synecdoches are figures of speech, so according to the Synecdoche View, scientists’ ascription of psychological predicates to the brain is a figurative way of saying what humans literally do. The psychological predicates can be understood literally, but the use of ‘brain’ is figurative and refers to the human being.<sup>73</sup> This does not imply that every time cognitive scientists use the word ‘brain’ they are referring to a human being and not a brain. For example, when a scientist says that ‘our fMRI scan detected activity in the brain’, then ‘brain’ refers to the brain. But in the controversial locutions under discussion, it is a live possibility that ‘brain’ refers to the human being.

It should be noted that sometimes the human being is not conscious of the psychological states that cognitive scientists ascribe to the brain, that is, the human being does not have any higher-order awareness of these psychological states. The predictive processing literature – where the brain is said to predict, guess, estimate or expect the causes of sensory input using Bayesian models – supplies several examples. Consider the following passage from Körding and Wolpert (2004):

---

<sup>73</sup> Because synecdoches and other metonyms are figures of speech, it is hard to understand Machamer and Sytsma’s (2009) view that the ascription of psychological predicates to the brain is *both* a form of metonymy where a part refers to a whole *and* a literal extension of psychological predicates to describe brain behaviour (2009: 358-360). One and the same scientist cannot use the word ‘brain’ both literally – referring to the brain, and metonymically – referring to the human being – at one and the same time.

the central nervous system [...] employs probabilistic models during sensorimotor learning (2004: 244).

This hypothesis is supposed to explain how tennis players are able to estimate the velocity of a tennis ball. But the fact that tennis players estimate the velocity of a tennis ball is no less evidence that the tennis player, i.e. the *human being*, employs probabilistic models than it is evidence that the *brain* does. In fact, the authors oscillate between ascribing Bayesian estimation to the brain and the ‘subjects’ of their study (2004: 244, 246) who plausibly are human beings.

True, the subjects do not *consciously* estimate anything using Bayesian models. But this does not mean that the brain does it consciously. Körding and Wolpert say themselves that ‘subjects *implicitly* use Bayesian statistics’ (2004: 246, my emphasis), and Andy Clark – one of the main figures in this literature – emphasizes that the hypothesized prediction is not meant to be conscious (2016: 2, 27). So even though human beings are not conscious of the psychological states that scientists ascribe to the brain, the Synecdoche View can still accommodate the relevant locutions – as long as the grounds for ascribing those psychological states to the brain are simultaneously grounds for ascribing the same states to the human being.

This brings out an important point about the scope of the Synecdoche View. The view applies to the locutions of cognitive scientists documented by philosophers, such as those quoted in §1, which prompted the philosophical debate. These locutions all involve ascriptions of psychological predicates to the brain made on grounds that are *simultaneously* grounds for ascribing those predicates to the human being. Perhaps there are also examples of cognitive scientists ascribing psychological predicates to the brain on grounds that are *uniquely* grounds for ascribing those predicates to the brain. If so, the Synecdoche View does *not* apply to these examples. But no such examples have been documented in the debate. The locutions discussed above are not cherry-picked by me, they are the examples discussed by proponents of the Nonsense View and the Literal View.

One may think that the restricted scope of the Synecdoche View makes it a less powerful view compared to a view that offers a unified account – regardless of whether grounds for assertion are uniquely grounds for asserting that the brain thinks or simultaneously grounds for asserting that the human being thinks. Perhaps the Nonsense View and the Literal View offer such unified accounts. But it is plausible that grounds for assertion matter for semantics. ‘My new set of wheels was expensive’ should be interpreted literally when asserted on the basis of

the new wheels I have bought to replace the old ones of the old car, and it should be interpreted synecdochically when asserted on the basis of the new car I have bought. It would be strange to insist on one single semantics for the sake of ‘unity’.

Finally, even though the Synecdoche View is restricted to ascriptions of psychological predicates to the brain made on grounds that are simultaneously grounds for ascribing those predicates to the human being, it is *not* restricted to ascriptions of psychological predicates to the brain made on the basis of human behaviour. Human beings *can* be ascribed psychological predicates on the basis of non-behavioural grounds, such as neural activity or computational simulations of neural activity. For example, it is common to ascribe psychological states to someone who is paralysed (e.g. someone with locked-in syndrome) on the basis of neural activity. So there may be cases in which brain activity is a ground for ascribing psychological predicates to both the brain and the human being. And in these cases, the ascription of psychological predicates to the brain *can* be accommodated by the Synecdoche View, since this view applies to any ascription of psychological predicates to the brain made on grounds that are *simultaneously* grounds for ascribing the same predicates to humans – regardless of what those grounds are. What it does *not* apply to are ascriptions of psychological predicates to the brain made on grounds that are *uniquely* grounds for ascribing such predicates to the brain.<sup>74</sup>

## 7. The Synecdoche View vs. the Nonsense View and the Literal View

Unlike the Nonsense View and the Literal View, the Synecdoche View is agnostic about (P1) and (P2) of the Nonsense Argument, which say that it only makes sense to ascribe psychological

---

<sup>74</sup> We are now in a position to see why (as mentioned in footnote 66), Figdor’s (2018) view is ambiguous and may not be a defence of the Literal View at all. On the one hand, she explicitly opposes Bennett and Hacker’s Nonsense View and says that there are similarities between brains and human beings captured in mathematical models that warrant ascriptions of psychological predicates to brains (2018: 96). But on the other hand, none of the examples Figdor discusses from the scientific literature involve ascriptions of psychological predicates to brains. Rather, they involve ascriptions of psychological predicates to *other entities* discussed in biology, such as neurons, bacteria, plants, fruit flies, infants and cells. And in *these* examples, ascriptions of psychological predicates *are* made on the basis of the behaviour of these entities, in which case ‘neurons’, ‘bacteria’, ‘plants’ etc. should plausibly be taken literally and refer to neurons, bacteria, and plants respectively. This suggests an alternative interpretation of Figdor’s view, according to which her view is *not* a defence of what I have called the ‘Literal View’. For on my understanding, the Literal View is a view about the ascription of psychological predicates to *the brain*. But Figdor’s view may be restricted to ascriptions of psychological predicates to *other entities* in biology – made on the basis of the behaviour of those entities – such as bacteria, plants etc. Her discussion of Bennett and Hacker’s Nonsense View then, is perhaps only meant to address a possible opponent to *that* view and is therefore not strictly relevant to the debate with which this chapter is concerned. Of course, the fact that a literal interpretation is plausible for the ascription of psychological predicates to the entities Figdor discusses provides no support for what I have called the ‘Literal View’, and a literal interpretation of the examples Figdor discusses is perfectly consistent with the Synecdoche View, which – like the Literal View – is a view about examples other than those Figdor discusses.

predicates literally to an entity that (under normal circumstances) has the behaviour of an animal or human being. In other words, the Synecdoche View is agnostic whether brain behaviour is a legitimate ground for ascribing psychological predicates to the brain – in a literal sense of ‘brain’. The Synecdoche View is compatible with both Bennett and Hacker’s view that brain behaviour is *not* a legitimate ground, and Dennett, Crane and Figdor’s view that brain behaviour *is* a legitimate ground. But to repeat, this question of whether brain behaviour is a legitimate ground for ascribing psychological predicates to the brain is not relevant to the semantics of the controversial locutions of cognitive scientists, since these locutions are not based on the observation of (or mathematical model of) brain behaviour.

But the Synecdoche View is *not* agnostic about (P3) of the Nonsense Argument – that cognitive scientists make literal ascriptions of psychological predicates to the brain. This is something the Synecdoche View rejects but which the Nonsense View and the Literal View accept. The Synecdoche View accepts the assumption that *psychological predicates* should be taken literally, but it rejects the assumption that the word ‘brain’ demands the same literal interpretation. It is a live possibility that cognitive scientists are not referring to brains, but rather human beings in the controversial locutions where ‘brain’ is a grammatical subject of psychological predicates.

Once we see that cognitive scientists’ grounds for ascribing psychological predicates to the brain are simultaneously grounds for ascribing those predicates to the human being, the Synecdoche View becomes available as a compelling option that can take on board aspects of both the alternative views. It agrees with the Nonsense View, *contra* the Literal View, that the relevant discoveries of cognitive scientists do not give us reason to believe that brains possess psychological properties, or that scientists have discovered that psychological predicates extend more widely than often assumed. This is how the Synecdoche View, which is primarily an answer to a semantic question, has implications for metaphysics. Yet, it agrees with the Literal View, *contra* the Nonsense View, that cognitive scientists’ locutions both report empirical truths and make good sense, which a charitable interpretation of the linguistic practice of a successful science should try to maintain.

The Synecdoche View offers a plausible semantics of the controversial locutions of cognitive scientists. But as I said above, I only defend this modest claim and not the further claim that it offers the best semantics. The reason is that, while there is reason to prefer the Synecdoche View over the Nonsense View, the Literal View and the Synecdoche View seem to be on par in the sense that there is no compelling reason to prefer one over the other. Let me elaborate.

Even if proponents of the Nonsense View tried to accommodate the fact that cognitive scientists' grounds for ascribing psychological predicates to the brain are simultaneously grounds for ascribing such predicates to human beings, the Synecdoche View would be more plausible. The main reason is that it is more charitable to say that cognitive scientists synecdochically ascribe psychological predicates to the brain, rather than saying that their claims make no sense or are false. The Synecdoche View can maintain that the locutions of cognitive scientists make good sense and express truths, which a charitable interpretation of the linguistic practice of a successful science should aim for. In this respect, it is more plausible than the Nonsense View.

Now let's consider the Literal View. Remember the first two assumptions identified in §4 – that (1) we are human beings and not brains, and (2) we possess psychological properties. If the human being literally thinks, and the brain literally thinks, and the grounds for ascribing thinking to both are the same, then both of them think simultaneously. This is a consequence of the Literal View, once it accommodates the fact that cognitive scientists' grounds for asserting that the brain thinks are simultaneously grounds for asserting that the human being whose brain it is thinks. The Synecdoche View, by contrast, only commits to one thinker in the literal sense, namely the human being. Eric Olson has argued that it is implausible to hold that both the brain and the human being think simultaneously, because he thinks it follows that I cannot know who I am. His argument is the following. Given that the brain thinks when and only when the human being thinks, and what and only what the human being thinks, both the brain and the human being think they are me. The brain thinks it falsely and the human being thinks it truly, because by assumption I am the human being. But I cannot know whether I am the human being or the brain inside me that falsely thinks it is me (Olson 2007: 80-81, 215-216, 2015: 46-47).

Perhaps proponents of the Literal View agree that we do not know who we are but deny that this is a problem. However, they need not. Derek Parfit (2012: 20-22) points out an assumption that can be questioned in Olson's argument, namely that first person pronouns ('I', 'me') always refer to the thinker of a first person thought.<sup>75</sup> A proponent of the Literal View can claim that those pronouns only refer to me, i.e. the human being, and not my brain, whenever my brain thinks thoughts involving first person pronouns. Consequently, it is not the case that my brain mistakenly thinks that it is me. The brain does not think that *it* is anything, because first person pronouns involved in its thoughts refer to me – the human being. So I can

---

<sup>75</sup> Parfit (2012) thinks that we are brains, so I have rephrased his argument in a way consistent with the assumption that we are human beings.



know who I am, namely the human being. And there is no question of the brain knowing who *it* is, for when it thinks that I am me, it does not (falsely) think that *it* is me, it (truly) thinks that I – the human being – am me – the human being. Thus, defenders of the Literal View can maintain that I know who I am.

To my knowledge, there is no further reason to prefer the Synecdoche over the Literal View. But is there any reason to prefer the Literal View? In discussion I have come across three arguments but neither of them are convincing.

The first argument is that the brain actually possesses psychological properties: if the brain thinks, then it is plausible to interpret ‘The brain thinks’ literally. But it is not clear why we should believe that the brain possesses psychological properties independent of the Literal View. If it is true that the brain possesses psychological properties, then that is not something we know *a priori*, but presumably something we know from science. How can we know whether science teaches us that brains possess psychological properties? We can begin by looking at what scientists say. And some cognitive scientists do ascribe psychological predicates to the brain and say things like ‘The brain thinks’. But this is only evidence that the brain possesses psychological properties *if those locutions should be interpreted literally*. I have argued above that locutions like those documented and discussed in the literature can plausibly be interpreted non-literally and that proponents of the Literal View have yet to provide evidence of other locutions that cannot plausibly be interpreted this way. The claim that the brain possesses psychological properties cannot be the support for the Literal View if the Literal View is the support for the claim that the brain possesses psychological properties, since that would be circular.

The second argument to prefer the Literal View is that interpreting the locutions of scientists literally is the default interpretation (cf. Figdor 2018: 87-88). But it is not clear why this is so. Scientific language often involves figurative elements, such as metaphors (Boyd 1979). So even though the default may be to interpret the locutions of scientists as making sense rather than being nonsensical, and reporting empirical truths rather than falsehoods, it is not clear why a *literal* interpretation should be the default interpretation.

The third argument to prefer the Literal View is that locutions with relevant similarities to the controversial locutions of cognitive scientists are literal. We say both ‘the human being kicks’ and ‘the foot kicks’ because the foot is functionally relevant to kicking, and the grounds for ascribing kicking to the foot are simultaneously grounds for ascribing kicking to the human being. Similarly, we say both ‘the eye sees’ and ‘the human being sees’ because the eye is functionally relevant to seeing, and the grounds for ascribing seeing to the eye are

simultaneously grounds for ascribing seeing to the human being. Given that these locutions are all literal and have several relevant similarities to the controversial locutions of cognitive scientists, it is plausible to interpret the latter locutions literally. But a defender of the Synecdoche View need not accept that locutions like ‘the foot kicks’ and ‘the eye sees’ are literal. These locutions have a well-established use but that does not imply that they are literal. They may be figurative, but in lack of a commonly accepted criterion for literalness or figurativeness it is not clear how to determine the issue.

It therefore seems to me that the Literal View and the Synecdoche View are on par and that there is currently no compelling reason to prefer one rather than the other. Hence my modest claim that the Synecdoche View is a plausible (not the ‘best’) alternative.

## 8. Conclusion

It is a live possibility that cognitive scientists’ ascriptions of psychological predicates to the brain should be understood synecdochically and report the psychological properties of human beings – at least those ascriptions made on grounds that are simultaneously grounds for ascribing psychological predicates to human beings. This claim rests on the assumptions that (1) we are human beings and not brains, (2) we possess psychological properties, and (3) cognitive scientists use psychological predicates in an ordinary literal sense. If one rejects (1) or (2), perhaps the best interpretation of the controversial locutions is a literal interpretation and that ascriptions of psychological predicates to human beings should be interpreted synecdochically instead – with ‘human being’ referring to the brain. If one rejects (3), perhaps the best interpretation is one according to which cognitive scientists’ ascriptions of psychological predicates to the brain is just a colourful way of talking about neural activity. I have not addressed these alternative views because my aim was to argue that the Synecdoche View is a plausible view on the *assumption* that (1)-(3) is true.

Even though I do not make the further claim that the Synecdoche View is the *best* view, the claim that it is a *plausible* view is still an interesting result, both for semantics and metaphysics. One consequence is that, contrary to both the Nonsense View and the Literal View, there is no reason to take the controversial locutions literally. A second consequence is that, contrary to what e.g. Figdor (2018: 5-6, 61) claims, there is no reason to believe that the controversial locutions indicate empirical support for the claim that brains possess psychological properties, or that scientists have discovered that psychological predicates extend

to a wider range of entities than often assumed. These are substantive and controversial consequences, which indicate the significance of the Synecdoche View.

# Chapter 8

## Conclusion

### 1. Introduction

In chapter 1, I introduced the linguistically sensitive approach, which consists of (a) committing to semantic externalism, conceptual freedom and conceptual pluralism, and (b) trying to avoid conceptual confusions (i.e. verbal disputes and conceptual conflations). I said that in the main body of the thesis (i.e. chapters 2-7) that engages directly or indirectly with first-order debates, I would not charge anyone with being subject to conceptual confusions. In these chapters, the idea of conceptual confusions only had a bearing on how I navigated in the relevant philosophical debates, in the sense that I distinguished senses of key expressions and tried to avoid conceptual confusions when engaging with philosophical questions. In this final chapter, I address the elephant in the room: are there any examples of conceptual confusions in the debates with which I have engaged?

I think it is plausible that there are examples of this. To show that this is a live possibility, I offer a speculative error theory, according to which some of the views I argue against in the previous chapters are based on conflating different meanings associated with the same expression. In other words, I propose that some of my opponents' views are based on conceptual conflations. This theory is 'speculative' in the sense that I only provide light evidence for it, and it is an 'error theory' in the sense that it explains why my opponents came to hold (according to me) mistaken philosophical views. Three clarifications are in order before I outline the error theory.

Firstly, the claim that some of my opponents' views are based on conceptual conflations is consistent with these views being justified by philosophical arguments. The idea is that conceptual conflations explain why philosophers came to hold the views they hold, or that conceptual conflations caused them to hold these views, which is consistent with philosophical arguments justifying the views. To put it simply, the idea is that philosophers came to hold that *p* because of a conceptual conflation, and after that they developed some arguments designed to justify *p*.<sup>76</sup>

---

<sup>76</sup> The distinction between that which explains why people hold a view and that which justifies the view corresponds to Reichenbach's distinction between the 'context of discovery' and the 'context of justification' (1938).

Secondly, whether the speculative error theory is true does not matter for the first-order debates with which I have engaged, since all that matters for those debates are the philosophical arguments, or the evidence for, the various views. This is also why I have not charged any of my opponents with conceptual confluations in the previous chapters. Still, it is interesting to consider the possibility that conceptual confluations occur in these debates, since if they do, they might explain why philosophers hold some of the views they hold.

Thirdly, the speculative error theory only says that *some* of the views I engage with in this thesis are based on conceptual confluations, it does not say that all the views I argue against are based on conceptual confluations, nor that all mistakes in philosophy are based on conceptual confluations. Conceptual conflation is only one of many sources of mistaken philosophical views, so it would be extremely implausible to claim that all mistakes in philosophy are based on conceptual conflation. But several of the debates I have engaged with arguably involve cases of conceptual pluralism, so there is at least room for conceptual conflation in these debates.

## **2. The Speculative Error Theory**

The speculative error theory I propose offers a model of how certain philosophical views come about. The first step is that a person is familiarized with two or more meanings of an expression. The second step is that the person conflates some of these meanings. The third step is that the person adopts a philosophical view based on this conflation. In chapter 1, I stipulated a non-philosophical example of this. I refer to the banana example, where I consistently use ‘banana’ to refer to apples and thereby create a distinct sense of ‘banana’ but conflate my novel sense of ‘banana’ with the ordinary sense, and therefore claim that bananas in the ordinary sense of the word are apples. We see how the above model applies to this banana example. I shall now describe how the model applies to the philosophical views I have in mind.

In chapter 2, I argued against the non-technical view, which says that the ‘what-it’s-like’ phrase has an ordinary meaning in the literature on phenomenal consciousness. According to the speculative error theory, the non-technical view is based on a conceptual conflation. The conflation in question is of the technical meaning of the ‘what-it’s-like’ phrase relevant to phenomenal consciousness, and one or more of the ordinary meanings discussed in chapter 2, i.e. the evaluative-descriptive sense, the non-evaluative-descriptive sense, the resemblance sense, and the ability sense (see chapter 2, §§2-3 for more on these senses). For example, Hellie (2004) thinks that when The Beatles sing ‘I know what it’s like to be dead’, their use of the ‘what-it’s-like’ phrase is the same as that which is characteristic in the literature on phenomenal

consciousness (2004: 369). But given the technical view, lay people's use of the 'what-it's-like' phrase, such as that of The Beatles, is not the same as that which is characteristic in the literature on phenomenal consciousness. Rather, their use of the phrase involves one of the four ordinary senses of the phrase mentioned above, in this case, plausibly the ability sense or the non-evaluative-descriptive sense (see chapter 2, §4). The idea then, is that Hellie holds the non-literal view because he conflates the sense relevant to phenomenal consciousness and the ordinary senses involved in examples like that of The Beatles, and that the same applies to other defenders of the non-technical view.

In this example, the conflation is of an ordinary sense and a technical sense peculiar to philosophy and cognitive science. The claim that a philosophical view is based on conflating a technical with a non-technical sense is reminiscent of Wittgenstein's claim that 'philosophical problems arise when language goes on holiday' (PI §38). On one interpretation of this claim, language on holiday is ordinary expressions that are used in a new way in which they are not ordinarily used, i.e. ordinary expressions that are used with a technical sense. So understood, Wittgenstein's claim is that philosophers are misled by their own technical use of ordinary expressions.

I am not sure what Wittgenstein would say about why language on holiday misleads philosophers. But on my view, a plausible explanation of why language on holiday can cause misunderstanding is the combination of the ideas discussed in chapter 1, namely conceptual freedom, semantic externalism and conceptual pluralism. Given conceptual freedom, there is nothing wrong about language on holiday. But given semantic externalism, language on holiday creates a distinct technical sense of the expression that is on holiday, since externalism says that things external to the speaker's mind, such as the speaker's use of words, determines the meaning of the words of the speaker. This gives rise to conceptual pluralism with respect to the relevant expression, and thereby the danger of conceptual conflation. In other words, the idea is that when philosophers use an ordinary expression with a technical meaning – i.e. when language is on holiday – they sometimes do not realize, or forget, that they use the expression with a technical meaning and simultaneously associate the relevant expression with an ordinary meaning with which they are *not* using the expression. The result is a conceptual conflation.

This is what the speculative error theory says about the non-technical view about the 'what-it's-like' phrase: the non-technical view is an effect of the 'what-it's-like' phrase being on holiday. That is, philosophers use the phrase with a technical sense but forget or do not realize that they do, hence they conflate the technical sense with an ordinary sense and claim that the phrase has a non-technical meaning.

The second example is from chapter 5, where I argued against the Experiential Theory of pain, on which pains in the locatable sense are experiences or mental states located in the mind or brain. According to the speculative error theory, the Experiential Theory is based on a conceptual conflation, namely that of the experiential sense and the locatable sense of ‘pain’ (see chapter 5, §2.1. for the distinction). For ‘pain’ in the experiential sense refers to experiences and ‘pain’ in the locatable sense (given the Bodily Theory) does not, so the conflation of these two senses explains why some philosophers hold that pains in the locatable sense are experiences.<sup>77</sup>

Does the conflation of the experiential and locatable senses of ‘pain’ involve language on holiday? It depends on whether any of these senses is a technical sense. The locatable sense is clearly not technical, since to say one has a pain in a body part is not peculiar to a defined theoretical community. But the experiential sense may well be technical. As I said in chapter 5, I doubt that lay people call the experience of pain ‘pain’, in which case the experiential sense of ‘pain’ is a technical sense peculiar to philosophy and cognitive science. It does not really matter whether it is technical, however, since a conflation with the locatable sense is a conflation either way. But if it is technical, then it involves language on holiday, just like in the above example with the ‘what-it’s-like’ phrase.

The third example is from chapter 7, where I discussed the Nonsense View and the Literal View, both according to which cognitive scientists’ ascription of psychological predicates to the brain is literal with ‘brain’ referring to the brain. According to the speculative error theory, the claim that the word ‘brain’ is literal in the controversial locutions of cognitive scientists is based on a conceptual conflation. The conflation in question is of the standard sense of ‘brain’, which is the sense with which one uses the word when one refers to the brain, and (assuming the Synecdoche View) the synecdochical sense of the word with which it is used in the controversial locutions of cognitive scientists. This conflation explains why defenders of the Nonsense View and the Literal View believe that ‘brain’ is used in the literal sense in the controversial locutions of cognitive scientists.

---

<sup>77</sup> In my view, this conceptual conflation can also explain another curiosity in the pain literature, which I did not discuss in chapter 5. Several philosophers discuss the ‘paradox of pain’: how can pains be experiences located in the head and things that are located in the body at one and the same time (Bain 2007, Hill 2014, Aydede 2020, Bradley 2021)? But once we distinguish pain in the locatable sense from pain in the experiential sense there is no more a paradox of pain than there is a paradox of bank (how can the bank be a financial institution *not* located next to the river and be a thing that *is* located next to the river?). In fact, the sense of paradox may be based on a conceptual conflation, i.e. the conflation of the locatable sense and the experiential sense of ‘pain’.

Unlike the above examples with the ‘what-it’s-like’ phrase and ‘pain’, the synecdochical use of ‘brain’ is arguably not technical, since there are common expressions like ‘my brain doesn’t understand this’ or ‘my brain has to make up its mind’ that can plausibly be interpreted synecdochically and are not peculiar to a defined theoretical community like philosophy and cognitive science. If this is right, the synecdochical use of ‘brain’ does not involve language on holiday. But insofar as the sense of ‘brain’ operative in the controversial locutions of cognitive scientists is different from the more standard sense of ‘brain’, the claim that ‘brain’ is literal in the controversial locutions of cognitive scientists might be based on a conceptual conflation.

This is how the speculative error theory works in practice. It claims that in these three cases, philosophers are misled by there being different senses of certain expressions. In the case of the ‘what-it’s-like’ phrase, and arguably in the case of ‘pain’, philosophers and scientists use ordinary expressions with a technical meaning peculiar to philosophy and cognitive science. In the case of ‘brain’, cognitive scientists arguably use the word with a meaning that is non-technical, but which is nevertheless not the most standard. Given conceptual freedom, there is nothing wrong with this. Indeed, conceptual freedom is the thesis that we can speak however we like, which implies that we can use these expressions with technical or non-standard meanings if we want to. But given semantic externalism, according to which things external to a speaker’s mind, such as the speaker’s use of an expression, determine facts about the meaning of the words of the speaker, the non-standard uses of these ordinary expressions give rise to conceptual pluralisms with respect to these expressions. If we are not careful then, we can end up with conceptual conflations, i.e. conflating the meanings with which these expression are used with meanings with which they are *not* used.

I said that it does not matter for the first-order debates whether the speculative error theory is true, since all that matters for those debates are the philosophical arguments, or the evidence for, the various first-order views – regardless of how those views came about in the first place. Nevertheless, it is clear that *if* the speculative error theory is true, then it would also be true that conceptual conflations can lead to dubious claims in first-order debates. This is because I have argued that all three claims that by hypothesis are based on conceptual conflation – that (a) the ‘what-it’s-like’ phrase is non-technical in the literature on phenomenal consciousness, (b) pains in the locatable sense are experiences, and (c) ‘brain’ is literal in the controversial locutions of cognitive scientists – are dubious claims. And at least in the case of (a) and (c) we have seen that these claims may form the basis for further dubious claims in first-order debates, such as the claim that the definition of phenomenal consciousness in terms of



what it is like to be in a mental state is informative, that lay people believe in phenomenal consciousness, and that brains possess psychological properties.

### **3. Two Objections**

One possible objection to the speculative error theory is to point out that the claim that philosophers conflate meanings is itself in need of explanation. Why do competent speakers conflate meanings? One explanation is that those who use the relevant expressions in the philosophical and cognitive scientific literature typically do not make clear what they mean with, or which of the meanings they have in mind with respect to, the relevant expressions.

In the literature on phenomenal consciousness, those who use the ‘what-it’s-like’ phrase typically neither do much to clarify what sense of the phrase they have in mind, nor contrast whatever sense they have in mind with other senses of the phrase. As Block (1978) says when imagining someone asking what is meant by common terms in the literature on phenomenal consciousness: ‘If you got to ask, you ain’t never gonna get to know!’ (1978: 241). Similar points apply to the pain literature, which contains a lot of discussion on the metaphysics of pain, but very rarely contains any discussion of what sense of ‘pain’ one has in mind. Finally, the same is also true of the use of ‘brain’ in the controversial locutions of cognitive scientists, as cognitive scientists do not say what they mean by ‘brain’ when they say things like ‘The brain thinks’. This lack of clarification explains why people (according to the speculative error theory) conflate meanings. Indeed, given my arguments that the ‘what-it’s-like’ phrase is technical, that ‘pain’ in the locatable sense refers to bodily occurrences, and that the Synecdoche View is plausible, this lack of clarification is the main evidence in favour of the error theory, as it suggests that theorists believe key expressions only has one sense rather than multiple senses.

Of course, this explanation of why people conflate meanings raises a further question, namely why theorists do not clarify what they mean with key expressions. One explanation is that they are neither aware of the different meanings of key expressions, nor are they always aware of what in fact are the meanings of key expressions they use. In the case of the ‘what-it’s-like’ phrase and ‘pain’, my opponents do not seem to be aware of what the ordinary meanings of these expressions are, and in the case of ‘brain’, they do seem to be aware of what the standard literal meaning is, but they do not seem to be aware of what the less standard, non-literal meaning with which cognitive scientists use the word, is.

The reason theorists are not aware of these things is plausibly that they do not take the linguistically sensitive approach outlined in chapter 1. Rather, they take what one might call the ‘what is X approach’ and start engaging with the metaphysics of mind without first clarifying key expressions. As noted in chapter 1, Chalmers warns that the ‘what is X approach’ might lead to verbal disputes (2011: 538). My hypothesis is that it might also lead to conceptual confluences.

Another possible objection to the speculative error theory is to point out that we need an explanation for why philosophers and scientists started using ordinary expressions with a technical or non-standard sense. But for two of the three examples under consideration, I have already offered explanations.

In chapter 2, I suggested that philosophers started using the ‘what-it’s-like’ phrase in connection with phenomenal consciousness because the ordinary expression ‘you do not know what it is like before you have experienced it’ seems to capture the epistemic gap inherent in the non-functional concept of phenomenal consciousness. I suggested that the above expression as used by non-philosophers – perhaps unbeknownst to philosophers – involves the ability sense of the ‘what-it’s-like’ phrase and is not concerned with phenomenal consciousness. Thus, when philosophers started using the ‘what-it’s-like’ phrase in connection with phenomenal consciousness, they introduced a technical meaning of the ‘what-it’s-like’ phrase. In chapter 7, I suggested that the synecdochical use of the word ‘brain’ in ‘The brain thinks’ is motivated by the functional relevance of the brain for thinking, similar to how the functional relevance of wheels might motivate the synecdochical use of ‘wheels’ in ‘my new set of wheels’. This functional relevance explains why cognitive scientists are using ‘brain’ with a sense that is not the most standard.

What about the experiential sense of ‘pain’, which I suggested is a sense peculiar to philosophy and cognitive science and thus a technical sense? Unlike the examples with the ‘what-it’s-like’ phrase and ‘pain’, I see no theoretical motivation to introduce this sense of ‘pain’ and call the experience of pain ‘pain’. This does not imply that there is anything wrong with this use of ‘pain’ (I am still a defender of conceptual freedom.) Perhaps it was introduced as a linguistic shortcut for theorists working on the experience of pain, I am not sure. In any case, we see that there are possible explanations available for why philosophers and cognitive scientists started to use ordinary expressions with non-standard meanings, which answers the above-mentioned worry with the speculative error theory.

#### 4. Conclusion

This thesis started by outlining the linguistically sensitive approach, which (a) commits to semantic externalism, conceptual freedom and conceptual pluralism, and (b) tries to avoid conceptual confusions. It then applied this approach to debates that are directly or indirectly about three first-order topics, namely phenomenal consciousness, pain and the extension of psychological predicates. In the chapters that engaged with these first-order debates, I did not charge other theorists who engage in first-order debates with being subject to conceptual confusions, however. The only use I made of the idea was for myself when navigating in first-order debates. This final chapter has addressed the elephant in the room and considered whether there are any examples of conceptual confusions in the debates with which I have engaged. To this end, I offered an error theory, which points out possible examples of conceptual confluations. This is a *speculative* theory, and I have not defended it in detail here. But if nothing else, I hope the connections it points out makes sense of why I have taken the linguistically sensitive approach and why I have spent so much space talking about words and their use, even though the aim has been to contribute to first-order philosophy of mind.

## Bibliography

- Alles S. R. A. and Smith P. A. (2018) Etiology and Pharmacology of Neuropathic Pain. *Pharmacological Reviews* 70.
- Apkarian V. (2017) Advances in the Neuroscience of Pain. In Corns J. (ed.) *The Routledge Handbook of Philosophy of Pain*. New York: Routledge.
- Arendt-Nielsen L. and Svensson P. (2001) Referred Muscle Pain: Basic and Clinical Findings. *The Clinical Journal of Pain* 17.
- Armstrong D. (1968) *A Materialist Theory of the Mind*. London: Routledge.
- Aydede M. (2017) Pain. Perception or Introspection? In Corns J. (ed.) *The Routledge Handbook of Philosophy of Pain*. New York: Routledge.
- (2020) What is a Pain in a Body Part? *Canadian Journal of Philosophy* 50.
- Bain D. (2007). The Location of Pains. *Philosophical Papers* 36.
- (2014) Pains That Don't Hurt. *Australasian Journal of Philosophy* 92.
- Baker G. P. and Hacker P. M. S. (2005) *Wittgenstein: Understanding and Meaning*. 2<sup>nd</sup> ed. extensively revised by P. M. S. Hacker. Oxford: Blackwell.
- Baliki M. N. and Apkarian V. (2015) Nociception, Pain, Negative Moods, and Behaviour Selection. *Neuron* 87.
- Balmer A. (2020) Soft-Wired Illusionism vs. The Meta-Problem. *Journal of Consciousness Studies* 27.
- Bennett M. R. (2007) Philosophy and Neuroscience. In Bennett M. R. et al. *Neuroscience and Philosophy*. New York: Columbia University Press.
- Bennett M. R. and Hacker P. M. S. The Conceptual Presuppositions of Cognitive Neuroscience: A Reply to Critics. In Bennett M. R. et al. *Neuroscience and Philosophy*. New York: Columbia University Press.
- (2003) *Philosophical Foundations of Neuroscience*. Oxford: Blackwell.
- Block N. (2002) The Harder Problem of Consciousness. *Journal of Philosophy* 99.
- (1995) On a Confusion about a Function of Consciousness. *Behavioural and Brain Sciences* 18.
- (1978) Troubles with Functionalism. *Minnesota Studies in the Philosophy of Science* 9.
- Bogduk N. (2009) On the Definitions and Physiology of Back Pain, Referred Pain, and Radicular Pain. *Pain* 147.

- Borg E. et al. (2020) Is the Folk Concept of Pain Polyeidic? *Mind and Language* 35.
- Boyd R. (1979) Metaphor and Theory Change: What is 'Metaphor' a Metaphor For? In Ortony A. (ed.) *Metaphor and Thought*. Cambridge: Cambridge University Press.
- Bradley A. (2021) The Paradox of Pain. *The Philosophical Quarterly*.
- Brown R. (2010) Deprioritizing the A Priori Arguments Against Physicalism. *Journal of Consciousness Studies* 17.
- Burge T. (1979) Individuation and the Mental. *Midwest Studies in Philosophy* 4.
- Byrne A. (2001) Intentionalism Defended. *Philosophical Review* 110.
- (2004) What Phenomenal Consciousness is Like. In Gennaro R. J. (ed.) *Higher-Order Theories of Consciousness: An Anthology*. Philadelphia PA: J. Benjamins.
- Byrne A. and D. R. Hilbert (2011) Are Colours Secondary Qualities? In Nolan L. (ed.) *Primary and Secondary Qualities: The Historical and Ongoing Debate*. Oxford: Oxford University Press.
- Cain S. (1969) Thorns Grow in Garden of Eden. *The Boston Globe*.
- Carruthers P. (2000) *Phenomenal Consciousness: A Naturalistic Theory*. Cambridge University Press.
- Carruthers P. and Veillet B. (2017) Consciousness Operationalized, a Debate Realigned. *Consciousness and Cognition* 55.
- Carter J. A. (2012) Recent Work on Moore's Proof. *International Journal for the Study of Scepticism* 2.
- Casser L. C. And Schiller H. I. (2021) A Hole in the Box and a Pain in the Mouth. *The Philosophical Quarterly*.
- Chalmers D. J. (2020a) How Can We Solve the Meta-Problem of Consciousness? *Journal of Consciousness Studies* 27.
- (2020b) Is the Hard Problem of Consciousness Universal? *Journal of Consciousness Studies* 27.
- (2018) The Meta-Problem of Consciousness. *Journal of Consciousness Studies* 25.
- (2011) Verbal Disputes. *Philosophical Review* 120.
- (2010) *The Character of Consciousness*. Oxford: Oxford University Press.
- (2003) Consciousness and its Place in Nature. In Stich S. P. and Warfield T. A. (eds.) *Blackwell Guide to the Philosophy of Mind*. Blackwell.

——— (2002) Does Conceivability Entail Possibility. In Gendler T. S. and Hawthorne J. (eds.) *Conceivability and Possibility*. Oxford: Oxford University Press.

——— (1996) *The Conscious Mind: In Search of a Fundamental Theory*. Oxford: Oxford University Press.

Chekhov A. (1912) *The Sea-gull*. New York: Charles Scribner's Sons.

Clark A. et al. (2019) Bayesing Qualia: Consciousness as Inference, not Raw Datum. *Journal of Consciousness Studies* 26.

Clark A. (2016) *Surfing Uncertainty*. Oxford: Oxford University Press.

Conee E. (1994) Phenomenal Knowledge. *Australasian Journal of Philosophy* 72.

Coppes O. J. M. and Sang C. (2018) Phantom Limb Pain – a Review of Mechanisms, Therapy, and Prevention. *Annals of Anesthesia and Pain Medicine* 1.

Corns J. (2020) *The Complex Reality of Pain*. New York: Routledge.

Crane T. (2015) The Mental States of Persons and their Brains. *Royal Institute of Philosophy Supplement* 76.

Crick F. (1995) *The Astonishing Hypothesis*. New York: Touchstone.

Damasio A. (1996) *Descartes' Error – Emotion, Reason and the Human Brain*. London: Papermac.

Damo O. (2012) Gaza: 'If I'm afraid, what is it like for the children?' URL: <http://blogs.savethechildren.org.uk/2012/11/gaza-if-im-afraid-what-is-it-like-for-the-children/>

Dennett D. (2007) Philosophy as Naïve Anthropology: Comment on Bennett and Hacker. In Bennett M. R. et al. *Neuroscience and Philosophy*. New York: Columbia University Press.

——— (1988) Quining Qualia. In Marcel A. J. and Bisiach E. (eds.) *Consciousness in Contemporary Science*. Oxford: Oxford University Press.

Diaz R. (2021) Do People Think Consciousness Poses a Hard Problem? Empirical Evidence on the Meta-Problem of Consciousness. *Journal of Consciousness Studies* 28.

Dretske F. (2003) How Do You Know You Are Not a Zombie? In Gertler B. (ed.) *Privileged Access: Philosophical Accounts of Self-Knowledge*. New York: Ashgate

Farrell B. A. (1950) Experience. *Mind* 59.

Farrell J. (2016) 'What it is Like' Talk is not Technical Talk. *Journal of Consciousness Studies* 23.

- Figdor C. (2018) *Pieces of Mind. The Proper Domain of Psychological Predicates*. Oxford: Oxford University Press.
- Flor H. (2002) Phantom-limb pain: Characteristics, Causes, and Treatment. *The Lancet Neurology* 1.
- Frankish K. (2019) The Meta-Problem is the Problem of Consciousness. *Journal of Consciousness Studies* 26.
- (2016a) Illusionism as a Theory of Consciousness. *Journal of Consciousness Studies* 23.
- (2016b) Not Disillusioned: Reply to Commentators. *Journal of Consciousness Studies* 23.
- Frith C. (2007) *Making Up the Mind: How the Brain Creates Our Mental World*. Oxford: Blackwell.
- Funkhauser E. (2017) Belief as Signals: A New Function for Belief. *Philosophical Psychology* 30.
- Gaskin R. (2019) A Defence of the Resemblance Meaning of ‘What It’s Like’. *Mind* 128.
- Gazzaniga M. (1995) Principles of Human Brain Organization Derived from Split-Brain Studies. *Neuron* 14.
- Gibb B. and Gibb R. (1967) To Love Somebody. Recorded by Bee Gees on *Bee Gees 1st*. London: Polydor/Atco.
- Gottlieb S. and Lombrozo T. (2018) Can Science Explain the Human Mind? Intuitive Judgements about the Limits of Science. *Psychological Science* 29.
- Grahek N. (2007) *Feeling Pain and Being in Pain*. 2<sup>nd</sup> ed. Cambridge MA: MIT Press.
- Graziano M. S. A. (2019) We Are Machines That Claim to Be Conscious. *Journal of Consciousness Studies* 26.
- (2016) Consciousness Engineered. *Journal of Consciousness Studies* 23.
- Gregory D. et al. (submitted) Who Knows What Mary Knew.
- Grice H. P. (1989). *Studies in the Way of Words*. Cambridge, MA: Harvard UP.
- Hacker P. M. S. (2013) *The Intellectual Powers: a Study of Human Nature*. Chichester: Wiley Blackwell.
- (2002) Is there Anything it is Like to be a Bat? *Philosophy* 77.
- (1990) *Wittgenstein – Meaning and Mind. An Analytic Commentary on the Philosophical Investigations*. Vol. 3. Oxford: Blackwell.

- Hall R. J. (2008) If it Itches, Scratch! *Australasian Journal of Philosophy* 86.
- Hellie B. (2004) Inexpressible Truths and the Allure of the Knowledge Argument. In Nagasawa Y. et al. (eds.) *There's Something About Mary*. Cambridge, MA: MIT Press.
- Hill C. (2017) Fault Lines in Familiar Concepts of Pain. In Corns J. (ed.) *The Routledge Handbook of Philosophy of Pain*. New York: Routledge.
- (2014) *Meaning, Mind, and Knowledge*. Oxford: Oxford University Press.
- Horgan T. (2011) From Agentive Phenomenology to Cognitive Phenomenology: A Guide for the Perplexed. In Bayne T. and Montague M. (eds.) *Cognitive Phenomenology*. Oxford: Oxford University Press.
- Humphrey N. (2011) *Soul Dust: The Magic of Consciousness*. Princeton: Princeton University Press.
- Hyman J. (2003) Pains and Places. *Philosophy* 78.
- IASP (1979) IASP Pain Terms – 1979. *Pain* 6.
- Irvine E. (2012) *Consciousness as a Scientific Concept: A Philosophy of Science Perspective*. Dordrecht: Springer.
- Jackson F. (1982) Epiphenomenal Qualia. *The Philosophical Quarterly* 32.
- Janzen G. (2011) In Defense of the What-it-is-Likeness of Experience. *Southern Journal of Philosophy* 49.
- Kahan D. (2017) Misconceptions, Misinformation, and the Logic of Identity-Protective Cognition. *SSRN Electronic Journal*.
- Kammerer F. (2021) The Illusion of Conscious Experience. *Synthese* 198.
- (2018) Can You Believe It? Illusionism and the Illusion Meta-Problem. *Philosophical Psychology* 31.
- Kim et al. (2016) Where is Your Pain? A Cross-cultural Comparison of the Concept of Pain in Americans and South Koreans. *Studia Philosophica Estonica* 9.
- Kirsch J. (1987) And There Was Light by Jacques Lusseyran. Trans. Cameron E. R. In *Los Angeles Times*, 19 July. URL: [http://articles.latimes.com/1987-07-19/books/bk-4757\\_1\\_blind-man](http://articles.latimes.com/1987-07-19/books/bk-4757_1_blind-man)
- Klein C. (2015) *What the Body Commands. The Imperative Theory of Pain*. Cambridge, MA: MIT Press.
- Knobe J. and Prinz J. (2008) Intuitions About Consciousness: Experimental Studies. *Phenomenology and the Cognitive Sciences* 7.



- Körding K. P. and Wolpert D. M. (2004) Bayesian Integration in Sensorimotor Learning. *Nature* 427.
- Kriegel U. (2015) *The Varieties of Consciousness*. Oxford: Oxford University Press.
- (2006) Consciousness, Theories of. *Philosophy Compass* 1.
- Kripke S. (1980) *Naming and Necessity*. Harvard University Press.
- (1977) Speaker's Reference and Semantic Reference. *Midwest Studies in Philosophy* 2.
- Lakoff G. and Johnson M. (1980) *Metaphors We Live By*. Chicago: The University of Chicago Press.
- LeDoux J. (1998) *The Emotional Brain*. London: Phoenix.
- Lee Y. C. et al. (2011) The Role of the Central Nervous System in the Generation and Maintenance of Chronic Pain in Rheumatoid Arthritis, Osteoarthritis and Fibromyalgia. *Arthritis Research and Therapy* 13.
- Lennon J. And McCartney P. (1966) She Said She Said. Recorded by The Beatles on *Revolver*. London: Parlophone.
- Levine J. (2001) *Purple Haze: The Puzzle of Consciousness*. Oxford: Oxford University Press.
- Lewis D. (1995) Should a Materialist Believe in Qualia? *Australasia Journal of Philosophy* 73.
- (1988) What Experience Teaches. *Proceedings of the Russellian Society* 13.
- Liu M. (2021) Pain, Paradox, and Polysemy. *Analysis*.
- (2020) The Intuitive Invalidity of the Pain-in-Mouth Argument. *Analysis* 80.
- Lycan W. G. (1996) *Consciousness and Experience*. Cambridge, MA: MIT Press.
- Machamer P. and Sytsma J. (2009) Philosophy and the Brain Sciences. *Iris. European Journal of Philosophy and Public Debate* 1.
- Mandik P. (2016) Meta-Illusionism and Qualia Quietism. *Journal of Consciousness Studies* 23.
- Massin O. (2017) Bad by Nature: An Axiological Theory of Pain. In Corns J. (ed.) *The Routledge Handbook of Philosophy of Pain*. New York: Routledge.
- McClelland T. (2016) Gappiness and the Case for Liberalism About Phenomenal Properties. *The Philosophical Quarterly* 66.
- McDowell J. (1986) Singular Thought and the Extent of 'Inner Space'. In McDowell J. and Pettit P. (eds.) *Subject, Thought, and Context*. Oxford: Clarendon Press.

- Mehta N. (2021). The Fragmentation of Phenomenal Character. *Philosophy and Phenomenological Research*.
- Melzack R. and Wall P. D. (1965) Pain Mechanisms: A New Theory. *Science* 150.
- Menzies P. and Price H. (1993) Causation as a Secondary Quality. *British Journal for the Philosophy of Science* 44.
- Newton N. (1989) On Viewing Pain as A Secondary Quality. *Noûs* 23.
- Nida-Rümelin (2016) The Illusion of Illusionism. *Journal of Consciousness Studies* 23.
- Noordhof P. (2001) In pain. *Analysis* 61.
- (2002) More in pain. *Analysis* 62.
- (2005) In a state of pain. In Aydede M. (ed.) *Pain: New Essays on the Nature of Pain and the Methodology of Its Study*. Cambridge, MA: MIT Press.
- Olson E. (2007) *What Are We?* Oxford: Oxford University Press.
- (2015) On Parfit's View That We Are Not Human Beings. *Royal Institute of Philosophy Supplement* 76.
- Papineau D. (2011) What Exactly is the Explanatory Gap? *Philosophia* 39.
- Parfit D. (2012) We Are Not Human Beings. *Philosophy* 87.
- Pereboom D. (2019) Russellian Monism, Introspective Inaccuracy, and the Illusion Meta-Problem of Consciousness. *Journal of Consciousness Studies* 26.
- (2016) Illusionism and Anti-Functionalism About Phenomenal Consciousness. *Journal of Consciousness Studies* 23.
- Peressini A. (2014) Blurring Two Conceptions of Subjective Experience: Folk Versus Philosophical Phenomenality. *Philosophical Psychology* 27.
- Putnam H. (1975) The Meaning of 'Meaning'. *Minnesota Studies in the Philosophy of Science* 7.
- Reichenbach H. (1938) *Experience and Prediction. An Analysis of the Foundations and the Structure of Knowledge*. Chicago: The University of Chicago Press.
- Reuter K. and Sytsma J. (2020) Unfelt Pain. *Synthese* 197.
- Reuter K. et al. (2014) Hallucinating Pain. In Sytsma J. (ed). *Advances in Experimental Philosophy of Mind*. London: Bloomsbury.
- Reuter K. et al. (2019). Putting Pain in its Proper Place. *Analysis* 79.

Rey G. (1995) Towards a Projectivist Account of Conscious Experience. In Metzinger T. (ed.) *Conscious Experience*. Paderborn: Ferdinand Schöningh.

Rosenthal D. (2019) Chalmers' Meta-Problem. *Journal of Consciousness Studies* 26.

Russell B. (1926) Philosophical Consequences of Relativity. *Encyclopaedia Britannica*. Evanston IL: Benton Foundation.

Salomons T. et al. (2021) Is Pain 'All in Your Mind'? Examining the General Public's View of Pain. *Review of Philosophy and Psychology*.

Schechter E. (2015) The Subject in Neuropsychology: Individuating Minds in the Split-Brain Case. *Mind and Language* 30.

Schrody E. F. (1998) What It's Like. Recorded by Everlast on *Whitey Ford Sings the Blues*. Tommy Boy Records.

Schwarz W. (2019) From Sensor Variables to Phenomenal Facts. *Journal of Consciousness Studies* 26.

Searle J. (2007) Putting Consciousness Back in the Brain: Reply to Bennett and Hacker, *Philosophical Foundations of Neuroscience*. In Bennett M. R. et al. *Neuroscience and Philosophy*. New York: Columbia University Press.

Shabasson D. (2021) Illusionism About Phenomenal Consciousness: Explaining the Illusion. *Review of Philosophy and Psychology*.

Shoemaker S. (1994) Self-Knowledge and "Inner Sense": Lecture II: The Broad Perceptual Model. *Philosophy and Phenomenological Research* 54.

Smythies J. (2009) Philosophy, Perception, and Neuroscience. *Perception* 38.

Snowdon P. (2010) On the What-it-is-Like-Ness of Experience. *Southern Journal of Philosophy* 48.

Sperry R. (1974) Lateral Specialization in the Surgically Separated Hemispheres. In Schmitt F. O and Worden F. G. (eds.) *The Neurosciences Third Study Programme*. Cambridge, MA: MIT Press.

Sprigge T. L. and Montefiore (1971) Final Causes. *Proceedings of the Aristotelian Society, Supplementary Volumes* 45.

Stoljar D. (2016) The Semantics of 'What it's Like' and the Nature of Consciousness. *Mind* 125.

Strawson G. (2017) One Hundred Years of Consciousness. *Isaiah Berlin Lecture*, Wolfson College, Oxford, 25 May 2017.

——— (2006) Realistic Materialism: Why Physicalism Entails Panpsychism. *Journal of Consciousness Studies* 13.

Sytsma J. (2010) Dennett's Theory of the Folk Theory of Consciousness. *Journal of Consciousness Studies* 17.

Sytsma J. and Machery E. (2010) Two Conceptions of Subjective Experience. *Philosophical Studies* 151.

Sytsma J. and Ozdemir E. (2019) No Problem: Evidence that the Concept of Phenomenal Consciousness is not Widespread. *Journal of Consciousness Studies* 26.

Tye M. (2018) Qualia. The Stanford Encyclopedia of Philosophy (Summer 2018 Edition). Ed. Zalta E. N. URL: <https://plato.stanford.edu/archives/sum2018/entries/qualia/>

——— (2017) Are Pains Feelings? *The Monist* 100.

——— (2005a) Another Look at Representationalism about Pain. In Aydede M. (ed.) *Pain: New Papers on its Nature and the Methodology of its Study*. Cambridge, MA: MIT Press.

——— (2005b) In Defense of Representationalism: Reply to Commentaries. In Aydede M. (ed.) *Pain: New Papers on its Nature and the Methodology of its Study*. Cambridge, MA: MIT Press.

——— (2002) On the Location of Pain. *Analysis* 62.

——— (1995a) A Representational Theory of Pains and their Phenomenal Character. *Philosophical Perspectives* 9.

——— (1995b) *Ten Problems of Consciousness: A Representational Theory of the Phenomenal Mind*. Cambridge, MA: MIT Press.

Unknown (1891) *The Eagle* 16.

Unknown (1938) Sight to the Blind. *Bath Weekly Chronicle Herald*.

Willard M. B. (2014) Against Simplicity. *Philosophical Studies* 167.

Williamson T. (1994) *Vagueness*. London: Routledge.

Wilson D. and Sperber D. (2002) Truthfulness and Relevance. *Mind* 111.

Wittgenstein (2009) *Philosophical Investigations*. Trans. Anscombe G. E. M. et al. Oxford: Wiley-Blackwell.

——— (1980) *Remarks on the Philosophy of Psychology* Vol. I. Chicago: University of Chicago Press.

Zakkou J. (2018). The Cancellability Test for Conversational Implicatures. *Philosophy Compass* 13.

