

**AT THE GATES OF CONSCIOUSNESS: PHYSICALISM AND PHENOMENAL
CONCEPTS**

by

Erhan Demircioglu

B.S., Istanbul Technical University, Turkey, 2002

M.A., Bogazici University, Turkey, 2005

Submitted to the Graduate Faculty of
Arts and Sciences in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy

University of Pittsburgh

2011

UNIVERSITY OF PITTSBURGH

ARTS AND SCIENCES

This dissertation was presented

by

Erhan Demircioglu

It was defended on

June 10, 2011

and approved by

Robert Brandom, Ph.D., Department of Philosophy

Peter K. Machamer, Ph.D., History and Philosophy of Science

John McDowell, M.A., Department of Philosophy

Dissertation Advisor: Anil Gupta, Ph.D., Department of Philosophy

Copyright © by Erhan Demircioglu

2011

AT THE GATES OF CONSCIOUSNESS: PHYSICALISM AND PHENOMENAL CONCEPTS

Erhan Demircioglu, Ph.D.

University of Pittsburgh, 2011

In this work, I have three related aims. First, I attempt to show that none of the popular responses to the Knowledge Argument works. After presenting the Argument as it is originally proposed by Jackson and Nagel, I clarify the distinctions between two proposals, and identify the core of the challenge it raises for physicalism. After that, I show that “no-new-knowledge” and “new-knowledge/old-facts” responses to the Argument fail.

My second aim is to develop an adequate account of phenomenal concepts. A widely held and influential idea is that phenomenal concepts, concepts under which we grasp the qualitative properties of our experiences, are *epistemically and semantically special*. I provide an account that, I believe, captures the core of this special nature. On my account, phenomenal concepts are epistemically special in that our beliefs formed by the application of phenomenal concepts to our experiences are *non-inferentially* justified, and they are semantically special in that they are *simple* concepts that represent certain phenomenal properties without using any property other than those properties. I also draw the logical implications of my account of phenomenal concepts with respect to the conceptual/non-conceptual distinctions, the “Myth of the Given”, and the infallibility of phenomenal beliefs.

My third aim is to provide a novel argument for property dualism on the basis of my account of phenomenal concepts. Instead of asking “how are phenomenal experiences possible in

an entirely physical world?”, I raise the question “how are phenomenal concepts possible in an entirely physical world?” This question accepts the invitation of our physicalist to ascend from the level of properties to the level of concepts but inquires whether such an ascent is of any help to the physicalist. I argue that ascending to higher levels only displaces the problem: phenomenal concepts are as problematic for physicalism as are phenomenal properties. I conclude that phenomenal concepts are not possible in an entirely physical world and, since we have phenomenal concepts, our world cannot be entirely physical.

TABLE OF CONTENTS

PREFACE.....	VIII
1.0 INTRODUCTION.....	1
2.0 THE KNOWLEDGE ARGUMENT	6
2.1 JACKSON’S MARY	7
2.2 WHAT MARY LEARNS: A CLOSER LOOK	13
2.3 MARY VERSUS MARIANNA	17
2.3.1 Two Knowledge Arguments?	19
2.3.2 The Generic Argument: Marianna Wins	24
2.4 THE KNOWLEDGE ARGUMENT: NAGEL’S BATS	28
3.0 DOES MARY GAIN ANY NOVEL KNOWLEDGE?	36
3.1 DENNETT AND “FIGURING OUT” WHAT IT IS LIKE.....	37
3.2 LEWIS AND THE ABILITY HYPOTHESIS	48
4.0 THE PHENOMENAL CONCEPT STRATEGY.....	58
4.1 INTRODUCTION	58
4.2 THIN ACCOUNT OF PHENOMENAL CONCEPTS	63
4.2.1 Two Types of Direct Reference	66
4.2.2 The Case of Blindsight	69
4.2.3 “The Phenomenological Intuition” Misconceived	73

4.3	THICK ACCOUNT OF PHENOMENAL CONCEPTS	76
4.3.1	The Fundamental Claim	78
4.3.2	Problems with the Thick Account.....	81
4.4	CONCLUSION	86
5.0	ACQUAINTANCE AND PHENOMENAL CONCEPTS	88
5.1	PHENOMENAL CONCEPTS, STATES, AND QUALITIES	88
5.2	ACQUAINTANCE AS AN UNMEDIATED EPISTEMIC RELATION.....	91
5.3	PHENOMENAL CONCEPTS	101
5.4	THE CONCEPTUAL/NON-CONCEPTUAL DISTINCTION(S).....	108
5.5	THE “MYTH OF THE GIVEN”	118
5.6	THE INFALLIBILITY OF PHENOMENAL BELIEF	123
6.0	AN ARGUMENT FOR PROPERTY DUALISM.....	127
6.1	ASCENDING LEVELS.....	130
6.2	A POSSIBLE REJOINDER	133
6.3	WHERE IS WHAT-IT-IS-LIKE?.....	135
6.4	CONCEPTS, PROPERTIES, AND CARTESIANISM	136
7.0	CONCLUSIONS (STILL MORE PROMISSORY NOTES).....	140
7.1	THE MYSTERIES OF DIRECT EPISTEMIC RELATION	140
7.2	LOGICAL ATOMISM	143
7.3	CONCEPTUALIZATION	148
	BIBLIOGRAPHY	153

PREFACE

I am grateful to my dissertation advisor Anil Gupta. I have benefited more than I can put into words from his gentle guidance, and I have witnessed first-hand his art of being critical in a constructive way. Without the discussions I had with Anil, this work would be poorer than it is. I would also like to thank Robert Brandom and John McDowell, whose presence and works have always been an invaluable motivation and inspiration to me. Finally, my thanks go to Peter Machamer who kindly accepted to be a member of my dissertation committee.

I dedicate this dissertation to my love, friend, and wife, Duygu: her caring attitude was an unending source of encouragement and sustenance to me.

1.0 INTRODUCTION

There is something it's like for us to be in certain mental states such as smelling a skunk, tasting a lemon, seeing red, hearing a symphony, and feeling pain. These experiential or phenomenally conscious states are essential to our lives: without them, our lives would be irredeemably impoverished. A central question in the philosophy of mind is whether there is any room for conscious experience in an entirely physical world. The question is difficult because consciousness is *prima facie* non-physical: the distinctive type of feeling associated with, say, smelling a skunk appears to be so radically different from anything physical going on in my body (or my environment) that any claim about the latter's being identical to the former appears to commit a category mistake.

In the second chapter of my dissertation, I focus on Frank Jackson's *Knowledge Argument*, which begins with the premise that complete physical knowledge of the world does not provide us with complete knowledge about our experiences and goes on to conclude that physicalism is false. I clarify the argument by distinguishing between recognitional and substantial aspects of the knowledge one gains through a certain experience. I argue that the substantial aspect can be expressed by the phrase "it is like Q to have E", where E is a certain experience and Q is the pertinent quality with which one gets acquainted through E. It is this substantial aspect of the knowledge gained, and not the recognitional aspect, that, I claim,

renders the Knowledge Argument problematic for physicalism. The chapter ends with an elaboration of Thomas Nagel's challenge to physicalism from which an important insight is distilled: *phenomenal* concepts (e.g., *experience of red*, *lemony taste*, *skunkish smell*, *excruciating pain*) are essential to an understanding of the nature of experiences.

There are physicalist views which reply to the Knowledge Argument by rejecting its premise. Two versions of this sort of reply are discussed in chapter three. On one version, defended by Daniel Dennett, what one knows when one knows what an experience is like is just a *very* rich physical knowledge. This "no-new-kind-of-knowledge" version holds that from complete physical knowledge, one can deduce what experience is like. Contra this, I argue that we *know* that the knowledge we have of our experiences is not just physical knowledge rich-and-condensed: no amount of physical information captures the feeling of pain. On another version, called *the Ability Hypothesis* which is defended by David Lewis, what one knows when one knows what an experience is like is a sort of knowing-how or ability like recognition and discrimination. This version grants the irreducibility of phenomenal knowledge to physical knowledge but proposes to construe it as an instance of the allegedly general irreducibility of knowing-how to knowing-that. I argue against the Ability Hypothesis that some of the abilities one gains after having a certain experience such as the ability to entertain *new thoughts* are conceptually-driven abilities which require an antecedent possession of concepts, and this shows that not all phenomenal knowledge is knowing-how since concepts are characteristic ingredients of knowing-that. I conclude that both versions fail to recognize what I called *Nagel's Insight* in chapter chapter: the centrality of a distinctive sort of concepts of our experiences (phenomenal concepts) to our understanding of their nature.

Chapter four addresses the physicalist strategy of resisting the Knowledge Argument through an acceptance of an irreducible category of phenomenal concepts. This strategy, which I call, following Stoljar, the *Phenomenal Concept Strategy* (PCS), accepts Nagel's Insight and offers an account of phenomenal concepts. PCS comes in two varieties. On the thin account, phenomenal concepts are "blind" concepts which pick out their physical references directly, i.e., without any kinds of modes of presentation associated with the concept. I argue that the thin account cannot accommodate the idea that introspection of our conscious states is what grounds our applications of phenomenal concepts. On the thick account, the mode of presentation of a phenomenal concept is the quality to which it refers. Appreciation of an unusually intimate sort of relation between phenomenal concepts and their referents is a virtue of the thick account; however, it does not, I argue, sit well with an antecedent commitment to physicalism. The thick account consists of three claims: (i) phenomenal concepts refer to physical/functional properties, (ii) phenomenal concepts are isolated from all physical/functional concepts, and (iii) phenomenal concepts are substantive concepts whose modes of presentation are constituted by their references. I argue that (i)-(iii) are not a happy triad.

A moral of my discussion of physicalist responses to the Knowledge Argument is that a theory of our epistemic access to our own experiences is required by a satisfactory account of phenomenal concepts. A subject's perceptual phenomenal states (i.e., the phenomenal states the subject is in when he or she perceives external objects) provide reasons for holding beliefs about external objects. But what reason does the subject have for holding beliefs about those states themselves? In chapter five, I defend the claim that phenomenal beliefs are immediately justified. I argue that one's epistemic relation to one's experiences is direct in that one's beliefs about those experiences are (*prima facie*) justified by merely having those experiences, while

one's beliefs about external objects are epistemically based on the effects they have on one's mental states. Following Russell, I call the special type of epistemic relation subjects bear to their phenomenal states *acquaintance*, and I argue that some of the philosophical mystery surrounding the notion can be dispelled if a distinction is drawn between acquaintance and knowledge by acquaintance. Furthermore, an account of acquaintance enables us to formulate an adequate account of phenomenal concepts. I argue that phenomenal concepts are simple in that they represent phenomenal properties without using any property other than those properties, and also that they are the only simple concepts we human beings possess. The chapter ends with an elaboration of some of the implications of my account of phenomenal concepts, which are these: (1) that experiences have contents the having of which does not require the possession of any concepts, (2) that contents of experiences are fully conceptualizable, (3) that my account of phenomenal concepts is not an instance of what Sellars calls *the Myth of the Given* in at least one of its standard interpretations, and finally (4) that there is a sense in which our knowledge of our experiences is infallible.

Drawing on the material from previous chapters, chapter six proposes a novel argument for property dualism. Instead of asking "how are phenomenal experiences possible in an entirely physical world?", I raise the question "how are phenomenal concepts possible in an entirely physical world?" A physicalist might answer the first question by arguing that the what-it-is-like aspect is something to be associated with concepts of experiences rather than with experiences themselves, and thus resisting property dualism through an embrace of conceptual dualism. The second question accepts the invitation of our physicalist to ascend from the level of properties to the level of concepts but inquires whether such an ascent is of any help to the physicalist. I argue that ascending to higher levels only displaces the problem: phenomenal concepts are as

problematic for physicalism as are phenomenal properties. I conclude that phenomenal concepts are not possible in an entirely physical world and, since we have phenomenal concepts, our world cannot be entirely physical.

2.0 THE KNOWLEDGE ARGUMENT

Arguments against one materialist theory or another are never very conclusive. It is always possible to adjust the details. But the knowledge argument, if it worked, would directly refute the bare minimum that is common to *all* materialist theories.

David Lewis, "What Experience Teaches"

A well-known and much discussed argument against physicalism is what is known as "the Knowledge Argument" in the literature.¹ In this chapter, I proffer a detailed analysis of the argument, an analysis by which I hope to clearly state and explain the premises of the argument and critically evaluate possible misunderstandings about its structure and the metaphysical conclusion it is purported to establish. A through analysis of the argument will function as groundwork for the later parts of this work, and is designed to provide a solid position from which we can distinguish the crucial differences among various physicalist replies to it.

¹ There are some other arguments developed against physicalism: the Conceivability Arguments (that is, the Zombie Argument and the Inverted Spectrum Argument) which are clearly presented and thoroughly argued for by Chalmers (1996), and the Modal Argument which is forcefully introduced by Kripke (1980). Starting with one of these arguments instead of another is almost always a matter of personal taste. But I think the Knowledge Argument has a certain advantage over other forms of anti-materialist arguments. At least on the face of it, the Knowledge Argument presents a very clear case against physicalism without necessarily invoking heavily-loaded discussions about the intricate relations between conceivability and possibility and between contingency and necessity. I take its straightforwardness in this sense as one of its virtues. Moreover, certain physicalist responses to the Knowledge Argument can without much effort be translated as responses to the Conceivability and Modal Arguments. Given its broad implications, a study of the Knowledge Argument (and possible responses to it) is thus not merely a study of one of the arguments against physicalism. My reason for taking the Knowledge Argument as my central focus is that it motivates an apparatus of phenomenal concepts which sheds some light on the question of why consciousness is, as Chalmers coins the term, "the hard problem." Besides, though the Knowledge Argument is central, I will make references to other arguments whenever convenience suggests.

2.1 JACKSON'S MARY

Although the generic intuition on which the argument rests has been voiced by several philosophers,² it is generally accepted that the Knowledge Argument has been presented in its most provocative and direct form in a paper of Jackson's named "Epiphenomenal Qualia." In this (1982) paper, Jackson invites us to consider a perfect scientist with perfect logical skills called Mary who knows everything there is to know about the physics of sensory perception of red. She knows everything about the relevant surface reflectance properties of those objects that their fellow citizens call "red," the physical processes through which light is reflected by those objects and reaches their eyes, and the physical processes that occur in their heads when they perceive red things. However, Mary has never experienced as of red, and she gains her complete physical knowledge in a black and white room, studying only those books which are all printed black and white, and watching an old-fashioned black and white television. One day, the physically omniscient Mary decides to get out of her room in the pursuit of having experiences of things her fellow citizens call "red." And there it is. Mary learns something new about the experiences of others when they see red things. She now knows what it is like to have red experiences,³ though before she lacks it.

² See, for instance, Broad (1925), Dunne (1958), Feigl (1958, 1967), Meehl (1966), Robinson (1982), Russell (1912). The introduction written by Stoljar and Nagasawa to *There's Something About Mary* (2004) gives a nice summary of philosophical accounts sharing the dualist intuition with Jackson's Knowledge Argument.

³ A terminological point. As I use the term, "a red experience" means "an experience as of seeing something red." I take it as a *prima facie* plausible hypothesis that the sense in which experiences are red is not the sense in which physical objects are red. So, phenomenal redness is different from physical redness. The claims that there is no

How are we to account for the apparently new piece of knowledge Mary acquires? Jackson argues that since, by hypothesis, Mary's physical information about sensory experiences as of red is complete, the only plausible way to explain our intuition as to Mary's new piece of information is to acknowledge that, by undergoing the relevant experience, she gets acquainted with some non-physical phenomenal properties or qualia.

A convenient way of presenting the Knowledge Argument is this:

- (1) Before her release, Mary has all physical knowledge about red experiences.
- (2) After her release, Mary will learn some new knowledge about red experiences.

From (1) and (2),

- (3) Before her release, Mary does not have all knowledge about red experiences.
- (4) Therefore, there is non-physical knowledge about red experiences.

We also independently have:

- (5) If physicalism is true, then complete physical knowledge is complete knowledge *simpliciter*.

From (4) and (5),

- (6) Physicalism is false.

There is no question regarding the validity of the argument. It is plain that given that physicalism implies that all knowledge is physical knowledge and Mary who has all physical knowledge gains non-physical knowledge after her release, physicalism is false. The central problems with the argument turn on how one should interpret some of the notions included in the premises like "physical," "physicalism," and "knowledge," but not on its formal validity.

phenomenal redness and, more generally, there are no phenomenal properties are non-trivial claims that need to be argued for. Representationism in its various forms is a currently popular way of philosophically substantiating these claims.

Moreover, the premises in the Knowledge Argument enjoy at least *prima facie* plausibility. The first premise appears to be true because it seems that being in an entirely monochromatic environment and lacking certain types of color experiences do not present any principled obstacles to the acquisition of the completed physicalist story about human color vision. As a general rule, a physicalist account of things in the world does not presuppose the presence of certain types of sensory experiences in order to be fully comprehended, and there seems to be little reason, if any, to think that a physicalist account of sensory experiences themselves is an exception. Moreover, it surely appears that Mary will learn something new about color experiences after her release. Intuitively, post-confinement Mary learns some sort of new information or knowledge. After she has her first visual experience of ripe tomatoes, it seems that she will be surprised and realize how impoverished her conception of red experiences has been all along.

And, finally, the truth of physicalism appears to imply that all knowledge is physical knowledge. For if there is a piece of knowledge even someone with complete physical knowledge and perfect logical acumen may lack, then one may plausibly say that there is a certain fact about the world that that piece of knowledge is about. Given that complete physical knowledge exhausts all physical facts, then the fact that the piece of knowledge in question is about is a non-physical fact. And this is just what physicalism denies.⁴

It has been pointed out by various philosophers that the specific conditions Jackson originally builds into his version of the argument are not sufficient to guarantee that Mary does

⁴ Another way to argue for the idea that physicalism implies that all knowledge is physical knowledge is given by Jackson: “Physicalists must hold that complete physical knowledge is complete knowledge simpliciter. For suppose it is not complete: then our world must differ from a world, $W(P)$, for which it is complete, and the difference must be in nonphysical fact; for our world and $W(P)$ agree in all matters physical. Hence, physicalism would be false at our world” (1986/2004, p. 51). Though the two arguments are very similar, an advantage of the one I give above over Jackson’s version is that the main assumption for both arguments – that is, the kind of connection alluded to between pieces of knowledge and facts – is stated more clearly in the former.

not have any experiences of red before her release. For the beginning, not only the walls of Mary's room but also her skin should be painted black and white. Moreover, the argument needs to be formulated in such a way as to guarantee that the relevant parts of Mary's color vision system do not atrophy from disuse. Another condition to be added is that in her room, Mary does not rub her eyes so that the possibility of her having after-image and phosphene experiences of red is excluded. (I will later argue that there is a version of the Knowledge Argument which shows that even when we grant that *before* her release Mary has visual experiences of red (e.g., certain after-images) and thus knows what it is like to experience as of red, there are remaining problems with physicalism.) In order to bypass some of these complications, it is sometimes suggested that Mary is color-blind from birth, and just before her release, she goes through a surgical operation after which she gains the capacity to have polychromatic color experiences. Whatever their intricacies are, it seems plausible to take for granted that nothing of substantial importance hangs on the complications that can in principle be added to the argument.

The Knowledge Argument plausibly generalizes in two directions. First, there is nothing relevantly special about Mary on which the premises of the Argument rest, and therefore one can replace Mary with any being which has the specific cognitive capacities required to possess the completed physical information about the experience of red and to figure out the logical implications of that information. It does not matter for the argument to proceed whether the individual in question who is confined in a black-and-white environment and subject to only visual experiences of monochromatic type is a creature possessing the recognitional and imaginative abilities typical of a human being or a visitor from a distant planet, a Martian scientist perhaps, whose corresponding abilities are of an entirely different kind. If the argument had been originally presented by Jackson so as to clearly rely upon the latter possibility, then it

would be close to Farrell's (1950) and Feigl's (1967) and, in a roundabout way, Nagel's (1974) formulations of it – about the latter of which I will make certain comments later. But nothing essential about the soundness of the argument draws upon this distinction.

Second, there is nothing relevantly special about the visual experience of red for the argument to proceed and therefore analogous arguments can be constructed for other types of visual experiences, e.g., the experience of green, yellow, and blue. Moreover, it seems that the soundness of the argument is not confined solely to visual experiences and therefore analogous arguments can be constructed for other types of perceptual experiences including, e.g., auditory, tactile and olfactory experiences, and also for bodily sensations such as pains, tickles, itches, and headaches. If the argument is good, it shows that phenomenal consciousness as a whole is left out by a complete physicalist account of the world.

A further point that deserves some clarification at the outset concerns what exactly the universal quantification “all” amounts to when we say Mary has all physical knowledge. It admits three different interpretations. Under the first interpretation, what Mary knows when she has all physical knowledge comprise only *general* facts (i.e., physical laws) – that is, she knows all and only the relevant physical science.⁵ On this construal, the knowledge Mary obtains in her room is of the form “whenever the light of a certain frequency X which is reflected from the surface of an object having the property Y meets the human eye, the following occurs in the human brain:____” where X and Y are physical properties of light and the surface in question respectively and the blank is filled with a universally quantified proposition expressed by

⁵ The kind of general scientific knowledge directly relevant to the Knowledge Argument is the knowledge of physical laws pertaining to human perception of red things. But, the Knowledge Argument would still have its intuitive force if we took it that Mary has the knowledge of all physical laws, not only laws about human perception of red things and not only laws about human perception *per se*, there are in the actual world. Nothing essential to the Knowledge Argument turn on how one makes this choice.

mentioning certain neurophysiological properties. Under the second interpretation, what Mary knows when she has all physical knowledge consists of *particular* physical facts, which have either the form of singular statements or the form of existentially quantified statements. On this construal, Mary knows every particular physical thing that is going on in the actual world. Finally, under the third interpretation, what Mary knows when she has all physical knowledge consists of general physical laws *and* particular physical facts. The third interpretation is equivalent to the second one given that fixing the totality of particular physical facts necessarily determines the totality of physical laws, and is stronger than the second if otherwise.⁶

In what follows, I will deploy the third interpretation. The central intuition that supports the Knowledge Argument, which is that Mary will gain a new piece of information after her release, persists even when the relevant universal quantification is taken in its strongest sense. So, there seems to be no reason to favor a weaker premise which corresponds to either the first or the second interpretation, and thereby to diminish the strength of the conclusion to be derived from the argument. Besides, anti-physicalists are generally willing to make the strongest case possible against physicalism and maintain that the argument goes through even when we take the third interpretation.⁷

The intuitive force of the Knowledge Argument needs to be granted. It certainly appears that nothing of a physical sort captures what are individually characteristic to the visual

⁶ Jackson seems to employ the first interpretation in his original article “Epiphenomenal Qualia” (1982) and the third interpretation in his later work “What Mary Didn’t Know” (1986/2004). In (1982), Jackson writes: “She [Mary] specializes in the neurophysiology of vision and acquires, let us suppose, all the physical information there is to obtain about what goes on when we see ripe tomatoes, or the sky, and use terms like ‘red’, ‘blue’, and so on. She discovers, for example, just which wave-length combinations from the sky stimulate the retina, and exactly how this produces *via* the central nervous system the contraction of the vocal chords and the expulsion of air from the lungs that results in the uttering of the sentence ‘The sky is blue’.” (p. 130). And, in his (2004), he says: “She knows all the physical facts about us and our environment, in a wide sense of ‘physical’ which includes everything in *completed* physics, chemistry, and neurophysiology, and all there is to know about the causal and relational facts consequent upon all this, there is to know” (p. 52, *italics original*).

⁷ See, for instance, Chalmers (1996, p. 33).

experience of red, the smell of a rose, the taste of a lemon, or the hurtfulness of pain – their respective phenomenal characters. Similarly, it certainly appears that Mary will learn something new about experiences after her release. The apparent force of the argument may be an illusion which can be explained away by showing that it rests on a failure to truly appreciate what it is to know everything physical there is to know or a subtle equivocation on the meanings of the notions “knowledge” and “physical” or something to a similar effect. Showing that its apparent force is *merely* apparent is the burden of the physicalist.

2.2 WHAT MARY LEARNS: A CLOSER LOOK

A physicalist may be inclined to offer the following rough-and-ready rejoinders to the Knowledge Argument. First, even though the phenomenal information Mary comes to apprehend after her release cannot in principle be gained in her confinement days, this raises no special problems for physicalism simply because Mary’s experience of red is a *new* particular fact and since the totality of physical knowledge Mary obtains in her room does not by itself enable her to foresee which particular facts will take place in the future, there is nothing problematic in the idea that she learns something new when a new physical fact occurs: surely, our physicalist would add, she cannot know what it is like for her to have red experiences given that there is no relevant fact to be known in the first place.

Secondly, our physicalist might go on to argue, what Mary learns after confinement includes an indispensable indexical element: what she comes to learn at the first instance can be

conveniently taken to be what the sentence “this is what it is like for me to have my present experience” expresses, which includes two evidently indexical components expressed by ‘me’ and ‘this’, the latter of which refers to the particular red experience she has when she looks at ripe tomatoes. However, there are familiar problems with capturing the cognitive content of an indexical statement in non-indexical terms which do not by themselves have any metaphysical implications.⁸ Therefore, the physicalist concludes, there is no challenge whatsoever for physicalism in the idea that there are pieces of indexical information which cannot in principle be read off from Mary’s physical knowledge which is expressed in non-indexical language.

However tempting they may initially seem, these attempts at quickly refuting the Knowledge Argument do not work. The first worry raised by the physicalist can be answered in two different ways depending on the range of knowledge Mary has in her room. One can think of the physical knowledge she has as consisting of either all particular physical facts which take place only prior to her release or all particular physical facts which take place both prior to her release and after her release. In the former case, we can say that her knowledge of particular facts combined with her exhaustive knowledge of physical laws enables her to make veridical predictions about the physical characteristics of the particular experience she will have when she looks at ripe tomatoes: she can give full physical descriptions of the brain states of normal perceivers when they have red experiences, and she can explain in a deductive-nomological manner why those descriptions are fulfilled by pointing out the initial physical conditions that obtain and the physical laws that are operative. It then seems that she will also know the state her brain would be in if she were subjected to the objects that have the surface reflectance properties of the objects her fellow citizens call “red”: her knowledge about the physical characteristics of

⁸ The cognitively distinct features of indexical knowledge are brought to attention in the now-classical papers on the topic including Castaneda (1967), Perry (1979) and Lewis (1979).

her brain and the physical characteristics of red objects combined with her knowledge of physical laws that relate those two sets of characteristics gives her the information about the physical state her brain will be in when she sees red objects. But still she does not know what it will be like for her to have experiences of red in her confinement days.

Things get easier if we simply suppose that Mary's knowledge of particular facts is exhaustive across time. On this reading, she knows every particular physical fact there is to know in the actual world irrespective of the time it takes place. That is, she has complete physical knowledge of the actual world in the fullest sense of the term. Then, under this construal, she already knows the physical characteristics her brain instantiates after she leaves her room and sees red objects. However, this knowledge does not seem to give her the information of what it will be like for her to see red objects. Hence, the intuitive force of the Knowledge Argument remains unaffected.

The second objection from the physicalist concerning the indexicality of new knowledge Mary acquires after her release deserves a more sustained attention, which I will attempt to provide in the following parts of this work. For now, let me point out a disanalogy between Mary's epistemic deficiency with respect to phenomenal knowledge and typical cases of being short of indexical knowledge. Intuitively, a person might know that Frederic Jameson is the author of *Marxism and Form* even though he does not know that *that* [demonstrating Frederic Jameson in a certain context of utterance] is the author of *Marxism and Form*. Why do we not think that his lacking the knowledge that *that* is the author of *Marxism and Form* does not have metaphysical implications? Why do we not think that the fact that Jameson is the author of *Marxism and Form* is ontologically distinct from the fact that *that* is the author of *Marxism and Form*? The reason is, I submit, that we think that what sentences containing demonstratives are

about – their worldly contents as one may call them – are not essentially tied to particular features of certain contexts of utterances: what they are about can be expressed by sentences containing only non-demonstrative expressions. Surely, that *that* is the author of *Marxism and Form* is not, ontologically speaking, a fact only those who have an antecedent grasp of the features of the context in which the sentence ‘that is the author of *Marxism and Form*’ is uttered can understand. However, how can one express the worldly content of Mary’s thought that *this* is what it is like to see red in a way that can be understood by those who do not have a prior grasp of the features of the experiential context in which it is tokened? Try to do it as much as you please, I bet you will face an insurmountable difficulty. The most promising attempt on this score seems to be to identify what ‘*this*’ picks out with a physical property and then argue that a sentence containing a physical concept referring to that property can be understood even by someone who is incapable of having red experiences, hence the analogy between two cases. But still the intuition persists: what one lacking the relevant experiential perspective understands when one understands the said non-indexical physical sentence is essentially incomplete with respect to what-it-is-likeness of red experiences in a way which what one who does not know the contextual features of the demonstrative utterance about Jameson understands when one understands the corresponding non-indexical sentence is not with respect to the relation between Jameson and the book *Marxism and Form*. This suggests that Mary’s unavoidable usage of demonstratives to express her new knowledge feeds the wrong impression that her epistemic progress after her release is solely a matter of gaining some indexical knowledge, and consequently that her lack of phenomenal knowledge cannot be modelled on the standard cases of lacking indexical knowledge.

2.3 MARY VERSUS MARIANNA

Some misunderstandings regarding the nature and content of the new knowledge Mary gains after her release can be avoided by recognizing different epistemic stages one might undergo in the process of obtaining information about experiences. When it comes to pointing out these differences, Nida-Rumelin's Marianna case (1996) is more helpful than Jackson's Mary. Like Mary, Marianna lives her entire life in a monochromatic environment. The central difference is, however, that when the happy day comes, rather than seeing ripe tomatoes and bananas and grass and the sky, she is randomly visually presented with four slides showing clear cases of blue, red, green and yellow but she is not told the names of the colors. Now Marianna does not know which of the four types of color experiences she has is, say, red nor does she know that having red experiences is like *that*, where the indexical in question picks out the type of experience she has when presented with the red slide. But still one can plausibly say that now that she has the experience of red, there is a clear sense in which she knows what it is like to see red. Marianna's gaining this piece of information is analogous to the case in which someone gets acquainted with Jameson without recognizing him as Jameson and hence without knowing that, referring to him, *that* is Jameson. However, as claimed above, the analogy breaks down when pressed further.

There may be worries about Marianna's knowing what it is like to see red. One may say that she has no justification *from the inside* for attributing herself the knowledge of what it is like to see red: she does not know that she knows what it is like to see red. And, if knowing that one knows p is a necessary condition for one's knowing p (the KK principle),⁹ then Marianna does

⁹ For a defense of the KK principle, see Hintikka (1970). As a side note, let me say that I myself find the KK principle particularly suspect. Suppose that the principle is true. Then, one's knowing that p requires one's knowing that one knows that p . But, one's knowing that one knows that p also requires that the truth of one's knowing that p

not know what it is like to see red. After all, she does not know whether or not what she is shown are orange, violet, purple, and ochre (or any other combination of four colors) instead of red, green, blue, and yellow. What she knows is only that she has the visual experiences of four polychromatic colors. Even though I think that there is still a perfectly cogent sense in which Marianna knows what it is like to see red, I am ready to slightly revise the thought experiment to obviate the concerns regarding the relevance of internal justification for knowledge. Suppose that Marianna is told that the four colors she is presented with are red, green, yellow, and blue but she is not told which one is which. In this case, she knows that she knows what it is like to see red. What she does not still know, however, is that, pointing at her red experience, *that* is a red experience.

But one may press even further: there is a sense in which Marianna does not know what it is like to see red because she cannot identify red experiences as red experiences and knowing what it is like to see red requires such a recognitional ability. When she experiences the red slide, she is not in a position to recognize that she experiences red and justifiably verbally report “that is red.” Does this mean that she does not know, in any epistemically respectable sense of the term, what it is like to see red? No, it only suggests that I should specify more carefully what I mean when I claim that after her red experience, Marianna knows what it is like to see red. What I mean is this: Marianna knows what it is like to have *this* experience, where the indexical picks out her red experience, and her experience is a red experience. (Compare: I meet Jameson, spend some time together talking over the intricacies of the Marxist theory but I do not know that

given that knowledge requires the truth of its content. There is a sort of circularity here, but I am not sure how serious an objection this is against the KK principle. Moreover, consider a person who lacks self-consciousness: he does not know that he is himself. Then, the KK principle seems to imply that this person cannot know anything because he cannot have the *de se* knowledge that he himself knows that *p*, whatever *p* is. This is a highly counterintuitive result but yet again I am not sure whether and how the KK principle can block this result or whether, if otherwise, it can live with it.

I am talking with Jameson. There is a sense in which I know the man I meet even if I lack the recognitional knowledge that, pointing at the person I meet, *that* is Jameson.)

Marianna's case shows that there is an intermediate epistemic stage which goes unnoticed in Mary's case. It seems obvious that Marianna learns something new when she is haphazardly shown different colors and hence she makes a sort of epistemic progress. However, she still misses various pieces of information the pertinent one of which is the information that *this* is what it is like to see red, where the indexical at hand picks out the type of experience she has when she is presented with the red slide. There is still room for Marianna to make further epistemic progress because she experiences red without knowing that what she experiences is red.

2.3.1 Two Knowledge Arguments?

An interesting question about Marianna's case is whether the epistemic progress she makes gives her the relevant item of knowledge which is crucial for the Knowledge Argument to work. After having the relevant visual experiences, Marianna evidently gains epistemic access to various new thought contents that are not available to her before. She is now in a position, for instance, to wonder whether ripe apples appear like *that* or *that*, where the indexicals respectively refer to the colors of the red and blue slides, and to entertain new hypotheses and make new guesses. What explains the fact that Marianna gains epistemic access to new thought contents is that she now knows what it is like to see red, blue, etc. Can an argument analogous to the original Knowledge Argument from Mary's case work in Marianna's case? Or does the property dualist need a

further step of epistemic progress to make a viable case for the intended metaphysical conclusion – the falsity of physicalism?

Despite her new ability to entertain the corresponding thought, Marianna does not come to know that ripe apples appear like the red slide she has the experience of. Nevertheless does she learn any item of knowledge? After her color experiences, she knows what it is like to see red; but it is not clear whether this knowledge poses any threat to physicalism because it is not clear that this knowledge has any propositional content, which can be expressed in a propositional form by a suitable that-clause. According to Nida-Rumelin, Marianna “has not gained any new item of the relevant propositional knowledge” (2004, p. 254) and hence there is nothing imperiling physicalism at this stage of her epistemic progress. After all, one may ask what the relevant proposition can possibly be given that she does not know *that that is what it is like to see red*, where the indexical in question picks out the relevant color experience.

However, contra Nida-Rumelin, the following construal of Marianna’s knowledge regarding what it is like to see red conveniently adapted from Lycan (1996, p. 92) appears plausible:

(P) Marianna knows that it is like *Q* to see red.

Here the expression ‘*Q*’ picks out the pertinent quality Marianna gets acquainted after having the red experience. Thinking of Marianna’s new knowledge in terms of (P) is supported by certain syntactic and semantic relations between the sentential constructions “S knows wh-...” and “S knows that...” For instance, “Jack knows who Mary is” is true in virtue of Jack’s knowing that Mary is so-and-so, and “Susan knows where Hector is” is true in virtue of Susan’s knowing that Hector is in such-and-such place. A plausible idea is that Marianna’s knowledge can be viewed as a special case which can be assimilated to this general scheme.

One may raise the worry about the proposal above that it does not work if Marianna does not introduce a linguistic expression that stands for the relevant quality in question. However, just as Susan may know that Hector is in such-and-such place without introducing a name for him, Marianna may know that it is like *Q* to see red without introducing a linguistic expression that stands for *Q*. Surely, in either case, a mental item that stands for the relevant entity needs to be present in the subject's mind but no ordinary linguistic expression which is antecedently tokened is required. The symmetry between the two cases thus goes unaffected.

Furthermore, it may be objected that (P) can be true even when Marianna does not know what it is like to see red and hence the paraphrasing attempt at hand fails. One may argue that Marianna already knows something relevantly similar to (P) before she experiences red since she already knows that it is like *Q* to see red, where '*Q*' stands for the experience of seeing ruby, even though she does not know what it is like to see red.¹⁰ However, the locution "it is like" does not mean "it resembles" within the context of the relevant that-clause just as the locution "it is like" does not mean "it resembles" in the context of "what it is like to see red."¹¹ Notice that if "what it is like" meant "what it resembles," then contrary to what the objection says, Marianna would also know what it is like to see red before she experiences red. The sense in which Marianna does not know what it is like to see red is identical with the sense in which she does not know that it is like *Q* to see red, where '*Q*' is properly understood. The objection at hand equivocates on the locution "it is like" in the contexts of the relevant that-clause and "what it is like," and thus fails to draw a wedge between the two cases.

¹⁰ Tye (2002, p. 8) rejects Lycan's proposal along these lines.

¹¹ Nagel writes: "[The] analogical form of the English expression "what it is *like*" is misleading. It does not mean "what (in our experience) it *resembles*," but rather "how it is for the subject himself.'" (1974, p. 440, fn. 6). Reading "what it is like" as "what it resembles" is what Lewis (2004) calls "the first way to miss the point" of the Knowledge Argument among the six ways he specified.

It is plausible to say that (P) captures what Marianna comes to know after having red experiences. Then, an interesting result appears to follow. There are indeed two knowledge arguments, rather than one, which can be distinguished with reference to the two items of knowledge that someone situated in a monochromatic environment appears to inescapably lack. First, such a person lacks the piece of knowledge expressed by the *that*-clause in (P): the knowledge that it is like *Q* to see red. The first Knowledge Argument then runs roughly like this: One who knows everything physical there is to know may still lack the knowledge that it is like *Q* to see red, and hence physicalism is false. Second, she does not also have the knowledge that *that* is what it is like to see red, where the indexical picks out an antecedent red experience of her own. The second Knowledge Argument then goes roughly like this: One who knows everything physical there is to know may still lack the knowledge that *that* is what it is like to see red, where the indexical picks out an antecedent experience of her own, and hence physicalism is false.

Having the second piece of knowledge mentioned above implies having the first but not *vice versa*. It is clear that, just like the first piece of knowledge, the second piece of knowledge requires having an antecedent experience of red since the indexical within the relevant *that*-clause picks out an antecedent experience of the subject. But there is more to the second item of knowledge than what is required to have the first. After her experience of the red slide, Marianna knows that it is like *Q* to see red but does not know that *that* is what it is like to see red, where the indexical picks out her red experience, while after her experience of ripe tomatoes, Mary knows both. The difference between the two stems from the fact that unlike Mary, Marianna is not in a position to recognize her red experience-token *as* an instance of the red experience-type because such a recognition requires either having experiences of paradigmatically red objects like ripe tomatoes or being told by others that pointing at the red slide, *that* is red.

A suggestive model which explains how Mary happens to recognize her experience of red as an experience of red goes like this: when Mary sees ripe tomatoes, she learns that it is like *Q* to see red. What is common to the cases of Mary and Marianna is this stage of epistemic progress. However, Mary also knows that what she sees are ripe tomatoes (after all, she can tell by the way they look to her), and combined with her knowledge that her perceptual apparatus works normally and ripe tomatoes induce red experiences in normal perceivers, she infers that *that* is what it is like to see red, where the indexical picks out her experience.¹² Hence, she gains an item of knowledge which Marianna lacks.

Mary's knowledge that her experience when she sees ripe tomatoes is of red cannot be separated from her knowing that others (normal perceivers) have experiences of the same type when they see ripe tomatoes and that they call them "red." In other words, Mary's knowing that this is what it is like for her to see red is inextricably bound up with her knowing that this is what it is like for others (normal perceivers) to see red, where the indexical 'this' picks out her experience in both usages. The tight connection between two pieces of knowledge can be thought of as a sort of broadly Wittgensteinian insistence on the constitutive role of social constraints for the proper usages of linguistic terms and the appropriation of corresponding concepts. This does not imply, however, the infamous "Private Language Argument" according to which individuals cannot coherently or justifiably introduce certain concepts and terms that express those concepts *solely* on individualistic or first-person perspectival grounds.

¹² One may argue that Mary's knowledge is non-inferential because she does not go through any conscious inferential process in her mind. I am inclined to reply that there are unconscious inferences as well as conscious ones and Mary's inference can be the former if not the latter. Nothing much hangs on this, however. What is important is that there are different stages of epistemic progress in the cases specified and we need an account which explains how one passes from one stage to another.

2.3.2 The Generic Argument: Marianna Wins

Jackson argues that the type of epistemic progress Marianna makes after her experiences is unproblematic for physicalism. He writes:

[The] knowledge Mary lacked which is of particular point for the knowledge argument against physicalism is *knowledge about the experiences of others*, not about her own. When she is let out, she has new experiences, color experiences she has never had before. It is not, therefore, an objection to physicalism that she learns *something* on being let out. Before she was let out, she could not have known facts about her experience of red, for there were no such facts to know. That physicalist and nonphysicalist alike can agree on. After she is let out, things change; and physicalism can happily admit that she learns this; after all, some physical things will change, for instance, her brain states and their functional roles. The trouble for physicalism is that, after Mary sees her first ripe tomato, she will realize how impoverished her conception of the mental life of *others* has been *all along*. (2004, p. 52, italics original)

In this passage, Jackson talks in terms of his Mary case but it is evident that what he is concerned with is essentially what I have pointed out above as the epistemic progress Marianna makes after seeing the red slide. Physicalism can, according to Jackson, account for Mary's new knowledge regarding her *own* experience of red with reference to the novel physical state that takes place in her brain when she first sees ripe tomatoes. The physicalist is supposed to say roughly something like this: surely, Mary learns something new about her own experience because that particular experience is a physical state of which she can have the knowledge only when she is in that state. The problem for physicalism, Jackson claims, rears its head with the recognition that Mary's knowledge is not confined to the knowledge of her own experience but extended to the knowledge of others' experiences of red. What is problematic for physicalism, Jackson claims, is Mary's realization that her complete physical knowledge about red experiences of *others* is not

complete knowledge simpliciter: even given her complete physical knowledge, there is still more to know about *those* experiences.¹³

If Jackson is right, then contrary to the suggestion that I have made in the previous section, the knowledge Marianna acquires cannot be thought of as a premise plausibly deployable in a knowledge argument against physicalism. Marianna does not appear to learn something about others' red experiences given that she does not know that, pointing at the red slide, that is red. She knows what it is like to have *this* experience, where the indexical picks out her particular red experience, but she does not know that that experience is what it is like for others to see red.

However, Jackson is mistaken. The problem for physicalism originates with Marianna's new knowledge rather than Mary's. What a normal perceiver learns when she first has the visual experience of a red slide is not confined to gaining a piece of particular information about that experience-token but is a general information about an experience-type: she learns what it is like to see red, though she would not express what she learns in those words given that she does not know that what she sees is red. I am not sure exactly how to strictly argue for this but it seems all too obvious to me, and hopefully to you too, that Marianna gets acquainted with a certain subjective character which is not merely of a particular experience but which is the common denominator of all red experiences.

Moreover, notice that if one's acquaintance with a certain subjective character were not a problem for physicalism, the whole argumentative force of Jackson's Mary case, whatever it comes to, would evaporate. If Mary's particular red experience were unproblematic for physicalism, then particular visual experiences of each and every individual, whoever he or she

¹³ Perry (1999) endorses Jackson's thesis that the problem with Mary's case for physicalism is the problem of her lacking certain items of knowledge regarding others' visual experiences.

is, would also be unproblematic for physicalism simply because we could have done the very same experiment with any of them instead of Mary. But if physicalism could exhaustively account for particular experiences of each and every individual, then what would remain to be explained? In that case, the complete physical story would account without any residue for each individual's present experiences including Mary's and yours and, in principle, all those particular experiences that occur at a certain time. Whatever remains to be explained even under those circumstances, if anything, it is clear that it is not something physicalism needs to be seriously worried about. Hence, the antiphysicist has no choice but to argue, if he wants to appeal to the Knowledge Argument at all, that the problem for physicalism presented by Jackson's case starts with Mary's new knowledge about her own experience rather than her new knowledge about others' experiences.

A question that needs to be answered at this point of the dialectic is this: what exactly is it that Mary knows in addition to what she and Marianna both know about the experience of red? How substantial is the *extra* epistemic progress Mary makes, that is, does she come to know a new fact about the world Marianna does not know? My contention is that Mary's extra knowledge is a kind of recognitional knowledge and hence does not have antiphysicist implications which are analogous to those her and Marianna's common knowledge about red experiences is supposed to have. Consider the following case. Jordan may know (2) and, at the same time, fail to know (1) even when (1) is uttered by Jordan herself:

(1) I am 56 years old.

(2) Jordan is 56 years old.

In this case, what explains Jordan's knowing (2) but failing to know (1) is her lack of the recognitional knowledge that she herself is Jordan. But this failure of recognition does not imply

that (1) and (2) are distinct ontological facts – neither does it imply that the fact that I (uttered by Jordan) am Jordan is distinct from the fact that Jordan is Jordan. Now compare Jordan’s case with the following:

(3) *Q* is what it is like to have this experience.

(4) *R* is what it is like to see red.

(3) is uttered by Marianna when she is shown the red slide. And, suppose that she also knows (4) in her room: she knows that there is something it is like to see red, and she calls it ‘*R*’. What she does not know is:

(5) *Q* is *R*.

On the other hand, Mary knows (5) in virtue of knowing that this experience (the experience type demonstrated by Marianna) is an experience of red. That is, what she knows but Marianna does not know is the following:

(6) This experience is an experience of red.

However, in accordance with the well-known features of recognitional failures, Marianna’s failing to know (6) does not have ontological implications. She already knows a fact identical to (6) under a different guise (for instance, she knows that an experience of red is an experience of red) just as Jordan knows a fact, which is (2), identical to (1) under a different guise.

There have been various attempts in the literature (Loar 1990/1997, Perry 1999) to think of Mary’s entire new knowledge as merely a kind of recognitional knowledge. I will later argue that these attempts are mistaken. For now, it should suffice to say that they rest essentially on a failure to distinguish different pieces of knowledge Mary gains after she sees ripe tomatoes: the one that she shares with Marianna and the other one she has but Marianna lacks. What I have

argued above is that the latter is to be viewed as recognitional; however, the former is not recognitional.

If Mary's extra knowledge (i.e., the knowledge that, demonstrating an experience of red, that is a red experience) is recognitional, then since lacking recognitional knowledge cannot be treated as lacking knowledge about worldly facts, the generic Knowledge Argument against physicalism rests on the knowledge Mary and Marianna both gain after their experiences of red (i.e., the knowledge that it is like *Q* to see red). Neither Mary's learning that her experience is of red nor her learning that others' red experiences are like *this* (demonstrating the experience she has when she sees ripe tomatoes) are problematic for physicalism. This means that Mary's case involves certain complications which are irrelevant from the point of view of the question as to the truth of physicalism, and it thus provides a fertile ground for otherwise easily avoidable confusions. For this reason, I submit, Marianna's case would be more effective to bring the main point of the Knowledge Argument into focus.

2.4 THE KNOWLEDGE ARGUMENT: NAGEL'S BATS

There is much to be appraised in Nagel's wonderfully rich "What Is It Like to Be a Bat?" (1974). Among other things, it goes beyond Jackson's Mary in having a clear account of *why* physical theories of conscious experiences appear to be destined to be incomplete, containing an important insight which will be central to our ensuing discussion. However, like Jackson's Mary case, Nagel's case from bat experiences is presented in a way that feeds certain misunderstandings about the central moral of the Knowledge Argument, and separating what is

relevant from what is irrelevant here also requires substantial work. In what follows, I will develop a rational reconstruction of Nagel's account of phenomenal consciousness which I believe is loyal to the spirit if not the letter of his paper.

Nagel's argument from bat experiences, in one of its simple and intuitive forms, can be formulated as follows:

(1) There is something it is like to have bat experiences.¹⁴

(2) What it is like to have bat experiences cannot be known by us.

From (1) and (2),

(3) There is something about bat experiences ("what it is like to have them") which cannot be known by us.

We independently have the following two:

(4) Everything about bat neurophysiology can be known by us.

(5) Physicalism implies that complete knowledge about bat neurophysiology gives us complete knowledge about bat experiences.

From (3) and (4),

(6) There is something about bat experiences beyond what we can know about bat neurophysiology.

From (5) and (6),

(7) Physicalism is false.¹⁵

¹⁴ Nagel uses the phrase "what it is like to be a bat" rather than "what it is like to have bat experiences." However, this usage gives rise to certain problems which are irrelevant to the truth of physicalism (see Jackson 1982). Moreover, it has an air of mysticism because it appears to conflate the issue about bat experiences with the issue of whether there is a bat-self just like, as some believe, we have human-selves. It is clear, however, from Nagel's writing that it is the former issue that he is concerned with. So, for present purposes, we can take what it is like to be a bat as analyzable in terms of what it is like to have all (and only) bat experiences.

¹⁵ In his paper, Nagel formulates his argument more cautiously, using the locution "appears (that)" instead of "is (that)" in the premises where suitable. Hence, the final conclusion he derives from his own reasoning is not that

Nagel's argument from bat experiences thus formulated is very similar to Jackson's argument from the super-scientist Mary except for a certain shift in perspective: while, in Jackson's argument, we are asked to consider someone who appears to be incapable of knowing what *our* experiences are like, Nagel invites us to appreciate that we ourselves appear to have no cognitive access to what fundamentally *alien* experiences are like.

The reason why Nagel thinks that premise (2) in the argument above is correct requires some attention. Premise (2) is, Nagel thinks, implied by the following two theses: (a) "at present," Nagel writes, "we are completely unequipped to think about the subjective character of experience without relying on the imagination – without taking up the point of view of the experiential subject" (1974, p. 449), and (b) "nothing in [our] present constitution enables [us] to imagine" (1974, p. 439) what the subjective characters of fundamentally alien experiences (e.g., bat experiences) are like.¹⁶ Nagel argues that if the only method we could deploy in order to know the subjective characters of others' experiences – the method of imagination – does not help in the case of fundamentally alien experiences, then what those experiences are like cannot be known by us.

Nagel's appeal to humans' imaginative abilities strikes one as bizarre in the first instance. For one thing, imagination appears to be a notoriously unreliable faculty for knowledge. Knowledge implies truth and since we can imagine something that is not the case we can imagine something we cannot know. Even when we imagine something that is the case, we cannot know whether it is the case merely in virtue of that imagination. For another thing,

physicalism is false but that we do not have any conception of how physicalism might be true and hence do not really understand what physicalism means. I will have more things to say about this conclusion later but for present purposes, it is convenient to take his argument as if it attempts to establish the falsity of physicalism.

¹⁶ Nagel uses the phrases "what it is like to have an experience *E*" and "the subjective character of an experience *E*" interchangeably.

physicalism does not appear to be committed to any special claims about the imaginative powers of human beings. It is unclear how one can argue from the deficiency of our imaginative abilities to the falsity of a metaphysical thesis which is silent on those abilities.

An instructive way to begin to make sense of Nagel's talk of imaginative abilities is to ask the question how a person knows what human pain is like. The problem Nagel wants to raise is not simply the problem of other human minds or the privacy of experience to its possessor. For Nagel, we *can* know what other human minds are like. "One person can know or say," Nagel writes, "of another what the quality of the other's experience is" (1974, p. 442). His question is: *how* do we know that quality? His answer is that we know what the experiences of other humans are like in virtue of extrapolating to their inner lives from our own case, and this does not involve any special difficulties parallel to the ones about knowing the experiences of bats because, being humans, we already occupy and thus adopt the relevant subjective point of view. And, given that adopting the relevant subjective viewpoint *is* (or at least requires) imagination (1974, p. 449), the epistemic connection between knowing subjective characters and imaginative abilities turns out to be less controversial than it may initially sound.

However, the worry as to what exactly our less-than-ideal imaginative abilities have to do with the truth of a metaphysical thesis like physicalism still remains untouched. Nagel's appeal to imagination gives the impression that if we were able to extrapolate to the inner lives of bats from our own case, then there would be no serious threats to physicalism. However, this is false. The problem is that even if we knew the subjective qualities of bat experiences, the source of our knowledge would still be problematic for physicalism because we would not be knowing those qualities on the basis of what we know when we have all the relevant physical knowledge. The point can be driven closer to home: the subjective characters of our own experiences of which we

undoubtedly have the knowledge are also in the category of what seem to escape the physicalist story. Hence, the question is not simply what we can know given the capacities of the cognitive and imaginative faculties we actually have but what we can know on the basis of what physical sciences tell us.

Even though Nagel's allusion to imaginative or extrapolative abilities obscures the intent and force of his version of the Knowledge Argument, his answer to the question "why are items of knowledge about the subjective characters of experiences inaccessible *from the physical point of view?*" involves a fundamental insight. A characterization of Nagel's insight demands an articulation of how physical explanations typically proceed for the things in the objective world. The standard model of explanation we find in physical sciences is of a reductive form, which rests on a central distinction between appearance and reality. Seeking a more comprehensive understanding of the external world, physical sciences move in the direction of greater objectivity by omitting the individual or species-specific points of view. For instance, in discovering heat to be, in reality, a form of molecular motion, the physical point of view leaves behind the phenomenal point of view we have on heat: heat sensations. Leaving the phenomenal viewpoint aside is legitimate in this familiar case of scientific reduction because even though the concepts we employ in thinking about the external world are inescapably initially formed from a particular point of view that is shaped by our experiences, they are used by us to refer to things other than those experiences. The physical phenomenon of heat may induce completely different types of experiences in some non-human creatures, and those creatures may never be in a position to have a complete grasp of the specifically human concept of heat. But they may still know what heat is because it has an objective nature which can be separated from *our* heat sensations.

What happens when we direct our reductionist weapons on experiences themselves? Are they equipped with the appropriate type of ammunition? It is clear that in the case of experiences, the relocation trick is no longer open to us. Heat sensations are experiences we have of heat by means of which we fix the reference of the concept *heat*. By moving from heat sensations to the referent of the concept *heat*, one leaves behind what does the reference-fixing and attends to the thing out there in the external world. However, experiences cannot be ignored permanently, and a disturbing fact for physicalism is, Nagel maintains, that there is nowhere to turn to if the problematic properties are already placed within the phenomenal mind. The familiar reductionist shift does not work here and phenomenal consciousness thus appears to be unique in its being resistant to the sort of reductions found elsewhere in science.

This reasoning from omitting experiences in well-known cases of reductions to the irreducibility of experiences is a *non sequitur* and it is important to see why. There are causes and there are effects, and the exclusion of effects in an attempt to understand causes is a methodological triviality. There is nothing mysterious in itself with Nagel's observation: certainly, if your aim is to provide a physical account of the nature of what causes our heat sensations, you are well advised to leave heat sensations aside. And, surely, when it comes to understanding the nature of heat sensations themselves, you are equally well advised not to take them out of the picture; but this does not by itself imply that experiences defy the familiar reductive strategy of explaining higher level phenomena in terms of lower level phenomena: what it tells us is that physicalists searching for reductive explanations for experiences cannot pass the buck and have to approach them head on.

The objection above against the Nagel-type reasoning makes a plausible point: from the fact that experiences are left out in paradigmatic cases of scientific reduction, it does not follow

that they cannot be reduced. And, this takes us to the heart of the matter: what reason do we have to believe that experiences are immune from scientific reductions once they are taken as the *target* of such reductions? The answer Nagel provides to *this* question is of central importance: while we can easily make sense of the idea of there being distinct concepts referring to the same physical thing in the external world, we do not really understand how there can be distinct concepts referring to experiences as they are experienced by the subject. As pointed out before, heat as a form of molecular motion can cause different types of experiences in different kinds of sensuous beings depending on their perceptual apparatuses, and those beings can develop distinct concepts picking out the same physical phenomenon of heat which are grounded in those experiences. However, can there really be different concepts of one and the same type of experience? Experiences do not appear to be things on which we can have different points of view analogous to the ones we can have on things in the external world. In contrast with external things, “in the case of experience,” Nagel writes, “the connection with a particular point of view seems much closer” (1974, p. 443). For instance, the idea of there being a non-human sensuous being that has a certain experiential or phenomenal perspective on what it is like to have human pain which is distinct from the one built in the human concept *pain* seems incomprehensible. As long as this being has any phenomenal perspective on human pains at all, it appears that it must be the very same perspective incorporated in the human concept *pain*. What we can conveniently call *Nagel’s Insight* (NI) can be stated as follows:

(NI) A phenomenal concept of a certain experience is essential to understand the nature of that experience.

As Nagel prefers to put it, phenomenal concepts give “the essence of the internal world” (1974, p. 445). Nagel’s idea is that an understanding of experiences without the deployment of the relevant phenomenal concepts is bound to be irredeemably incomplete.

NI enables us to give a new twist to the old appearance-reality gap. We can conceive the gap in question not in terms of a relation between how things look to us and how they really are but in terms of a relation between how our concepts represent things and how they really are. From this perspective, there is an appearance-reality gap, for instance, between the physical phenomenon of heat and the human concept *heat* in the sense that one can understand the objective nature of the former without having the ability to grasp the latter, and hence the human concept *heat* can be justifiably omitted in the search for an account of heat itself. By contrast, however, there is no corresponding gap, say, between the pain experience and the phenomenal concept *pain*. The path to understanding the nature of the pain experience is cognitively closed to those who are incapable to grasp the phenomenal concept *pain*.

I will later discuss NI and the uniqueness of phenomenal concepts in more detail. For now, let me just note that if NI is true, it appears that physicalism is in trouble. The thesis of the identity of phenomenal states to physical states seems false or at least deeply problematic given that one can fully comprehend the nature of physical states without having the capacity to grasp the concepts which are essential to understanding the essence of phenomenal states. NI gives us the beginnings of an answer to the question why it appears that complete physical knowledge is not complete knowledge *simpliciter*.

3.0 DOES MARY GAIN ANY NOVEL KNOWLEDGE?

When Mary leaves her room and experiences red, will she be surprised and exclaim “At last! So, *that* is what it is like to experience red”? Will she gain a new concept of (and, accordingly, entertain a new thought about) what it is like to experience red? According to one form of physicalism which will be addressed in this chapter, the answer is “no”: given her complete repertoire of physical concepts, she has all the conceptual resources that are required to entertain all the thoughts which can be possibly entertained about experiencing red. This form of physicalism holds that if one has all physical conceptual knowledge (knowledge which is expressed by physical concepts) then one has all conceptual knowledge.

The question that naturally arises is how this form of physicalism accounts for the strong intuition that upon seeing red, Mary will be learn something new that can be expressed by the sentence “so that is what it is like to experience red!”, where the demonstrative *that* expresses a novel concept of experiencing red. There are basically two sorts of reply to this question in the literature, and now I will investigate them and show that neither is plausible.

3.1 DENNETT AND “FIGURING OUT” WHAT IT IS LIKE

Dennett raises several objections to the general understanding of Mary’s case as it is presented by Jackson. Dennett thinks that the thought experiment “actually encourages us to misunderstand its premises” (1991, p. 398). The premise Dennett especially has in mind is the premise that Mary has all physical knowledge about experiencing red. This premise is, for Dennett, problematic not because it is impossible for a human being to have all physical knowledge but because imagining that one has all physical knowledge is “so preposterously immense” (1991, p. 399) for us that when we think we imagine Mary’s epistemic situation, what we really imagine is typically something much weaker. He writes:

What [the thought experiment] ask[s] you to imagine is so preposterously immense, you cannot even try. The crucial premise is that “She has *all* the physical information.” That is not readily imaginable, so no one bothers. They just imagine that she knows lots and lots – perhaps they imagine that she knows everything that anyone knows *today* about the neurophysiology of color vision. But that’s just a drop in the bucket, and it’s not surprising that Mary would learn something if *that* were all she knew. (1991, p. 399)

For Dennett, we just do not imagine what the thought experiment asks us to imagine but we take ourselves to imagine this. This “illusion of imagination”, according to Dennett, explains why we are so strongly inclined to think that Mary will learn something new when she first experiences red: since what we really imagine is merely that Mary knows a lot about the physics of experiencing red, we feel that what she learns in her room can be increased by experiencing red.

Dennett is certainly right that there is a sense in which it is difficult to imagine a human being who has all the physical information about the sensory perception of red. The difficulty stems from two facts: first, we do not have all the relevant physical information; second, all the

relevant physical information is presumably too much for us to digest in order to imagine the case at hand. Dennett makes use of both facts to support his contention that we typically fail to imagine what the thought experiment asks us to imagine. But I think the former fact, that is, the fact that our current knowledge is incomplete, is by itself sufficient to justify his contention: if we do not actually have all the physical information, how can we imagine someone who has that information? It is plausible that imagining such a person requires knowing what she knows. Thus, Dennett's observation points to a claim which is stronger than the one he makes: it is not only very difficult but in fact impossible for us to imagine a person who has all the physical information.

However, I think that Dennett's observation does not show what he intends to show. The difficulty (or even impossibility) of imagining a person who has all the physical information does not explain why we are so strongly inclined to think that Mary will acquire a new piece of information through seeing red. Our current knowledge about the neurophysiology of visual perception may be incomplete but we still have a good sense of the *sort* of account a neurophysiological inquiry into visual perception may provide. And, this is what grounds our inclination to think that Mary will learn something new through seeing red: a neurophysiological account of visual perception, whether it is complete or not, is just not of the *right* sort to enable one to know what an experience of red is like. It may be difficult or even impossible for us, in a certain sense, to imagine a person who has all the physical information; but if such a person were readily imaginable for us, our inclination to think that Mary will learn something new would still persist because Mary's neurophysiological knowledge, complete or not, is just not the right sort of knowledge. Why do we think that a blind neurophysiologist cannot know, on the basis of his scientific expertise, what seeing red is like? Is this because we think that his scientific knowledge

is incomplete? No. It is because his scientific knowledge is not of the right sort to enable him to know what seeing red is like.

There are two different senses of imagining a person who has all the physical information. According to the first one, imagining a person who has all the physical information is imagining a person who knows a very long conjunctive statement whose conjuncts are general physical truths like “such-and-such neurons fire in such-and-such frequencies when an object with such-and-such surface reflectance properties is perceived under such-and-such conditions”. This is the sense, according to Dennett, in which a person who has all the physical information is “not readily imaginable” (Dennett) for us. According to the second sense, however, imagining a person who has all the physical information is imagining a person whose knowledge of a certain sort is complete. I think that in this second sense, a person who has all the physical knowledge is imaginable for us; in fact, this is what we typically do when we consider Mary’s situation. In this sense, there is no “illusion of imagination”. Dennett’s error is to take it that what the thought experiment asks us is to imagine a person who has all the physical information in the first sense. The thought experiment would not even get off the ground if that were really the sense in question. But the thought experiment does not need the first sense of imagination to make the point it is designed to make. Imagining a person who has all the physical information in the second sense is enough to put one in a position to recognize that complete physical knowledge is not complete knowledge *simpliciter*.

Dennett argues that nobody has *showed* why Mary’s story must proceed as it is presented by Jackson and he points out that no reasons have been given for the claim that she will learn something new through experiencing red. Dennett puts forwards another scenario (“the blue

banana trick”) which is intended to demonstrate that Mary’s story could have proceeded in a different but equally plausible way:

And so, one day, Mary’s captors decided it was time for her to see colors. As a trick, they prepared a bright blue banana to present as her first color experience ever. Mary took one look at it and said “Hey! You tried to trick me! Bananas are yellow, but this one is blue!” Her captors were dumfounded. How did she do it? “Simple,” she replied. “You have to remember that I know everything – absolutely everything– that could ever be known about the physical causes and effects of color vision. So of course before you brought the banana in, I had already written down, in exquisite detail, exactly what physical impression a yellow object or a blue object (or a green object, etc.) would make on my nervous system. So I already knew exactly what thoughts I would have (because, after all, the mere disposition to think about this or that is not one of your famous qualia, is it?). I was not in the slightest surprised by my experience of blue (what surprised me was that you would try such a second-rate trick on me). I realize it is hard for you to imagine that I could know so much about my reactive dispositions that the way blue affected me came as no surprise. Of course it’s hard for you to imagine. It’s hard for anyone to imagine the consequences of someone knowing absolutely everything physical about anything! (1991, pp. 399)

In one of his later papers on the Knowledge Argument, Dennett explains the motivation for his alternative “blue banana” version:

My variant was intended to bring out the fact that, absent any persuasive argument that this could not be how Mary would respond, my telling of the tale had the same status as Jackson’s: two little fantasies pulling in opposite directions, neither with any demonstrated authority. (2007, p. 16)

Here Dennett does not claim that Mary’s story must proceed as the blue banana version depicts it, but he claims that the story can proceed that way *given what the main premise of the thought experiment* (that Mary has all the physical knowledge) *requires*. Dennett’s idea is that Mary’s

having complete physical knowledge is compatible with both versions and Jackson gives no reason to prefer his “little fantasy” over Dennett’s.

A question is: can Mary’s story proceed as Dennett’s variant describes? Notice that on Dennett’s variant, Mary is in a position to know, on the basis of her perfect physical knowledge, what it is like to have certain types of visual experiences (experience as of red, experience as of blue, etc.) without undergoing those experiences. On his variant, Mary “was not *in the slightest* surprised” (my italics) by her experience of blue because she already knew in advance what it is like to have that experience. Dennett is clear on this point:

I have found some readers – maybe most – just didn’t get my blue banana alternative. What was I saying? I was saying that Mary had figured out, using her vast knowledge of color science, *exactly what it would be like for her to see* something red, something yellow, something blue in advance of having those experiences. (2007, p. 16)

But how can Mary “figure out” what it is like for her to see red without having an experience as of red? Dennett tells us that just by looking at the banana, Mary can recognize whether it is yellow or blue because she already knows “in exquisite detail, exactly what effect a yellow object or a blue object ... would have” on her nervous system. But knowing which effects, described in neurophysiological terms, each particular color has on one’s nervous system does not provide one with the knowledge of whether an object one is *currently* perceiving is having *that* effect.¹⁷ Mary knows that perceiving blue under certain lighting conditions, described in physical terms (e.g., in terms of light waves), causes such-and-such neurophysiological effects. But when she first sees the blue banana, she does not know, immediately and directly, whether her perception is currently causing those effects. If she already knew that she is perceiving blue

¹⁷ This point is nicely put in H. Robinson (1993).

when she sees the banana, then on the basis of her perfect physical knowledge, she could figure out what kinds of neurophysiological states are taking place in her brain. Or, if she knew what kinds of neurophysiological states are currently taking place when she sees the banana, she could figure out that she is perceiving blue. But since no reason has been given by Dennett for the claim that she can have one of these pieces of knowledge about her *current* perceptual/neurophysiological states when she first sees the banana, no reason has been given by him for the claim that she can have the ability to tell, just by looking at the banana, whether it is blue or not: she surely knows that she is having a visual experience of a certain type, but she cannot tell whether it is an experience as of blue through her knowledge about her current perceptual/neurophysiological states because she simply does not have that particular knowledge.

But let us suppose, for the sake of the argument, that she has all the knowledge, in physical terms, about her current states when she is seeing the banana. In that case, she would be able to figure out that she is having an experience as of blue on the basis of that knowledge combined with her general scientific knowledge acquired in her room. It is a simple deduction: objects which others normally call “blue” provide such-and-such physical inputs of light waves on one’s perceptual system, this particular banana provides the very same input on my perceptual system, and hence this particular banana is what others normally call “blue”. However, this would not show that Mary is not “in the slightest surprised” when she experiences blue. This is because the ability to tell whether one is having an experience of blue or not *through* her physical knowledge is compatible with one’s being surprised and one’s exclaiming “so, *this* is what it is like to have an experience of blue”. One can have that ability without knowing what it is like to have an experience of blue. In order for Mary not to be surprised, she must figure out

what it is like to experience blue without having an experience of blue, and this she cannot do merely by having the ability to tell whether something is blue (or called 'blue') on the basis of her physical knowledge.

So, again, how can Mary know what it is like to see blue (or red) through her perfect physical knowledge? On Dennett's variant, Mary may know what it is like to see blue on the basis of her knowledge about what *thoughts* other people have (or she would have) when they see (or she sees) blue. Dennett assumes that Mary can have complete knowledge of thoughts about experiences through her complete physical knowledge because he assumes that thoughts are "mere dispositions" (presumably toward *inter alia* verbal responses) and dispositions are physically unproblematic. Now, if Mary can have complete knowledge of thoughts about experiences (e.g., the thought *that that is what it is like to see blue*) through her physical knowledge, then she can know what it is like to see blue. Thanks to her knowledge about thoughts, whose entire nature can be described by physical terms, she understands the thought *that that is what it is like to see blue*, and understanding that thought is not possible without knowing what it is like to see blue.

However, this argument works only if we assume that thoughts about experiences are "mere dispositions," and Dennett is mistaken in supposing that the friends of the Knowledge Argument will find that assumption unproblematic. In fact, the very same considerations which make experiences physicalistically problematic according to the friends of the Argument hold *mutatis mutandis* for thoughts about experiences. In her colorless captivity, Mary may know about the "reactive dispositions" (Dennett) standard perceivers have when they see blue (e.g., what they are disposed to say, how they are disposed to act), but she will not be able to entertain the thought *that that is what it is like to have an experience of blue*, where the concept expressed

by the demonstrative *that* is a concept of experiencing blue which she cannot possess without experiencing blue. The intuition is that in addition to her ignorance about the nature of experiences, Mary will not be able to entertain certain thoughts about experiences. The conditional is surely true: if one can entertain the thought *that that is what it is like to have an experience of blue*, then one can know what it is like to have an experience of blue. But, the antecedent of the conditional cannot be assumed in an argument against the Knowledge Argument, it is what is supposed to be proved. Dennett begs the question against the Knowledge Argument by assuming that *all* thoughts about experiences are mere dispositions, things whose nature is entirely amenable to a physicalist description.

Mary's story can proceed in the way Dennett portrays it only if Mary can figure out what it is like to see blue through her perfect physical knowledge. However, Dennett's proposal for how Mary can figure it out assumes what needs to be proved, and hence is not acceptable. Still, Dennett is right when he writes: "the usual [i.e., Jackson's] way of imagining [Mary's] story doesn't *prove* that she *does* [learn anything]" (1991, p. 400). The Knowledge Argument does not even try to prove, let alone prove, that Mary learns anything upon seeing red. It simply takes Mary's novel knowledge for granted, and argues from that premise (and some other premises) to the falsity of physicalism. But if the Knowledge Argument does not prove that in her captivity, Mary can not figure out what it is like to see red, what right do we have to request a proof from Dennett that she indeed can? It may look as if because neither side presents a proof to back up its own thesis, what we have is a stalemate: without a proof at hand to the effect that Mary does or does not learn anything, one may understandably say, we can simply forget about the Knowledge Argument.

However, the fact that Jackson does not present any formal proof for the claim that Mary learns something does not show that, contra Dennett, the claim is made merely “on the basis of ancient philosophical tradition going back at least to John Locke” (2007, p. 16) and is itself “a bit of folk psychology with nothing but tradition ... in its favor” (2007, p. 21). Dennett wants to give the impression that Mary’s epistemic progress is a “hypothesis” of one of the possible “theories” about human mind (“folk psychology”) supported by an “ancient tradition” which is suspicious by its “philosophical” character. But the impression is wrong. From my own experience, I can confidently say that even the most uninitiated with no information about the history of philosophy find the idea that Mary will learn something new quite plausible, and when asked how they support it, they are generally inclined to say something along the lines “it just seems obvious” or “how else could it be?” What grounds our intuitions about Mary’s epistemic progress is not “ancient philosophical tradition” but the fact that we do not even have a *clue* about *how* Mary might have deduced what it is like to see red without seeing red. We do not know *where* to start the deduction (e.g., which physical facts are the main premises which play a role in deducing what it is like to see red?), let alone know what *structure* the alleged deduction will take. This is because knowledge of experiences which is had on the basis of having those experiences (“phenomenal knowledge”) appears to be radically different from knowledge of experiences expressed in physical terms: they are just, or strongly appear to be, different *sorts* of knowledge. If someone like Dennett is tempted to say that there is somewhere to start and there is a foreseeable and promising way of making the deduction, then he better say what they are. If he cannot do any better than claiming that no formal proof can be given to the effect that Mary learns something upon seeing red, then I think the friends of the Knowledge Argument are justified in their presumption of Mary’s epistemic progress. After all, formal proofs cannot, and

indeed need not, be given for many things we are justified in believing; and, in our case, the appearance of the radical difference of physical knowledge from phenomenal knowledge is a sufficiently good reason to be justified in believing that Mary learns something.

Is it possible that when we say that physical and phenomenal knowledge are of different sorts, that is, when we say that no amount of physical knowledge can capture the content of a piece of phenomenal knowledge, we make the mistake of inflating, as Dennett says, “practical indescribability into something metaphysically more portentous” (2007, p. 19)? Dennett writes:

[If] what it is like to see triangles can be adequately conveyed in a few dozen words, and what it is like to see Paris by moonlight in May can be adequately conveyed in a few thousand words (an empirical estimate based on the variable success of actual attempts by novelists), are we really so sure that what it is like to see red or blue can’t be conveyed to one who has never seen color in a few million or billion words? (2007, p. 21)

Can it really be the case that while what it is like to see red is in principle physically describable, we think it is not because we are mistaking “practical indescribability” for in-principle indescribability? Consider Dennett’s example of the torn Jell-O box (Dennett 1998), half of which has shape property M, and the other half of which is the only *practical* M-detector: the shape *practically* defies description, but as Dennett writes (2007, p. 19), “it is not literally ineffable or unanalyzable; it is just extremely rich in information”. Perhaps, Dennett suggests, the indescribability of what it is like to see red is similar to that of the shape M; and if his suggestion is right, the presumption that physical and phenomenal knowledge are of different sorts is false.

However, the claim that what it is like to see red is merely practically, but not in principle, physically indescribable is plausible only if we have *some* idea, however vague and underdeveloped, of how we might *start* to describe what it is like to see red in physical terms.

But no physical term we possess affords us any clue as to how it might be a constituent of a description of what it is like to see red. This is in contrast with the case of describing the shape of the torn Jell-O box, which is at bottom a matter of articulating a long conjunction of descriptions of instantiations of different shape properties. Describing the shape of the box is admittedly *very tough* but it does not appear to be something which is in principle impossible. However, describing what it is like to see red in physical terms appears to be *impossible*; and, this appearance is grounded in the recognition that physical terms are simply non-starters for the job at hand (if only we knew how to start a physical description of what it is like to see red). Until then the friends of the Knowledge Argument are justified in holding that Mary cannot deduce what it is like to see red on the basis of her perfect physical knowledge.

To summarize: (1) Dennett's claim that there is a sense in which the thought experiment is "not readily imaginable" for us is plausible; however, Dennett is mistaken in thinking that that is the sense which is required by Mary's case and in claiming that our failure to imagine the case in that sense explains our inclination to think that Mary will learn something. (2) Dennett's idea that one can gain the ability to recognize what it is like to see colors on the sole basis of general scientific knowledge is implausible. Moreover, even if that were plausible, that ability can still be had without knowing, in a significant sense of the term, what it is like to see colors. (3) Dennett argues that Mary has complete knowledge of thoughts about experiences because thoughts are "mere dispositions," which are entirely amenable to physical descriptions; and, he claims that her complete knowledge of thoughts about experiences enables her to know what it is like to have those experiences. But by assuming that *all* thoughts about experiences are physically describable, Dennett begs the question against the Knowledge Argument. (4) Dennett's suggestion that what it is like to have experiences is merely "practically indescribable"

in physical terms would be plausible only if we had a clue about how to start the description (notice that the reason why we think the shape of the torn Jell-O box is *in principle* physically describable is that we have a sense, however rudimentary it may be, of how to start the description and of how it may go). But, as long as no physical terms we have appear to be even slightly promising for the job, we are justified in thinking that what we have is in principle indescribability.

3.2 LEWIS AND THE ABILITY HYPOTHESIS

The Ability Hypothesis, as it is articulated by David Lewis, is about what it is to learn what an experience is like. It is developed as an idea which rests upon two further ideas which appear to be in tension with each other. Firstly, contra Dennett, Lewis takes it as a *datum* to be accounted for that no amount of physical information can just by itself enable one to know what it is like to have certain experiences. Lewis writes:

No amount of the physical information that black-and-white Mary gathers could help her know what it was like to see colors; no amount of the physical information we might gather about bats could help us know what it's like to have their experiences; and likewise in other cases. (2004, p. 84)

Secondly, Lewis is in agreement with Dennett that there is no distinctive sort of information gained through having experiences (“phenomenal information”), and claims that all information

about experience is physical information and can be in principle conveyed in lessons. Lewis writes:

There is a natural and tempting explanation of why physical information does not help. That is the hypothesis that besides physical information there is an irreducibly different kind of information to be had: *phenomenal information*. The two are independent. (2004, p. 84)

Lewis tells us that the Hypothesis of Phenomenal Information is invoked by the Knowledge Argument to explain why no amount of physical information suffices to teach us what a new experience is like and he argues that it needs to be rejected by materialists.

The knowledge argument works. There is no way to grant the hypothesis of phenomenal information and still uphold materialism. Therefore I deny the hypothesis. (2004, p. 90)

How does Lewis account for the alleged datum that no amount of physical information could help Mary know what it is like to see colors if phenomenal information is, as he claims, “an illusion” (2004, p. 91)? In other words, if physical information is the only sort of information one can have about experiences, then since, by hypothesis, Mary has all the physical information, what is it that Mary gains through having the experience as of red? The Ability Hypothesis is presented as an alternative to the Hypothesis of Phenomenal Information, which answers to these questions. It is characterized by Lewis as follows:

The ability hypothesis says that knowing what an experience is like just *is* the possession of these abilities to remember, imagine, and recognize. It isn’t the possession of any kind of information, ordinary or peculiar. It isn’t knowing that certain possibilities aren’t actualized. It isn’t knowing-that. It’s knowing-how. Therefore it should be no surprise that lessons won’t teach you what an experience is like. Lessons impart information; ability is something else. Knowledge-that does not automatically provide know-how. (2004, p. 100)

Lewis's idea is, simply put, that abilities satisfy the two demands specified. First, abilities cannot be learned simply by scientific lessons. Ability is knowing-how; and as such it cannot be gained by any amount of physical information or knowing-that. Second, ability is not a distinctive sort of information because it is not a sort of information or knowing-that at all. Lewis's point is that while physical information is the only sort of information one can have, having information is not the only sort of knowing. The "illusion" of phenomenal *information*, for Lewis, derives from a failure to distinguish two sorts of knowing:

[W]e confuse ability with information because we confuse knowledge in the sense of knowing-how with knowledge in the sense of knowing-that. There may be two senses of the word 'know', but they are well and truly entangled. They mark the two pure endpoints of a range of mixed cases. The usual thing is that we gain information and ability together. If so, it should be no surprise if we apply to pure cases of gaining ability, or to pure cases of gaining information, the same word 'know' that we apply to all the mixed cases. (2004, pp. 100-1)

For Lewis, there is one sort of knowing-that (and it is physical) while there are (at least) two sorts of knowing (and they are knowing-that and knowing-how). The confusion from which the Knowledge Argument gains its intuitive appeal stems from being unclear about the distinction between sorts of knowing-that and sorts of knowing.

Let us assume for the sake of the argument that there are two irreducibly different sorts of knowing, knowing-that and knowing-how. That is, let's suppose that having no amount of knowing-that implies having a piece of knowing-how and vice versa. Of course, we need a qualification here. As Lewis is well aware, some abilities are entirely information-dependent. "If you want to know how to open the combination lock on the bank vault, information is all you need" (Lewis 2004, p. 100). The point of the Ability Hypothesis is that the abilities gained through having a certain type of experience (abilities like remembering, recognizing, and

imagining the experience) do not consist simply of possession of information. So, the Ability Hypothesis claims that there is a sort of knowing-how that cannot be acquired by any piece of knowing-that, and it argues that remembering, recognizing, and imagining an experience are of *that* sort of knowing-how. Let us assume that this is true. The question I will address is whether knowing what it is like to have experiences *consists in* the abilities Lewis mentions. My answer to this question is “no”, though not for the reasons proposed by some other opponents of the Ability Hypothesis. I will argue that the ability to entertain new thoughts, which is not on Lewis’s list, is typically acquired through experiencing, and this shows the falsity of the idea that what Mary acquires are merely pieces of knowledge-how. Assuming that some of those new thoughts are cases of knowing (Lewis does not deny that through experiencing, Mary comes to *know* something new), some of what Mary acquires are pieces of knowledge-*that*.

A famous argument against the Ability Hypothesis is provided by Conee (1994). Conee argues that the abilities to remember, imagine, and recognize a certain experience are not necessary for knowing what it is like to have that experience. He writes:

Suppose that Mary has no visual imagination. She is unable to visualize anything. Now the story continues as before... [W]hile she is intently gazing at the color of red ripe tomatoes, it is clearly true that she knows what it is like to see something red. She has made an exciting discovery. ‘Aha!’, she might well exclaim. Yet she is unable to imagine anything. *A fortiori*, she is not able to imagine, remember, and recognize the experience... In light of her incapacity to imagine, it is also true that she does not know how to visualize red at will. Hence, knowing what an experience is like does not imply having any such abilities. (2004, pp. 200-1)

Conee argues that Mary knows what it is like to see red *at the time she is staring at the ripe tomatoes* but she does not, by hypothesis, have the abilities to remember, recognize, and imagine

experiences as of red. So, Conee claims, knowing what it is like to see red cannot be identified with having those abilities.

Conee's argument works only if we can plausibly assume that just by "intently gazing at" the color of red ripe tomatoes (without having the relevant abilities), Mary can *know* what it is like to have experiences of red. It seems at least *prima facie* plausible that Mary can have experiences of red just by virtue of gazing at the color of red tomatoes, but the supposition that having experiences itself suffices for knowing what those experiences are like is very controversial. Certainly newborn babies can have some experiences, but it is not that certain whether they *know* what those experiences are like. The friends of the Ability Hypothesis can be plausibly taken as arguing that the gap between having an experience and knowing what it is like is bridged by the possession of certain abilities like imagination and recognition. Conee assumes what the friends of the Ability Hypothesis deny ("gazing at X suffices for knowing X"). He does not, strictly speaking, beg the question against the Ability Hypothesis just as the Ability Hypothesis does not beg the question against Conee; but because of its reliance on an assumption many will find problematic, Conee's argument against the Ability Hypothesis is dialectically weak.

Let us assume that abilities like imagination and recognition are acquired through experiencing, and they are necessary for knowing what it is like. An important question is whether these abilities are information-dependent, that is, whether their acquisition requires the antecedent possession of certain pieces of knowing-that. I will also concede to the Ability Hypothesis, just for the sake of the argument, that these abilities are not information-dependent: they are pure cases of knowing-how. Now, let us ask whether the Lewis abilities are the only abilities one acquires through experiencing. Consider Mary again as she remarks, "So, this is

what it is like to experience red.” Intuitively, in making this remark, Mary is expressing a thought that she could not entertain before. Before having the experience, she was wondering what it is like to experience red. Thanks to her perfect physical knowledge, she already knew that what it is like to see red resembles what it is like to see orange more than it resembles what it is like to see blue. She knows which experiences are called “similar” by subjects who have them; which ones involve roughly the same neurons firing in similar frequencies, and so on. As Lewis nicely put it:

Maybe Mary knows enough to triangulate each color experience exactly in a network of resemblances, or in many networks of resemblance in different respects, while never knowing what any node of any network is like. Maybe we could do the same for bat experiences. But no amount of information about resemblances, just by itself, does anything to help us know what an experience is like. (2004, p. 80)

Through experiencing, Mary comes to entertain a novel thought about what it is like to see red, a thought which goes beyond the thoughts she could have entertained in her room about its similarities to other experiences. The thought that this is what it is like to experience red, where *this* denotes the type of experience she is having when she is seeing the tomatoes, is a novel thought. It is clearly a piece of thinking-*that*, a truth-evaluable thought.

Now, as Lewis himself agrees, Mary comes to learn something new when she first experiences red. And learning is coming to know something one did not know before. (As Stoljar and Nagasawa puts it, “it is simply part of the notion of learning that if S learns p at t, then there is a previous time t^* such that S did not know p at t^* .” (2004, p. 9)) So, Lewis should agree that the novel thought she entertains is indeed a novel piece of knowledge. And, so, this novel piece of knowledge is a novel piece of knowing-*that* rather than knowing-*how*.

I do not claim that all the novel thoughts Mary can entertain through seeing red express novel pieces of knowing-that. For instance, she can entertain the thought that this is what it is like to see the sky, where *this* denotes the type of experience she had when she saw the tomatoes. This thought is not a piece of knowledge simply because it is false. There are indeed infinitely many false thoughts she can entertain through seeing red. But there is at least one thought which strongly appears to express a piece of knowledge, namely, the thought that this is what it is like to see red, where *this* denotes the type of experience she is having when she is seeing the tomatoes. This thought is clearly true, and Mary appears to be justified in believing it (we do not need a well-developed theory of (immediate) justification to note that Mary *appears* to be justified in believing the particular thought about seeing red which she forms through seeing red). Now, since this new piece of knowledge is a piece of knowing-that, the Ability Hypothesis is false.

Mary acquires the ability to entertain new thoughts through seeing red. What kind of ability is this? Given that concepts are ingredients of thoughts, new thoughts require new concepts. So, Mary acquires the ability to entertain new thoughts about experiencing red because she acquires a new concept of experiencing red. So, the ability to entertain new thoughts is not a pure knowing-how since it requires the possession of new concepts, which are characteristic components of what that-clauses express. According to the Ability Hypothesis, all the abilities acquired through seeing red are of that sort of knowing-how which does not require the antecedent possession of certain pieces of knowing-that. Now, we find an ability acquired through experiencing, the ability to entertain new thoughts, which does not fit the Ability Hypothesis: the ability to entertain new thoughts about experiencing red requires the antecedent

knowledge that this is what it is like to see red, where *this* denotes the type of experience when one sees red. This is another way of putting where the Ability Hypothesis goes wrong.

I assumed, for the sake of the argument, that the abilities Lewis mentions like imagination and recognition are pure cases of knowing-how. But, it seems that even these abilities require the antecedent possession of a new concept of what it is like to see red. Let us ask why in her room, Mary lacks the imaginative and recognitional abilities Lewis points out. A natural answer is that she does not have a certain concept of what it is like to see red. Why do I lack the ability to visually recognize the singer Justin Bieber? Well, because I do not have a visual concept of him, the possession of which typically requires a visual experience of him (or of his picture). Similarly, for Mary's lacking the abilities to recognize and imagine what it is like to see red. But, since Mary's acquisition of the novel concept is through her internally demonstrating her novel experience itself, it is accompanied with the novel thought that this is what it is like to see red, where the demonstrative *this* denotes the type of experience she is having. So, the abilities Lewis himself mentions are not pure cases of knowing-how. The Ability Hypothesis does not hold even for the abilities it takes to be identical with knowing what it is like.

Some of what Lewis writes exhibits a certain misunderstanding about the nature of the discussion between the proponents of the Ability Hypothesis and its opponents. When Lewis considers the suggestion that the causal origin of gaining new abilities is "a special kind of representation" of what it is like to experience red, he writes:

If the causal basis for those abilities turns out to be a special kind of representation of some sort of information, so be it. We need only deny that it represents a special kind of information about a special subject matter. Apart from that it's up for grabs what, if anything, it may represent. (2004, p. 102)

Lewis characterizes “information” in terms of eliminated possibilities. “When we get physical information,” Lewis writes, “we narrow down the physical possibilities, and perhaps narrow down all the way to one” (2004, p. 90). If we understand “information” in this sense, then surely a physicalist has no option but to deny the Hypothesis of Phenomenal *Information*. That is, a physicalist cannot accept that there is a special sort of representation about *a special subject matter* (a special domain of phenomenal, as opposed to physical, possibilities), while there is no obvious reason why he cannot accept that there is a special sort of representation, gained through experiencing, about the good old physical subject matter. But the important question I wish to address is whether a defender of the Ability Hypothesis as such (*not* qua a physicalist as such) can coherently hold that there is a special sort of representation of what it is like to see red, whether it represents a special subject matter or not. As we have seen above, acquiring special representations of experiences through experiencing enables one to have new pieces of knowing-that, a possibility which is denied by the Ability Hypothesis. By granting the possibility of acquiring novel representations of what it is like to see red, Lewis appears to concede what the Ability Hypothesis is originally designed to deny. Lewis is perhaps right that physicalists as such “need only deny that [there is] a special kind of information about a special subject matter”. But he is mistaken to think that a defender of the Ability Hypothesis can allow “a special kind of representation” of what it is like to see red. So, the above passage, as I understand it, displays Lewis’s confusion about what a physicalist as such can perhaps allow and what a defender of the Ability Hypothesis should reject.

The Ability Hypothesis is perhaps right in claiming that knowing what it is like is identical with having certain abilities; however, it is definitely wrong in claiming that the abilities in question are purely knowing-how, their possession does not require having certain

pieces of knowing-that. There is surely at least one piece of knowing-that which Mary acquires through seeing red, that is, the thought she would express by saying “so that is what it is like to see red!” Lewis’s talk about “special kind of representations” of experiences is, in effect, nothing less than a covert admission of the implausibility of the Ability Hypothesis. In the next chapter, I will examine a physicalist strategy held by many philosophers which clear-mindedly allows a special kind of representations (concepts) of experiences (and, accordingly, a special kind of knowing-that) and ask whether physicalist ontology can be plausibly married with a certain sort of dualism of representations (concepts).

4.0 THE PHENOMENAL CONCEPT STRATEGY

4.1 INTRODUCTION

There is something it's like for us to be in certain mental states such as smelling a skunk, tasting a lemon, seeing red, hearing a symphony, and feeling pain. These experiential states, often called "phenomenally conscious" states, are central to our lives.¹⁸ Without them, we would be biological machines exhibiting mere computational ("robotic") intelligence: in a world where the gates of "the senses" are closed, living would be reduced to just being there.¹⁹ If all were darkness within, being alive would be empty.

A central question in the philosophy of mind is whether there is any room for conscious experience in an entirely physical world, which many of us believe *our* world to be an example

¹⁸ Obviously, not all states which a subject is in are conscious (e.g., certain biological states such as digestion). Less obviously, some philosophers argue that there is nothing it is like for a subject to be in some *mental* states (e.g., being in intentional states such as believing and desiring). Let us grant this point and stick with uncontroversial examples of phenomenally conscious states. Moreover, some philosophers claim that putting all experiential states into a single baggage of consciousness is unfair to them: they argue that perceptually conscious states are intentional while bodily states are non-intentional. Surely, however, accepting such a distinction is no hinderance to accepting the phenomenal character (what-it's-likeness) of all these states.

¹⁹ Woody Allen once famously said that "ninety percent of life is just being there". Luckily then we have the remaining ten percent.

of.²⁰ Jackson's (1982) *Knowledge Argument*, as we assessed in chapter two, nicely articulates the problems with physicalism (roughly, the thesis that our world is entirely physical). The argument, as you remember, moves from the premise that complete physical knowledge about the world is not complete knowledge about our experiences to the falsity of physicalism. Consider Mary, a perfect scientist who knows everything there is to know about the physics of sensory perception of red. She knows everything about the relevant surface reflectance properties of those objects called "red", the physical processes through which light is reflected by those objects and reaches the eyes, and the neurophysiologic processes that occur in the head. However, Mary has never experienced red. Now, the intuition is that despite her complete physical knowledge, Mary does not know what it is like to have red experiences. And, since physicalism implies that given her complete physical knowledge, Mary knows everything about experiences of red, physicalism is false.²¹

²⁰ There are familiar problems about how to define "physical" and how to understand "an entirely physical world". Stoljar (2001) makes a distinction between theory-based and object-based conceptions of the physical. According to the theory-based conception, a property is physical if it is *either* a property which physical theory tells us about *or* else is a property which metaphysically supervenes on properties physical theory tells us about. And, according to the object-based conception, a property is physical if it's *either* a property required by a complete account of the intrinsic nature of paradigmatic physical objects (e.g., rocks, trees, planets) *or* else a property which metaphysically supervenes on properties required by a complete account of the intrinsic nature of paradigmatic physical objects. (Stoljar then goes on to argue that these two conceptions need not be co-extensional.) There are two further issues about the theory-based conception: First, how broadly should we take "physical theory" (just physics or physical sciences broadly conceived)? Second, which physical theory, the current one or the one in an idealized future (Crane and Mellor 1990)? For the purposes of this chapter, we need not get bogged down with these questions and can instead be very inclusive: a property is physical if it is either a theory-based (where theory includes all idealized physical sciences conceived as a recognizable descendant of current physical sciences (Jackson (1998)) or an object-based physical property. The problem of locating the place of conscious experience arises even under this very broad conception of the physical.

²¹ Other well-known anti-physicalist arguments include the *Conceivability Argument* (Chalmers 1996) and the *Explanatory Gap Argument* (Levine 1983). The Conceivability Argument moves from the conceivability of zombies (creatures physically identical to us but lacking consciousness) to the falsity of physicalism. The Explanatory Gap Argument moves from the premise that no physical account can constitutively explain why a certain subject has *this* phenomenal property (rather than that property) to the falsity of physicalism. The physicalist position addressed here holds that these arguments can be replied to in a way structurally similar to the reply it proposes against the Knowledge Argument. Hence, the lessons of my discussion of this position's reply to the Knowledge Argument can be plausibly generalized to its replies to other arguments.

The physicalist position I will address in this chapter accepts the premise of the Knowledge Argument but denies that the falsity of physicalism follows from it. I will call it, following Stoljar (2005), “the phenomenal concept strategy” (briefly, PCS).²² PCS acknowledges that one can have complete physical knowledge while still lacking some pieces of knowledge about experiences, and, accordingly, that Mary will gain a genuinely new piece of knowledge (i.e., one which is not entailed by her previous knowledge) when she first experiences red. However, it hastens to add that knowledge is a form of belief, and beliefs (i.e., *what* is believed or the *content* towards which one has the attitude of believing) have transparent and opaque readings. Opaquely, what is believed is specified in terms of the concepts the subject has of the things in the world; and, transparently, what is believed is specified in a way that is insensitive to those concepts. Since a subject may have two different concepts of the very same thing, she may have two genuinely different beliefs, construed opaquely, about that thing while her belief attitude is directed towards the very same content, construed transparently. Accordingly, PCS argues that Mary’s acquiring a genuinely new belief, in the opaque sense, about experiencing red is consistent with the possibility of her having a belief with the same transparent content, and that what is relevant to ascertaining the viability of physicalism is the latter but not the former: physicalism is a thesis, PCS says, about relations of necessitation among physical and phenomenal properties, not about relations of entailment among concepts of those properties.

According to PCS, Mary does not get acquainted with a novel property but acquires a new phenomenal concept of a property she already knew under a physical concept. PCS thus invites us to restrict the lessons of Jackson’s knowledge argument to the level of concepts. A

²² Proponents of PCS include Balog (forthcoming), Block (2007), Hill and McLaughlin (1999), Levin (2007), Loar (2004), McLaughlin (2001), Papineau (2004), Perry (1999), and Tye (2000).

question that naturally arises is what concepts are. The proponents of PCS generally operate with an informal understanding of concepts, and use “concept,” “conception,” and “mode of presentation,” interchangeably. A test which they apply to determine the identity of two concepts is Fregean, one which rests on considerations about cognitive significance;²³ and, the difference between co-extensional concepts is a difference in their cognitive *contents*. The cognitive content of the concept *water* is different from the cognitive content of the concept H_2O even though they denote the same property. Cognitive contents of concepts are usually expressed by (definite or indefinite) descriptions: the cognitive content of *water* is, roughly, what is expressed by the description “the colorless and tasteless stuff which actually fills the lakes.” This is what *water* contributes to the cognitive contents of complex concepts of which it is a constituent. In addition to descriptive concepts, there are perceptual concepts whose contents may not be exhaustively captured by definite descriptions. I may form a visual concept of gold upon seeing an instance of it, which guides me to recognize other instances in subsequent cases. This visual concept differs from my descriptive concept of gold which can be expressed by “the stuff with alloys of which people decorate themselves” (the example is from Loar (2004, p. 223)). Perceptual concepts are a sort of recognitional concepts, about which I will have more to say in what follows.

According to PCS, the cognitive contents of phenomenal concepts (e.g., *experience of red*, *lemony taste*, *skunkish smell*, *excruciating pain*) are not *in any way* analyzable in functional or causal terms. In a widely recognized exposition of PCS, Loar says:

²³ Cognitive significance may in turn be understood in terms of either informativeness or cognitive roles, giving two different versions of the Fregean test. First version: the concepts, *X* and *Y*, are not identical when the identity statement “*X* is *Y*” is informative. The concept *water* is different from the concept H_2O because “water is H_2O ” is informative. Second version: the concepts, *X* and *Y*, are identical when they play the same role in reason and recognition. The concept *water* is different from the concept H_2O because different consequences can be inferred from the statements “that is water” and “that is H_2O ” and also because the circumstances under which these concepts enable the subject to recognize water are different. I will not adjudicate between these two versions (they appear to be equivalent), as our discussion may proceed without such a decision.

It is my view that we can have it both ways. We may take the phenomenological intuition at face value, accepting introspective concepts and their conceptual irreducibility, and at the same time take phenomenal qualities to be identical with physical-functional properties of the sort envisaged by contemporary brain science. (2004, pp. 220-1)

The conceptual irreducibility here is of a particularly strong form:

Phenomenal concepts are conceptually irreducible in this sense: they neither a priori imply, nor are implied by, physical-functional concepts... A phenomenal concept [is] *conceptually independent* of all physical-functional concepts. (2004, p. 220, pp. 224-5)

The sort of conceptual dualism PCS advocates is fundamentally different from the garden-variety conceptual pluralism one typically encounters. It is frequently said that concepts of physics are distinct from concepts of biology or of economics: no physical concept, however complex it is, is conceptually equivalent to a concept of biology, e.g., *reproduction*, or one of economics, e.g., *money*. These points are likely true, but they are not directly relevant to the case at hand. The unavailability of complete conceptual analyses is not the issue here. With regard to biological and economic concepts, we at least have a priori access to their *partial* analyses: these concepts can be given rough-and-ready definitions in terms of the causal tasks encoded in their meanings. Reproduction is the function of an organism's producing another organism, and money is whatever serves the function of being a medium of exchange in a community. Given these characterizations, one can do some empirical research and determine the physical realizers on specific occasions of these high-level properties of reproduction and of being money. For instance, a species-specific explanation of human reproduction invokes cellular processes of certain sorts resulting from fertilization of a woman's ovum by a man's sperm. Similarly, one can figure out that in present communities, the property of being money is realized by coins of certain hard metallic sorts and banknotes. Thanks to partial analyses at hand, any sort of deep ontological or explanatory mysteries that might have surrounded these properties is removed.

As for phenomenal concepts, however, Loar holds that *conceptual irreducibility takes the strong form of conceptual independence*. Loar's idea is that "the phenomenological intuition" that phenomenal concepts cannot be functionally analyzed is compatible with their referring to physical-functional properties: there is no incoherence in subscribing to both physicalist property monism and a *strong* form of conceptual dualism. The combination of these two theses – "having it both ways" – is original.

In what follows, I will focus on how PCS cashes out the conceptual independence of phenomenal concepts. This will require us to take a close look at the cognitive contents of these concepts. A dilemma for PCS will arise: cashing out the notion of conceptual independence in a way compatible with physicalism requires abandoning some manifest phenomenological intuitions or cashing it out in a way compatible with those intuitions requires dropping physicalism. The upshot is that one cannot have it both ways.

4.2 THIN ACCOUNT OF PHENOMENAL CONCEPTS

There are two broadly distinct accounts of phenomenal concepts endorsed by the friends of PCS in the literature. First, there is the *thin* account according to which modes of presentation associated with phenomenal concepts are empty: phenomenal concepts are "characterless" or "blind" concepts which pick out their physical references directly, i.e., without any kinds of modes of presentation associated a priori with the concept. Secondly, the *thick* account argues that there are phenomenal modes of presentation associated with phenomenal concepts, which are to be defined by reference to a kind of an unusually intimate relation they have to their

referents. Thick accounts can basically be thought of as attempts to make sense of Loar's suggestion that an experience "serves as its own mode of presentation" (2004, p. 229). In this section, I will evaluate the thin account, and in the following, I will turn to the thick account.²⁴

The thin account of phenomenal concepts holds that phenomenal concepts pick out their referents in a purely direct way, i.e., without the mediation of *any* (descriptive or not) modes of presentation. It accounts for the directness of reference of phenomenal concepts by reference to the notion of type-demonstrative concepts.²⁵ Type-demonstrative concepts are concepts that designate in all contexts of utterances certain types in whose formation demonstration (or ostension) plays a central role. One sees a color token of a certain type one has not seen before, attends to its qualitative features, and forms the demonstrative concept *that type*. Intuitively, in this case, attending to certain salient features of that color token and (mentally) ostending to the type it belongs to is central to the formation of the concept *that type*. That a certain act of demonstration is required for the formation of the concept at hand is what makes the concept demonstrative, though the type that is designated need not be characterized *as* an object of demonstration.²⁶

²⁴ Chalmers (2007) also distinguishes between thin and thick accounts of PCS, though its definitions of those accounts are different from the ones given here. I agree with Chalmers that "no account of phenomenal concepts is both powerful enough to explain our epistemic situation with regard to consciousness and tame enough to be explained in physical terms" (p. 168). Chalmers goes through complicated arguments which bring in controversial zombie considerations to defend his thesis. I think the thesis itself can be argued for in a more direct manner, something which I attempt to accomplish here.

²⁵ By "the model of type-demonstrative concepts," I don't mean an account of phenomenal concepts which models their semantics fully on the semantics of ordinary demonstratives such as 'this' and 'that'. Such an account would be highly counterintuitive given that the reference of a demonstrative term changes from a context of utterance to another while the references of phenomenal concepts are context-independent. However, for a defense of the idea that phenomenal concepts are as context-dependent as indexical terms like 'I', see O'Dea (2002).

²⁶ Joseph Levine's following remarks are helpful here: "A token-demonstrative picks out an individual, a token of some type, though type information needn't be part of the content of the expression... A type-demonstrative is a demonstrative expression that picks out a type or a kind, as in 'that color' said while holding up and visually

On the thin account, phenomenal concepts are type-demonstratives which have the form “x is one of *that kind* [of experience].” The question is now what their reference-determining mechanisms are. Levin, an astute recent defender of the thin account, writes:

The denotation of a phenomenal type-demonstrative will be the property – presumably physical – that’s causally responsible for the application of that concept in the introspective *recognition* or *reidentification* of an experience as “that (kind) again” or “another of those.” These concepts are taken to refer “directly”; that is, to have *no* reference-fixing “modes of presentation” or Kaplanian “characters” that change reference from world to world... Rather, their references are determined solely by the causal and dispositional relations an individual has to her internal states that are effected by an introspective “pointing in”; that is, by the *fact* that she’s in causal contact with a certain property and is disposed to reidentify it on subsequent occasions. (2007, p. 89)

On this view, the references of phenomenal concepts are determined solely by the causal and dispositional relations a subject introspectively bears to her internal states. What determine the reference of a certain tokening of a given phenomenal concept are what physically causes one to do the tokening and one’s dispositions to recognize certain experiences as falling under the concept.^{27, 28}

attending to a paint chip. One isn’t referring to the actual paint on the surface of the chip, but to the color type of which that surface is a token.” (2010, pp. 176).

²⁷ Levin does not give many examples of what she takes to be phenomenal concepts. She says: “We can think of these type-demonstrative concepts as stored in memory in the form of an “indexed” demonstrative, so that what it’s like to see red can be stored as “that-r” or what it’s like to see green as “that-g.” (2007, p. 89). So, according to Levin, the concept *what it is like to see red* is a phenomenal concept which can be stored as “that-r”. I will sometimes use the concept *red sensation* instead of the concept *what it is like to see red* just for the sake of convenience. As I use the concept *red sensation*, it does not imply any kind of “reification” of phenomenal states, something like sense-data.

²⁸ Levine distinguishes between “ascriptive” and “non-ascriptive” modes of presentation: “An ascriptive mode is one that involves the ascription of properties to the referent, and it’s (at least partly) by virtue of its instantiation of these properties that the object (or property) is the referent... [Non-ascriptive modes] establish relations... “behind the scenes”, not by being cognitively grasped by the subject. The subject’s competence with the term, her “knowledge” of the meaning, consists entirely in her instantiating the requisite [e.g., causal] relation to something in the world” (1998, pp. 457-8). Levine’s non-ascriptive modes of presentation are what the thin account thinks are associated with phenomenal concepts.

It might appear that phenomenal concepts can plausibly be taken as type-demonstratives for the following reason. The acquisition of ordinary type-demonstrative concepts requires the adoption of a particular perspective, and this is surely reminiscent of how phenomenal concepts are acquired. That a condition for the acquisition of phenomenal concepts is having undergone the experiences to which they refer reflects the perspectival nature of phenomenal concepts. And, if phenomenal concepts are type-demonstratives, the epistemic gap between phenomenal and physical truths ceases to be a problem and can be viewed as a reflection of a more general phenomenon that there is an epistemic gap between perspectival and non-perspectival truths.²⁹

4.2.1 Two Types of Direct Reference

Consideration of how recognition proceeds in ordinary cases of demonstrative identification and the comparison of it to the proposed mechanism in cases of phenomenal recognition will provide us a better grasp of the thin account. This will allow us to see the implausibility of the thin account, especially with respect to the epistemic role played by the phenomenology of experiences involved in phenomenal recognition.

Consider the perceptual demonstrative concept *that kind* I have which is roughly formed by an act of mental pointing to a particular car. How do I recognize whether another token-car that I encounter after I form the concept is an instance of the kind designated by the concept? A natural suggestion is that I recognize the objects that fall under my demonstrative concept *as* instances of the type referred to by the concept in virtue of the visual mode of presentation (an

²⁹ See, for instance, Perry (1979) where it is argued that knowing every non-indexical fact about the world won't be enough for one to deduce what time it is *now*.

image, or something close to it) associated with it. There is an epistemically intermediary role played by the visual mode of presentation which enables one to proceed from visual perceptions to visual beliefs such as *that is one of that kind*. Recognitional dispositions, as Loar notes, “are typically linked with capacities to form images” (2004, p. 225). The role of images is to substantiate recognitional concepts in such a way as to ensure that the lifetime of those concepts is not restricted to the lifetime of demonstrative acts which are operative in their formation: by storing images in our memory, we retain recognitional concepts for future usages with which those images are associated. Roughly the natural picture goes like this: I visually perceive an object, (implicitly or explicitly) compare it to the mode of presentation which is associated with the demonstrative concept I already possess, and if there is a sufficient match between the two, I form the recognitional belief *that is one of that kind*.³⁰ My recognitional dispositions to classify things as falling under a perceptual demonstrative concept do not occur blindly or in an epistemic vacuum, but are epistemically grounded in a certain visual (or some other) mode of presentation which is associated with the concept.

The lesson can be generalized: in order for one to recognize an object as an instance of a type designated by a perceptual concept, there must be a mode of presentation associated with the concept. Perceptual recognition is always “recognition *as*”; and, if there is no mode of

³⁰ Admittedly, the picture is *very* rough but it will do for the present purposes. Here I ignore Wittgensteinian scruples about how to recognize images themselves. Moreover, one may be inclined to claim that there are no visual images in play in the recognition of external objects, that they are non-inferentially recognized. It is important to be clear on what is meant by “non-inferential” here. If it means “without a conscious process of making inferences from (visual images)”, then it is no doubt true that visual images do not generally intervene between perceptions and perceptual beliefs. But in order to grant an epistemically intermediary role for visual images, we need not assume that the inferences from them are consciously performed by the subject. If the subject needs something like visual images to *justify* his application of a certain concept to a particular object, then there is a good reason to assign visual images the epistemic role in question. Finally, I remain non-committal about whether visual images are picture-like (“analog”) or sentence-like (“digital”) objects.

presentation available to the subject, it is hard to make sense of how she can perceptually “recognize (anything) as”.

Moreover, it is worth noting that this idea is in accordance with Kaplan’s account of demonstrative reference: “We now have a kind of standard form for demonstrations: ‘The individual that has appearance *A* from here now’, where an appearance is something like a picture with a little arrow pointing to the relevant subject” (1989, p. 225). Kaplan claims that a “standard form” of a perceptual *token*-demonstrative involves an appearance of the individual to which it refers. A natural generalization of Kaplan’s proposal to cover type-demonstratives is this: a standard form of a perceptual type-demonstrative involves a (standard) appearance of the individuals that fall under the type which it refers.³¹

Ordinary perceptual type-demonstrative concepts do not refer directly if “direct reference” is taken epistemically, that is, if it means “without any (visual) modes of presentation, images, or (what Kaplan calls) appearing.” A perceptual type-demonstrative concept represents its referent in a certain way to the subject who possesses the concept, and hence its content is not epistemically empty. This need not imply, however, that the modes of presentation associated with perceptual demonstratives serve any *semantic* function in the determination of references. The idea that ordinary type-demonstratives are epistemically indirect is compatible with the idea that the relation between that concept and its referent is semantically direct. As Levine succinctly formulates, “though the relation between a demonstrative and the object or property demonstrated is semantically direct, the mechanism by which a demonstrative acquires its referent involves mediation by a perceptual representation” (2010, p. 169).

³¹ For an illuminating defense of the idea that perceptual demonstratives are never *pure* in the sense of being unassociated with any descriptive components, see Gertler (2001).

In contrast to ordinary perceptual demonstratives, however, the thin account argues that phenomenal concepts, in the words of Levin, “aren’t supposed to characterize phenomenal properties *as* anything at all” (2007, p. 105, emphasis original). They serve “merely as a pointer directed at (that is, differentially caused by) a type of experience” (2007, p. 90). On the thin account, phenomenal concepts are *pure* demonstratives in the sense that there are no modes of presentation associated with them: they are epistemically empty concepts which directly designate physical properties. On this view, when a subject tokens a certain phenomenal concept, the concept directly refers to the brain state-token B it does because B is what is causally responsible for the tokening and S is disposed to token that concept whenever B is present. Hence, the thin account implies that there is an important epistemic difference between ordinary type-demonstratives and phenomenal concepts, given that only the former are associated with contingent modes of presentation. So, on the thin account, the reference-determining mechanisms of phenomenal concepts are *doubly* direct: not only do they not pick out their referents through a relation of satisfaction of certain descriptive conditions; they are also not accompanied with any modes of presentation.

4.2.2 The Case of Blindsight

A reasonable objection to the thin account is that the idea of pure type-demonstrative concepts is incoherent. One may argue that by definition, concepts represent their referents, and that since representing is always representing as, there cannot be concepts which do not represent as. Alternatively, one may argue that cognitively blind dispositions to register certain pieces of information cannot give rise to concepts, properly so-called, because concepts require some kind

of reflexivity on the part of their possessors about those pieces of information. A reason why we do not ascribe the concept *temperature* to thermometers is that they *blindly* register pieces of information about temperature. So, one may argue, if we are like thermometers with respect to our phenomenal states, then the proper thing to say is not that we have purely demonstrative concepts of those states but that we have no concepts at all. In what follows, I will assume the coherence of the idea of purely demonstrative concepts. My thesis against the thin account is that *our* phenomenal concepts are not purely designative.

An instructive way to start to show what is wrong with the pure demonstrative account is to ask how it accounts for the epistemic difference between blindsight subjects and phenomenally conscious subjects. Blindsight subjects are those individuals who can pass certain perceptual tests by making surprisingly accurate guesses about what lies ahead of them, and even behave as if they are seeing objects, but deny that they are seeing anything at all. Now these subjects can equally well deploy demonstrative concepts in their “introspective” moments to pick out their non-phenomenal intentional states. For instance, a blindsight subject may form the concept *that_R* to pick out the state-type she is in when she is presented red objects, and accordingly develop a disposition to recognize instances of that state as “*that_R again!*” on later occasions. Now surely there is a great difference between the ordinary phenomenal concept *red sensation* and the blindsight subject’s concept *that_R* with respect to the cognitive command they provide to their possessors. The question is whether the pure demonstrative account can account for this difference given that it claims that phenomenal concepts are epistemically blank.

If phenomenal concepts are pure designators without any modes of presentation, then the only way to account for the difference between those concepts and other pure but non-phenomenal concepts like *that_R* seems to be to explain it in terms of differences in what they

refer to. On this view, the phenomenal concept *red sensation* differs from the blindsight subject's *that_R* simply because they pick out brain states with different neural properties. And, the non-identity of these brain states is guaranteed by the fact that on physicalism, phenomenal states are brain states and by hypothesis, the blindsight subject has no phenomenal states. So, there is, on this view, nothing about phenomenal concepts over and above the fact that they denote what they denote.³²

But this cannot be correct. There are surely epistemic differences between phenomenal concepts and non-phenomenal pure demonstratives which go beyond what they refer to if physicalism is true. On the view specified above, there can be no epistemic differences between phenomenally conscious and blindsight subjects with respect to the states they are in when they are presented paradigmatically red objects until they know what the physical referents of their concepts are. However, a phenomenally conscious subject may have absolutely no idea about the brain state which is picked out every time she tokens the thought *that is a red sensation* but she surely knows something about red sensations, namely, what it is like to have them, which is not known by the blindsight subject who merely has a grasp of an epistemically empty demonstrative concept *that_R*. This epistemic difference in what they know has nothing to do with a difference in their knowledge of physical referents.

Let me put the point slightly differently. The blindsight subject has a pure demonstrative concept *that_R* which picks out a certain brain state that occurs when she is presented with red objects. Moreover, the same subject can presumably have another pure demonstrative concept *that_G* which picks out the brain state that takes place when she is presented with green objects.

³² Levin explicitly accepts this: "[If] phenomenal concepts really function like introspectively deployed demonstratives, then all that's needed to distinguish them from introspectively deployed nonphenomenal demonstratives are differences in what they denote" (2007, p. 92).

Whenever there is a red object within her visual field, she will be disposed to token the thought *that is that_R*, while whenever there is a green one, she will be disposed to token the thought *that is that_G*. Moreover, she may also think that *that_R* is not *that_G* though she does not know what the difference consists in until she knows which brain states are designated by those concepts. Now if phenomenal concepts are pure demonstrative concepts as specified above, then from the epistemic point of view of the subject possessing those concepts, the difference between the phenomenal concept *red sensation* and *that_R* is the same difference between *that_R* and *that_G* – that is, a difference merely in referents about which the subject might have no piece of information. But this cannot be right. What the phenomenally conscious subject knows when she knows that red sensations are not *that_R* is not merely the hypothetical information that if physicalism is true, their referents cannot be identical but something much more substantive: namely, what the difference between them consists in, a difference in what it is like to have them. This epistemic difference must be a matter of a difference in the epistemic contents encoded in the concepts *red sensation* and *that_R*, which implies that phenomenal concepts are not pure demonstrative concepts which aren't supposed to characterize phenomenal properties *as anything at all*.

If the pure demonstrative view were correct, our epistemic situation with respect to our own experiences would be the same as the epistemic situation of the blindsight subject with respect to her peculiar non-phenomenal states. This does not necessarily mean that in that case, we would not *have* any conscious experiences or phenomenal states for which there is something it is like to have them. But it means that we would not *know* what those experiences or states are like because our phenomenal concepts by means of which we grasp them would be empty or

unassociated with any modes of presentation.³³ In other words, we would be oblivious to our own experiences. But we are not oblivious to them. We know, for instance, what it is like to have red sensations. Hence, the pure demonstrative view is wrong: phenomenal concepts present experiences in a certain way.

Both the blindsight subject's *that_R* and our phenomenal concept *red sensation* may perhaps be plausibly called “self-directed recognitional concepts” but after that point, they part company. The lesson is indeed simple and predictable: if we know that we are not blindsight subjects, then our phenomenal concepts are not pure demonstrative concepts. Differences in epistemic situations of two subjects must be underwritten by differences in contents of concepts they possess. And, we phenomenally conscious subjects do not only have phenomenal states, but we also know that we have those states and what they are like.³⁴ This is a substantive piece of knowledge which the blindsight subject lacks, and one has it only if one has the required phenomenal concepts. But, a condition for this substantive knowledge is that the concepts used to express it, which are presumably available to the subject who has the knowledge, are not empty.

4.2.3 “The Phenomenological Intuition” Misconceived

The thin account of phenomenal concepts is a version of PCS which aspires to “have it both ways”: physicalism married with “the phenomenological intuition” that phenomenal concepts are

³³ For Gertler (2001), this is indeed what happens in blindsight cases. She argues that they have experiences but they do not know what they are like because of a certain cognitive defect.

³⁴ I do not mean to imply that one cannot be phenomenally conscious without knowing that one is phenomenally conscious. My point is merely that, as a matter of fact, we are not only phenomenally conscious but we also know that we are phenomenally conscious.

conceptually independent of all physical-functional concepts. For the thin account, conceptual independence falls into the picture pretty straightforwardly given that there can be no a priori links between the contents of pure demonstrative and physical-functional concepts simply because the former are empty. Moreover, the cognitive blindness of phenomenal concepts guarantees that physicalist sentiments are not irritated: there are only causal and dispositional connections between physical brain states and tokenings of pure demonstratives. However, this route of having it both ways comes too cheap to be of real worth. Let me explain.

On the pure demonstrative account, phenomenal concepts have thin contents: only references but no modes of presentation. The idea is that phenomenal concepts are blind pointers that aim at our brain states with no substantive conceptualization at all: *all* there is to a phenomenal concept, say, *red sensation* is, on this account, the property it refers to, say, the brain state B_{14} . However, if this is the case, then once we learn that the referent of *red sensation*, which is all there is to it, is B_{14} , our wondering as to how red sensation can be B_{14} is expected to disappear. The causal link that connects the brain state B_{14} with our tokenings of *red sensation* can be ascertained by some scientific investigation; and once it is ascertained, it seems that there remains no intelligible question as to how red sensation (understood merely as the referent of a pure demonstrative concept) can be B_{14} . This is in contrast with what we find in our questions about the relation between our conscious states and brain states. The intelligibility of the question how the sensation as of red can be a brain state does not vanish even when we know everything there is to know about the causal links that connect our tokenings of *red sensation* with brain states. The problem with the thin account is that it misconceives the phenomenologically compelling intuition that it promised to take seriously: the intuition is not merely that phenomenal concepts are conceptually independent but that they are conceptually

independent while, at the same time, being cognitively substantive. Rendering phenomenal concepts cognitively empty does not respect the cognitive grasp we have of the phenomenology of conscious experiences.³⁵

Here are two objections against the claim that the thin account fails to keep the promise of doing justice to the phenomenological intuition. First, it may be argued that the thin account does not eliminate the possibility that one who has a phenomenal concept can still coherently wonder whether the phenomenal concept one has could refer to another brain state rather than what it actually picks out. For instance, one may, it may be claimed, reasonably ask “Might red sensation [understood merely as the referent of a pure demonstrative concept] not have been B₁₅ rather than B₁₄?” My reply is that the objection is a confused one. One cannot coherently wonder whether red sensation might have been something other than what it is just for the same reason why one cannot coherently wonder whether something might not have been identical to itself. Of course, one can reasonably think that the term ‘red sensation’ could be used to denote what it does not actually denote, but that thought has nothing to do with physicalism.

Second, it may be claimed that the phenomenological intuition in question ultimately derives from the perspectival character of phenomenal concepts; and, by taking phenomenal concepts as type-demonstratives, the thin account does justice to this feature of phenomenal concepts. My reply is that the thin account takes phenomenal concepts to be *pure* demonstrative concepts, and pure demonstrative concepts are not perspectival, while what I have said above suggests that perceptual demonstrative concepts, which are not pure (i.e., always associated with a (visual) mode of presentation), are perspectival. A pure demonstrative concept is not

³⁵ Loar explicitly distinguishes between phenomenal concepts we conscious subjects have and “phenomenally blank” concepts of blindsight subjects: “The difference between a self-directed blindsight recognitional concept and a phenomenal concept appears to be that the latter involves a phenomenal mode of presentation while the former conceives its referent in some other, odd, way.” (2004, pp. 230-1).

perspectival because, by definition, it is not associated with any modes of presentation.³⁶ There is no particular point of view whose adoption is required in order to possess pure demonstrative concepts (everybody with a basic conceptual capacity and ability to make demonstrations can have them), which is yet another way of putting the point that phenomenal concepts are not pure.

4.3 THICK ACCOUNT OF PHENOMENAL CONCEPTS

Thus, phenomenal concepts are substantive concepts which present their referents in a certain way. The uniqueness of phenomenal concepts cannot reside in their being purely referential because purely referential concepts cannot accommodate the idea that introspection of our conscious states is what grounds our applications of phenomenal concepts. One who deploys a phenomenal concept to make a reference to an experience does so via attributing certain properties to that experience: a phenomenal concept does not latch onto its referent in a way that might be completely beyond the cognitive reach of its possessor. An individual who knows that an experience falls under the scope of a phenomenal concept knows that it instantiates certain properties ascribed to it by the mode of presentation of the concept.

What makes a phenomenal concept phenomenal? In other words, what is the nature of modes of presentation of phenomenal concepts? Let us first ask an easier question: how do non-phenomenal concepts present their referents? It is instructive to recall here what we have said above about perceptual recognitional concepts. Roughly, a perceptual recognitional concept, say,

³⁶ Hence, Levin's holding pure demonstrative account conflicts with her assertion that phenomenal concepts "pick out their referents from a particular point of view" (2007, p. 90). One of the lessons of the discussion so far has in fact been that one cannot have both.

my recognitional concept *car* is associated with certain visual image(s) or “appearing(s)” (Kaplan) which guides me to recognize what falls under the concept, namely, cars. The descriptive content of the mode of presentation of my recognitional concept can, in first approximation, be expressed by the locution “that which causes me to have this (type of) image.”³⁷ It is easy to see that the same explanation also goes for paradigmatically recognitional concepts such as color and shape concepts. What about quasi-recognitional concepts whose epistemic content is partly constituted by the way they appear to us? Take, for instance, the ordinary quasi-recognitional concept *water*, whose content roughly corresponds to the description “the tasteless, colorless liquid stuff that actually fills the lakes, falls from the sky etc.” What grounds my dispositions to recognize water when I am presented with it is again its appearing to me to be a certain way: “that which causes me to have this (type of) taste experience” (tasteless) or “that which causes me to have this (type of) visual experience” (colorless), where the demonstratives refer to mental samples of the types of experiences in question, which I have previously acquired and thus already have.

In all these cases of recognitional dispositions, “appearing” (or the way things appear to us) plays a central intermediary role between recognition and objects in the external world. In a world in which things do not stand in the appearing relation to its inhabitant subjects, there would be no recognitional concepts as we know them.³⁸ And, what provides recognitional

³⁷ This does not of course mean that in order for me to have the recognitional concept *car*, I should be able to have the concepts incorporated in the description. I think even a small child who does not have the concepts *image* and *cause* can form a recognitional concept of cars. And, this is consistent with the idea that recognitional concepts are generally formed against a further conceptual background. Though a sufficiently thick conceptual background might be taken as a general requirement to have recognitional and demonstrative abilities, possessing the specific concepts that are used to articulate the descriptive contents of particular recognitional concepts is not required to have those recognitional concepts.

³⁸ As Gertler says, “Suppose that there is a man drinking a martini in the corner of a room down the hall. If that man does not *appear* to me in any way, my utterance “that man in the corner drinking a martini” will not refer to him” (2001, p. 316).

concepts with epistemic contents and, accordingly, what in a fine-grained way distinguishes a certain recognitional concept from another are the distinct ways things appear to us. The upshot is that the modes under which recognitional concepts present their referents are appearances.

4.3.1 The Fundamental Claim

Let us now turn to phenomenal concepts which are a kind of recognitional concepts. What is the nature of their modes of presentation? In the cases above, what we find is a difference between things in the external world and the ways they appear to us, which in turn grounds the distinction between the references and modes of presentation of non-phenomenal recognitional concepts. However, phenomenal concepts are concepts *of* experiences which are typically exercised when the subject becomes aware of what those experiences are like via introspection. And, when it comes to experiences, there is no corresponding distinction to be made between things and their appearances. The experience as of red and the way the experience as of red feels are one and the same thing. As Kripke famously stressed, how pain feels is pain. When it comes to concepts of things other than experiences, there is some distance between reference and mode of presentation, the former residing in the world out there and the latter provided to us *through the lens of* our experience of the former (the way it looks to us). However, when the subject matter is experiences, it seems that there is no distance between mode of presentation and reference, given that there is no lens through which we stand to our experiences as we stand to things out there,

that is, we do not have experiences of our experiences in the sort of manner we have experiences of things out there.³⁹

Given that phenomenal concepts are not empty but present their referents in a certain way, and given also that there is no distance between an experience and how an experience appears to the subject, what appears to follow is that the only candidate which can be the mode of presentation of a phenomenal concept is nothing but its referent itself. When the subject becomes aware of the phenomenal character of her experiences, her awareness is not mediated by experiences distinct from those experiences. A certain experience is, then, in some sense to be specified, both the reference and the mode of presentation of a phenomenal concept. The ontological identity of experiences and their appearances grounds an unusually tight semantic connection between phenomenal concepts and their references. Metaphorically, it is as if the commonplace Fregean duality of mode of presentation and reference collapses into a peculiar Russellian unity within the realm of phenomenal concepts.

The thick account essentially rests on the appreciation of this point: the special nature of phenomenal concepts does not derive from their being cognitively empty but from their contents' being loaded with their references.⁴⁰ Hence, the following remarks:

A phenomenal concept has as its mode of presentation the very quality it picks out. (Loar 2004, p. 231)

³⁹ What about the so-called inner-sense theory? Here is a conditional I hold: If the inner-sense theorists (Armstrong 1968, Lycan 1996) argue that we have experiences of experiences just in the same sort of way that we have experiences of things out there, then they are wrong. A good reason to think that there are no higher-order experiences is that we have no concepts of those experiences, and it is reasonable that if we had higher-order experiences, then we would have concepts of them just like we have concepts of first-order experiences. For powerful objections against the inner-sense theory, see Carruthers (2000).

⁴⁰ I don't claim that every proponent of the thick account holds the above account according to which there is no distance between experiences and their appearances. I use that account to present *at least* one way to motivate the thick account.

There is only one property in play when a phenomenal concept refers to a phenomenal property: namely the phenomenal property itself. No further property mediates between referring concept and referent. (Papineau 2004, pp. 104-6)

The mode of presentation associated with a phenomenal concept is the very phenomenal quality to which the concept refers. (McLaughlin 2001, p. 326)

Phenomenal concepts are partly constituted by the phenomenal experiences they refer to. On this view, a current phenomenal experience is *part* of the token concept applied to it. (Balog “Acquaintance and the Mind-Body Problem,” p. 18)

A phenomenal concept is individuated with respect to fundamental uses that involve the *actual occurrence* of phenomenal properties at the time of those fundamental uses. In these fundamental uses, a simultaneous actually occurring experience is used to think about that very experience... [An] instantiation of a phenomenal property is used in the concept to pick a phenomenal property (a type). (Block 2007, pp. 252-3)

Loar, Papineau and McLaughlin talk about phenomenal qualities where Balog talks about phenomenal experiences. This difference is not substantial: all these philosophers use talk of qualities and experiences almost interchangeably, which is especially clear in Block’s quotation. Another issue is what the mode of presentation of a phenomenal concept is: a property or its instantiation? Loar and McLaughlin explicitly say and Papineau’s quotation suggests “a phenomenal property”, while Balog’s and Block’s quotations appear to suggest “an instantiation.” I will not spend time to investigate this otherwise important point since my discussion may proceed without it,⁴¹ and will take the following thesis as the fundamental claim of the thick account:⁴²

⁴¹ Loar talks as if both claims are equivalent: “We might say that a phenomenal concept has as its mode of presentation the very phenomenal quality that it picks out. We might also say that phenomenal concepts have “token

(T) The mode of presentation of a phenomenal concept is the property it refers to.

I agree that phenomenal concepts are unique in the manner that is indicated by the remarks of the defenders of the thick account. And, the good news for PCS is that taking phenomenal concepts as cognitively substantive concepts appears to account for the intuition that there is a troubling sort of epistemic gap in the phenomenal case. Phenomenal concepts are not only radically different from other concepts, but they are different in a substantive way: the conceptual isolation of phenomenal concepts is grounded in the fact that their contents are loaded in a sort of manner the contents of physical-functional concepts cannot be.

4.3.2 Problems with the Thick Account

There are two problems about the thick account I would like to point out.

3.2a. Given the characterization above, we can take it that the thick version of PCS consists of three claims: (i) phenomenal concepts refer to physical/functional properties, (ii) phenomenal concepts are isolated from all physical/functional concepts, and (iii) phenomenal concepts are substantive concepts whose modes of presentation are constituted by their references, which is (T). Claim (i) makes the thick account physicalist, claim (ii) makes it a version of PCS, and

modes of presentation” that are non-contingently tied to the phenomenal qualities to which those concepts point: particular cramp feelings and images can focus one’s conception of the phenomenal quality of cramp feeling” (2004, pp. 231-2). A central question is how we should understand the idea of “non-contingent tie.”

⁴² There are important questions an adequate theory of phenomenal concepts should answer: how many kinds of phenomenal concepts are there? (For instance, Chalmers (2003) makes useful distinctions among what he calls pure, standing, and direct phenomenal concepts.) What is the relation between constitution and reference? Does constitution by itself determine reference (Balog) or do we need a teleological/causal account (Papineau)? Is “quotation” a good model for phenomenal concepts? Is an “experience operator” needed for the thick accounts to work? I will address these and similar questions in another work, as nothing in my present argument rests on them.

claim (iii) distinguishes it from the thin version of PCS. I will argue that (i)-(iii) are not a happy triad.

Holding (i) and (ii) while rejecting (iii) is a coherent position. In fact, a good example is the thin account. And, holding (ii) and (iii) while rejecting (i) is a form of property dualism, about the coherence of which there is no serious debate. Finally, accepting (i) and (iii) while rejecting (ii) is also coherent. In this case, phenomenal concepts are similar to what some philosophers take concepts like *hydrogen* to be: physical concepts whose modes of presentation are identical to their references.⁴³

Now, (i) and (iii) imply that both the modes of presentation and the referents of phenomenal concepts are physical/functional properties. But if this is so, then what reasons do the proponents of the thick account have to hold (ii)? If both the modes of presentation and referents of phenomenal concepts are physical/functional, then they are not isolated from physical/functional concepts but *are* themselves physical/functional concepts: if any concepts have a right to be called ‘physical/functional’, they are surely those concepts whose both modes of presentation and referents are physical/functional. If (i) and (iii) are true, then contra (ii) phenomenal concepts *are* indeed physical/functional concepts *par excellence*.

Instead of asking “what is the sense in which phenomenal concepts are isolated?” we may ask: What is the sense, according to the thick account, in which phenomenal concepts are *phenomenal*? Their ‘phenomenality’ is, on the thick account, not a matter of their referents (because of (i): the referents of phenomenal concepts are physical/functional properties) but is a matter of their modes of presentation. (If the phenomenality of phenomenal concepts were due to their referents, they would not be phenomenal in the first instance but be physical/functional.)

⁴³ For a defense of the idea see Chalmers (1996) and McGinn (2001), and for an attack against it, see Boyd (1980), Hill (1991), and Block (2007).

However, according to (iii), their referents *are* their modes of presentation. And, this is problematic: the thick account appears to claim that the phenomenality of phenomenal concepts both comes and does not come from their referents.

There are (at least *prima facie*) inconsistencies in the thick account. The proponents of the thick account want to combine certain intuitions about phenomenal concepts with an antecedent commitment to physicalism. But what the above reasoning appears to show is that this is not achievable. In order to appreciate anti-physicalist intuitions about phenomenal concepts incorporated in (ii) and (iii), one is forced to drop the physicalist thesis stated by (i).

A question: why do the friends of the thick version fail to notice these inconsistencies in their position? A main reason is, I suggest, suboptimal presentations of the thick version of PCS by its own proponents, presentations which admit a number of interpretations. Consider what McLaughlin says:

New wave materialists hold that (a) phenomenal concepts directly refer to phenomenal properties, that (b) phenomenal concepts are not physical-functional concepts, and, that (c) phenomenal properties are identical with physical-functional properties. These three are consistent. And new wave materialists hold that they are all true.⁴⁴ (McLaughlin 2001, p. 329)

In this formulation, (b) fails to adequately characterize the thesis the thick version is committed to because it does not distinguish the position from the thin version. Moreover, McLaughlin's (a) and (c) together amount to the idea that phenomenal concepts directly refer to physical-functional properties. This is almost equivalent to my (i) above except for McLaughlin's qualification "directly." Here the qualification "directly" is supposed to do the work of my (iii).⁴⁵

⁴⁴ The passage from McLaughlin is quoted from his response to Horgan and Tienson's (2001), where what I call the thick version of PCS corresponds to what they name "new wave materialism."

⁴⁵ Since McLaughlin clearly states that the mode of presentation associated with a phenomenal concept is the very quality it refers to, it is reasonable to construe his claim of referential directness of phenomenal concepts as (iii).

But again it fails to describe the thick version in a satisfactory manner given that the thin version also accepts the combination of (a) and (c) understood in a certain way. McLaughlin's claim that (a)-(c) are consistent is true but phenomenologically inadequate when they are interpreted "thinly," and it is phenomenologically adequate but false, as I have argued, when they are interpreted "thickly." So, a reasonable suggestion is that failure to notice the inconsistency can be traced back to a lack of a clear-eyed attention to the distinction between thin and thick versions of PCS.

3.2b. Now, let's focus on (iii): the idea of the identity of modes of presentation (or cognitive contents) and referents of phenomenal concepts. If (iii) is true, then the cognitive contents of phenomenal concepts mark a stark contrast to the cognitive contents of cognitively substantive concepts of an ordinary sort such as the concept *water*. The natural kind concept *water* is substantive in the sense that it presents what it refers to in a certain way to the subject who possesses the concept. But the cognitive content of the concept *water* does not involve as part of it water. It is even hard to see what it would mean for the cognitive content of a concept to involve water as a part: water just does not seem to be a sort of thing that can be a part of concepts as we know them. Rather, the mode of presentation of the concept *water* consists of contingent properties of water such as falling from the sky, filling the lakes and seas, being liquid at room temperature, and being colorless. That its mode of presentation consists of its contingent properties is what is responsible for what may be called the cognitive intransparency between the concept and its referent: one cannot read off what the essence of the reference of the concept *water* consists in merely on the basis of having the concept. The cognitive access one has to the features of water by virtue of having the concept *water* does not enable one to know that water is H₂O. However, when it comes to phenomenal concepts, there is a sort of unusually strong sort of

cognitive transparency that holds between them and their referents simply on account of the constitutive relation between the two. Metaphorically, it is as if the reference of the phenomenal concept *pain* lies before the mind's eye of the subject solely in virtue of his possession of the concept. So, what naturally seems to follow is that the cognitive access one has to the phenomenal property of being in pain by virtue of having the phenomenal concept *pain* enables one to know what the essence of that property consists in: the referent is *in* the concept the subject apprehends. And, the problem for physicalism is that the essence of the phenomenal property of being in pain so apprehended appears to be something different from neural or functional sorts.

Another way of putting this problem is by raising the question "what reasons do we have to think that mind-brain identities are a posteriori if a physical/functional property is a part of the mode of presentation of a phenomenal concept?" Suppose that physicalism is true, and the phenomenal property of being in pain turns out to be a posteriori identical to the brain state B_{52} . Then, the thick version implies, in view of (T), that the brain state B_{52} is a part of the mode of presentation of the phenomenal concept *pain*. But if this is so, why does not the possession of the phenomenal concept *pain* enable me to know that pain is B_{52} ? Given that the cognitive access to references which is provided by having phenomenal concepts is direct in the sense of being unmediated by contingent modes of presentation, then the natural expectation is that just in virtue of having the concept *pain*, the subject knows that pain is B_{52} . But the possession of the concept *pain* does not bestow such a piece of knowledge upon the subject. And, this is not because having the concept *pain*, the brain state B_{52} lies before the mind's eye of the subject but he lacks the physical concept B_{52} required to entertain the thought that pain is B_{52} . It is rather

because what lies before the mind's eye of the subject when he tokens the concept *pain* does not appear to be anything like B₅₂.

In the writings of the proponents of the thick version who are sensitive to this problem, one finds statements like the following:

Phenomenal concepts pick out their referents in a direct and substantive way; via exemplars of the very phenomenal properties they refer to... But if they pick out their referent in this way then, while they afford an insight into the essence of the referent (by exemplification), they will not afford any clue as to the fundamental nature of the referent. (Balog "In Defense of Phenomenal Concept Strategy," p. 9)

In this passage, Balog appears to be claiming that while in virtue of having a phenomenal concept, the subject gains an insight into the essence of its referent, she might still fail to know its fundamental nature. However, the difference between the essence of a phenomenal property and its fundamental nature is, mildly put, very puzzling. I do not see how it can be satisfactorily cashed out. Moreover, even if we succeed in making sense of the distinction between essence and fundamental nature, and take it that the essence of pain is, say, how it feels and its fundamental nature is, say, a brain state to be empirically discovered, then we end up with (a timid version of) a dual aspect theory of mental properties, a theory which the proponents of PCS regard as metaphysically extravagant and are initially determined to avoid.

4.4 CONCLUSION

I have argued against a novel physicalist position ("the phenomenal concept strategy" (PCS)) which holds that physicalism is consistent with the conceptual isolation of phenomenal concepts

from physical-functional ones. The conceptual isolation in question can be accounted for in two different ways, and hence there are two different versions of PCS: thin and thick versions. I have shown that neither option is viable, and hence PCS cannot help the physicalist cause.

Despite its failure as a response to anti-physicalism, PCS has an important insight: it clearly recognizes that we need a theory of the concepts under which we grasp our phenomenal experiences. While phenomenal concepts play an indispensable role in our understanding of the nature of experiences, a comprehensive theory of the distinctive semantic and epistemic features of these concepts is missing. And, since phenomenal concepts are formed on the basis of the epistemic relation one has to one's own experiences, an account of the former requires an account of the latter. If phenomenal concepts are fundamentally distinct from physical/functional concepts, then we can safely assume that one's cognitive access to one's own experiences is fundamentally different from one's cognitive access to physical/functional properties. This is an idea physicalists are uncomfortable with, but perhaps one they have to learn how to live with.

A moral of my discussion can be summarized by a disjunctive slogan: either deflationism (roughly, the thesis that consciousness is a priori analyzable in physical/functional terms (Block 2002)) or property dualism. A non-deflationist physicalism is a dead end, whose failure brings out an important lesson: the recognition of a category of concepts of phenomenal experiences whose cognitive contents bear no a priori links to the cognitive contents of physical/functional concepts is fatal to physicalism. But, it seems that we *have* such concepts; and if so, one may argue from the mere existence of phenomenal concepts to the necessity of property dualism, an argument the full articulation of which demands an essay on its own.

5.0 ACQUAINTANCE AND PHENOMENAL CONCEPTS

5.1 PHENOMENAL CONCEPTS, STATES, AND QUALITIES

What are phenomenal concepts? What is the sense in which some concepts are phenomenal? There are actually two different senses in which some concepts are called phenomenal. According to the first sense, phenomenal concepts are concepts of phenomenal states, and a state is phenomenal when it instantiates a certain phenomenal quality (“raw feeling”) or, equivalently, there is something it is like to have that state. Being in pain, for instance, is a phenomenal state which has a certain characteristic phenomenal quality, which renders it qualitatively different from other phenomenal states such as tasting a lemon or smelling a rose. As we shall see shortly, the second sense is more demanding.

Some philosophers argue that there are non-phenomenal mental states like believing and thinking for which there is nothing it is like to have them: believing that the capital of Italy is Rome, for instance, does not have any characteristic feel associated with it. Being in pain feels a certain way (it *hurts*), but what can possibly be the feeling associated with believing that the capital of Italy is Rome? On the first sense, phenomenal concepts are concepts of mental states like pain but, as some argue, not of states like believing. A question: is the concept *red* a

phenomenal concept in this sense of the term? In so far as it is a concept of external things but not of phenomenal states, it is not a phenomenal concept. A concept that is applied to apples or strawberries cannot be a phenomenal concept simply because apples and strawberries are not phenomenal states. But we also sometimes talk about red experiences; for instance, when a normal perceiver looks at a red apple and visually experiences its color, it is natural to say that he has a red experience. Of course, experiences or phenomenal states are not red in the same sense in which external objects are red. Phenomenal redness as a property of phenomenal states is a different property from physical redness as a property of external objects, and a corresponding distinction is to be made between phenomenal concepts which are applied to phenomenal states and perceptual concepts which are applied to external objects.

Phenomenal concepts so understood seem uncontroversial: does anybody really believe that pain does not hurt or we have no concept of pain? Phenomenal concepts in this sense commit one to a weak sense of phenomenal realism which acknowledges the existence of phenomenal states and their qualitative (not merely relational or functional) features. One may argue that phenomenal states or qualitative feels do not have the features which are attributed to them by some philosophers, e.g., they are introspectible or the subject's epistemic access to them is in some sense privileged; but that is another issue.

According to the second and more controversial sense of the term, all phenomenal concepts are concepts of phenomenal states but not all concepts of phenomenal states are phenomenal concepts: phenomenal concepts are, in some sense, special among concepts of phenomenal states. The state of being in pain, for instance, may be picked out by a phenomenal or a non-phenomenal (e.g., functional) concept. A functional concept of pain characterizes it as a state which has a certain (typical) cause and effect. For instance, the concept *a mental state*

which is typically caused by tissue damage and typically causes avoidance behavior is a functional but not phenomenal concept of pain. A phenomenal concept of pain, on the other hand, characterizes it as a state which feels a certain way. You accidentally hit your thumb with a hammer, you feel pain. You can think of that particular pain in at least two different ways: a state that is caused by hitting your thumb with a hammer or a state that has a certain feel; both of them are ways of thinking about that phenomenal state but only in the latter, you think of that pain under a phenomenal concept.

In what follows, I will use the term ‘phenomenal concept’ in the second sense. Phenomenal states have certain properties: they have functional properties and they have phenomenal qualities. And, just as we can use the functional properties of phenomenal states to pick out those states, we can also use their phenomenal qualities for the same job. When we think of pain as a mental state that has a typical cause and effect, we think of it with reference to its functional properties. And, when we think of it as a mental state that feels a certain way, we think of it with reference to its phenomenal property. Compare: water has the following properties, it is the liquid that fills our lakes and it boils at 100 degrees. And, I can think of water under the concept *the liquid that fills our lakes* or the concept *the liquid that boils at 100 degrees*. A familiar rule is that we can use a certain (unique) property of what we think of to think of it; and using a what-it-is-like property of a phenomenal state to think of that state is in conformity with that rule. (More on the idea of using properties later.)

So far I have talked about phenomenal concepts as concepts of phenomenal states. But, not only do we have concepts of phenomenal states, we also have concepts of phenomenal qualities instantiated by those states; and, a phenomenal quality is what gives the particular what-it-is-like character to the state that instantiates it. I will also call those concepts that pick out

phenomenal qualities by using those qualities themselves phenomenal concepts. A phenomenal concept of a phenomenal state has the complex form *a state that has P*, where *P* is the phenomenal concept of the phenomenal quality the state bears, and thus phenomenal concepts of states are derivative from phenomenal concepts of qualities. In this work, I am mainly interested in the special semantic and epistemic features of *P*, and accordingly my interest in concepts of phenomenal states is mainly an interest in concepts of qualities. For simplicity's sake, I will ignore the distinction between states and qualities here, but it is important to be aware that they are different.

5.2 ACQUAINTANCE AS AN UNMEDIATED EPISTEMIC RELATION

What is the epistemic relation the subject who is in a certain phenomenal state has to that state? It is important to note that that relation is of a different kind from the epistemic relation the subject bears to external objects. One's epistemic access to external objects is mediated in the sense that one's beliefs or knowledge of them are based on the effects they make on one's mental states. My belief that the pen I am now holding is red depends on the sensory experience I have when I look at that pen; and, similarly my belief that there is a man sitting at the table across me right now is *inter alia* dependent upon the sensory experience I have when I turn my perceptual attention to that particular table. These beliefs about external objects are in part justified by my experiences: my experiences are my pieces of evidence for them. To the question "why do you believe that your pen is red?" my answer would roughly be "because it is how it appears to me."

How things perceptually appear to us normally provides reasons for our beliefs about those things.

We do not normally ask ourselves whether our occurrent beliefs about external objects are justified or what pieces of evidence we have for those beliefs. Accordingly, there is generally no explicit process of inference that we go through when we form perceptual beliefs about objects surrounding us. Upon seeing a dog, Carl comes to believe that a dog is running towards him. How does he come to hold that belief? Does he infer it from the belief that he sees a dog that is running towards him or maybe from the belief that he seems to see such a dog? As a matter of fact, Carl's formation of his belief in question is automatic in that there is no explicit inference that Carl makes before he holds the belief. In this sense, his belief about the dog in the park is different from his belief about the particular angle of a certain triangle which he calculates given his information about the other two angles of the triangle. Carl does not seem to make any conscious inference to his belief about the dog analogous to the one he makes when he calculates the angle, nor does he question what evidence he has for his belief, nor does he ask whether his pieces of evidence are strong enough to hold it. But still this apparently automatic and, in this sense, non-inferential character of the formation of his belief does not show that Carl does not have any justification for it. If forced to answer why he believes that a dog is running towards him, Carl would appeal to his perceptual experiences to justify his belief and say something along the lines of "because it is how it appears to me". How it perceptually appears to Carl is his reason for his belief about the dog.

Carl's belief that a dog is running towards him is non-inferential in that he does not acquire the belief by making explicit inferences from other beliefs. His belief is not inferentially *acquired*. But this does not mean that his belief is not inferentially *justified*. In fact, I believe that

beliefs about external objects require inferences from beliefs about experiences in order to be justified, for which Carl's appeal to his relevant experience to justify his belief is presented as a case above. That Carl's belief is non-inferentially acquired is compatible with its being inferentially justified.

The justificatory support one's experiences provide for one's perceptual beliefs can be diminished or undermined by other considerations. If Carl had taken some hallucinatory pills before the relevant experience, then the evidential support his experience provides for his belief about the dog would not be as strong as it would otherwise be: a normal perceiver is more justified in taking his experiences at face value than the one who is prone to hallucinations. Or, the evidential support that seeming to see a dinosaur provides for believing that there is a dinosaur is undermined by the strength of the evidence I have for believing that dinosaurs are extinct: I would not be justified in believing, after seeming to see a dinosaur, that there is a dinosaur. Some philosophers believe that perceptual beliefs are never justified if one is not in a position to rationally exclude all skeptical scenarios like Descartes' Evil Demon or more recent scenarios about being a brain in a vat. Even if we accept skepticism about the epistemic status of all perceptual beliefs, we still need to make a distinction between the evidential role, say, seeming to see a red ball plays with respect to, say, believing that there is a red ball and believing that there is a green ball: even if we cannot ultimately exclude relevant skeptical scenarios, it is still plausible to say that other things being equal, seeming to see a red ball provides some evidence for believing that there is a red ball but not for believing that there is a green ball. So, even skeptics should draw a distinction here. What I will say is that certain experiences provide *prima facie* evidence for certain beliefs but not for others, even if there is no perceptual belief which is *ultima facie* justified because of the impossibility of eliminating skeptical possibilities.

Does the idea that beliefs about external objects are *prima facie* justified by experiences conflict with what is called direct realism? In a relevant sense of the term, direct realism is the position that the immediate objects of our perceptions are external objects but not our mental objects or states. According to direct realism, what I directly see when I look at my room are a computer and several books on the floor, but not the mental images I have of them or the perceptual states I am in. Direct realism is a position with some intuitive support, but its philosophical success depends on whether it can plausibly respond to certain challenges like the ones from illusion and hallucination: how is hallucination (e.g., an after-image) possible, for instance, if what I directly see when I have a perceptual experience is always an external object? Whether direct realism can keep its intuitive appeal without being philosophically naive is an issue that I will not even attempt to settle here. Let me just say that direct realism is consistent with the idea that experiences justify perceptual beliefs given that the temporal order of immediate awareness need not determine the evidential order of what justifies what: even if what I am directly aware of are external objects, what justifies my belief about those objects can still be my mental states. It is not a mental image or state that I directly perceive when I look at my room, but when asked why I believe that there is a computer on the table, I appeal to my mental state as a piece of evidence for my belief. I think that this is a coherent position.

What is the nature of the epistemic relation a subject has to his own perceptual states? My perceptual phenomenal states (that is, the phenomenal states that I am in when I perceive external objects) provide reasons for holding my beliefs about external objects, but what reason do I have for holding beliefs about those states themselves? What reasons do I have for believing, say, that I am now having a visual experience as of red? The obvious reply is that my beliefs about my visual states are justified by those visual states themselves. If asked what

evidence I have for believing that I am now having an experience of a certain quality, my reply would simply be “I am just having it”. Consider a curious interlocutor who is not satisfied by my answer. He says: “You say that your evidence for your believing that you are having a certain experience is your having it. But when you say “I am just having it”, what you do is merely to *restate* your belief in question that you are having a certain experience, but you are not providing any evidence for that particular belief. What I want to learn is more: please tell me what evidence you have for believing that you are having a certain experience, and please do it without merely restating that belief because I already know that you have that belief.” What can I say in response? I can appreciate the point of the interlocutor: while he waits for a piece of evidence for my phenomenal belief in question, I do not seem to provide one. But I also have a point: having a certain experience *is* my evidence for that belief. I think that this dialectic tie between me and the interlocutor is a manifestation of something philosophically interesting: the interlocutor's search for an evidence for phenomenal beliefs other than experiences takes for granted that my epistemic relation to my experiences is the same kind of relation I have to external objects – when it is not. If my epistemic relation to my experiences is direct in that my beliefs about those experiences are (*prima facie*) justified by having those experiences, then the interlocutor's search for evidence other than my experiences for my phenomenal beliefs is misplaced.

One may argue that if the only evidence I can bring forth for my beliefs about my phenomenal states is merely to restate them that shows that those beliefs are indeed unjustified. Consider someone who believes that there are ghosts, and when asked what evidence he has for believing that there are ghosts, his reply is “there just are”. In this case, we would not be inclined to say that his belief about ghosts is justified. The situation concerning beliefs about phenomenal states seems similar. When asked what reason I have for believing that I am having a certain

experience, I say “I am just having it”. One may then claim that given the similarity between the two, if the belief about ghosts is unjustified, the belief about phenomenal states is unjustified. The problem with this view is that it denies the obvious: having experiences justifies beliefs about having experiences. Surely, this is not to say that every phenomenal belief is justified. There may be a subject who is so confused that while he is not having a reddish experience, he believes that he is because he believes that Earth is secretly invaded by aliens. Or, there may be another confused subject who believes that he is having a reddish experience for the wrong reasons: while he is actually having such an experience, the reason he provides for his belief about his phenomenal state is, strangely enough, that two plus two is four. In these cases, it is natural to say that the beliefs these subjects have about their phenomenal states are not justified. But these cases are compatible with the point I am trying to make: one’s being in a certain phenomenal state itself provides *some* evidence for believing that one is in that state.

The epistemological questions “how do you know that there are external objects?” and “how do you know that you are in a phenomenal state (e.g., that you are in pain)?” are answered in different ways. The first question requires an appeal to one’s phenomenal states: I know that there is a red object in front of me because I know that I am having an experience of a certain quality and I know that this experience is normally caused by red objects. My knowledge of a red object in front of me is thus mediated by my knowledge of something other than that object, namely, my experience. I know that there is a red object in front of me in virtue of knowing that I am having an experience of a certain quality. For variety, take the case of hearing. I hear a sound, which happens to be the sound of my mother, but how do I know that it is the sound of my mother? I know that what I hear is the sound of my mother in virtue of knowing that I have a certain kind of auditory experience and also knowing that that kind of experience is normally

caused by my mother. In contrast, my knowledge of my auditory experience is not mediated by my knowledge of something other than my experience. I do not know that I am having an experience of a certain quality in virtue of knowing something else, I *just* know it. It is worth noting that the structural difference between the two replies is nicely reflected in skeptical scenarios concerning our knowledge of external objects and phenomenal states: it is widely agreed that skepticism about external objects is much more plausible than skepticism about phenomenal states. For all the evidence I have, I may be a brain in a vat which is triggered to think that it is surrounded with certain material objects while there may be no such objects. In other words, all the evidence I have for believing that there is a book before me is compatible with my being a brain in a vat. But, my knowledge that I am now having a certain experience is not affected by, hopefully counterfactually, my being a brain in vat. Even if I am a brain in a vat, I am still having this experience. It is not for no reason that skeptical scenarios generally take for granted one's knowledge of one's phenomenal states.

I will call the special type of epistemic relation conscious subjects bear to their phenomenal states *acquaintance*. Acquaintance is an unmediated epistemic relation which makes a subject's beliefs about his own phenomenal states justified not in virtue of his beliefs about other things but by having those states themselves. As I define the term, acquaintance itself is *not* a special type of belief or knowledge but an epistemic relation that provides the experiential basis for unmediated beliefs or knowledge; hence, acquaintance is *not* knowledge by acquaintance. A subject who is acquainted with his own phenomenal states is, in a familiar sense of the term, in a position to have epistemically unmediated beliefs about those states; and knowledge of a particular phenomenal state by acquaintance is knowledge which is had by being acquainted with that state.

Let me illustrate my notion of acquaintance by example. Consider two small children, Mark and Jennifer; both are having the same dental operation at the moment. Mark feels intense pain but Jenny suffering from a certain disease known as CIPA (Congenital Insensitivity to Pain with Anhidrosis) is physically incapable of feeling pain. I will say that Mark is acquainted with pain while Jenny is not: Mark is and Jenny is not in a position to have unmediated beliefs about a certain phenomenal state (that is, being in pain). But being in a position to have unmediated beliefs is not to have unmediated beliefs. A natural idea is that beliefs properly so-called require application of concepts, and small children may not have the required conceptual repertoire to form the belief that they feel pain. In order for Mark to believe that he feels pain or that pain has Q , where Q is a concept of the particular unpleasurable feeling he has while having the dental operation, he should have such concepts as *pain*, *feeling*, or Q . Typically, small children feel pain without knowing that they feel pain. If so, Mark is acquainted with the phenomenal state of being in pain but he may still not know pain by acquaintance, if he does not have the phenomenal concept *pain* to apply to that state.

I think some of the philosophical mystery surrounding the notion of acquaintance can be dispelled if the distinction between acquaintance and knowledge by acquaintance is suitably drawn. A common worry is that acquaintance forces one to acknowledge the dubious idea of non-conceptual knowledge, but this worry can be met by the distinction at hand. It is true that getting acquainted with a phenomenal state is non-conceptual in that it does not require the application of a particular concept, but acquaintance itself is not knowledge by acquaintance. A commitment to knowledge by acquaintance need not commit one to the existence of non-conceptual knowledge; knowledge by acquaintance like other types of knowledge is conceptual. The difference between knowledge by acquaintance and other types of knowledge does not rest

on the former's being non-conceptual while the latter is conceptual but on the kinds of justifications we have (that is, mediated or unmediated by other pieces of knowledge) for them.

It is worth noting that knowledge by acquaintance, as I define the term, is different from what is sometimes referred to as knowing things as opposed to knowing facts (or truths). Russell sometimes defines his original distinction between knowledge by acquaintance and knowledge by description in terms of knowing things and knowing truths. For instance, he says:

Knowledge of things, when it is of the kind we call knowledge by acquaintance, is essentially simpler than any knowledge of truths and logically independent of knowledge of truths. (1912, p. 46)

For Russell, the things that can be known by acquaintance are mainly mental particulars which he calls sense-data but not physical objects. For Russell, physical objects can only be known by a description which applies to them, that is, by knowing certain truths about them, but not by being acquainted with them. He writes:

All our knowledge of the table is knowledge of truths, and the actual thing which is the table is not, strictly speaking, known to us at all. We know a description and we know that there is just one object to which this description applies, though the object itself is not directly known to us at all. In such a case, we say that our knowledge of the object is knowledge by description. (1912, pp. 47-8)

In these passages, Russell appears to draw a fundamental distinction between knowledge of things and knowledge of truths, and also argue that knowledge by acquaintance is knowledge of things and knowledge by description is knowledge of truths. A natural interpretation is that for Russell, while only some things can be objects of knowledge by acquaintance (sense data but not physical objects), all knowledge by acquaintance is knowledge of things. Moreover, Russell

claims that knowledge by acquaintance is “essentially simpler than” and “logically independent of” knowledge by description. If knowledge by description involves application of concepts, a plausible interpretation of Russell's claim about “simplicity” and “logical independence” is that knowledge by acquaintance is non-conceptual.

Now, it is true that we sometimes talk about knowing things: I know Istanbul, I know the joy of watching soccer, and I know Emre (my brother). Istanbul, the joy of watching soccer, and Emre are not the kinds of things that can be true or false; but, it would be too fast to derive from this the conclusion that my knowledge of them is not knowledge of truths or it is non-conceptual knowledge. I know Istanbul *as* a city, I know the joy of watching soccer *as* one of my favorite pleasures, and I know Emre *as* my brother. That is, my knowledge of these particulars involves application of certain concepts (the concepts such as *city* and *brother*) to them. From the mere fact that the object of my knowledge is something that cannot be true or false, one cannot derive the conclusion that my knowledge is non-conceptual. A similar point also applies to one's knowing one's own phenomenal states given the distinction at hand between acquaintance and knowledge by acquaintance. The mere fact that the object of my phenomenal knowledge, which is a phenomenal state, is something that cannot be true or false, does not show that my knowledge is non-conceptual. Acquaintance itself does not require concepts, but this is not a reason to hold that knowledge gained on the basis of acquaintance is non-conceptual. There is no direct entailment from being acquainted with a phenomenal state to knowing what that state is like. The distinction I have in mind between knowledge by acquaintance and other types of knowledge does not rest on the former's being non-conceptual thing knowledge and the latter's being conceptual fact knowledge. We can accept that all knowledge is conceptual fact

knowledge, and still argue that among all our knowledge, there is a class which deserves to be called knowledge by acquaintance.

Moreover, as I draw the distinction, there is knowledge by acquaintance which is also knowledge of truths. My phenomenal knowledge about the particular pain I had when I hit my thumb is knowing a truth about it. What I know is that that pain had a particular phenomenal quality, which is a truth about it. I know not only a thing, the particular pain; I also know of it that it is so-and-so. This is another aspect of my concept of knowledge by acquaintance that differs from Russell's.

5.3 PHENOMENAL CONCEPTS

A subject's knowledge by acquaintance of his phenomenal states is special in that the justification he provides for it is not mediated by knowledge of other things; and, acquaintance is the epistemic relation a subject bears to his phenomenal states, which is the basis for knowledge by acquaintance. A subject's acquaintance with his phenomenal states explains why his knowledge of his own phenomenal states is unmediated. Now, I will try to provide an account of phenomenal concepts that are the constituents of contents of phenomenal beliefs; and, given that concepts of things are formed on the basis of the type of epistemic relation between the subject and those things, it should not come as a surprise that concepts of phenomenal states formed on the basis of acquaintance with them are expected to have some unique features.

Suppose that I am having a veridical visual experience of a ripe tomato. My present experience instantiates a certain phenomenal quality Q (phenomenal redness), which gives it its

distinctive nature and distinguishes it from other types of experiences, for instance, experiences as of green things. I now attend to that specific quality Q ; and form, on the basis of my attention, a phenomenal concept of it Q , which can be a constituent of my thoughts that I would verbally express by uttering sentences such as “I am having an experience of Q ” and “it is like Q to have this experience”. The question before us is: what is the kind of representative relationship between the phenomenal quality I am acquainted with, which is Q , and the phenomenal concept I form on the basis of my acquaintance, which is Q ? I claim that that I am in a direct epistemic relation with Q enables me to form Q which represents it not relatively (e.g., by means of other properties or the sensory effects it makes on us) but substantively (i.e., as a property in itself).

Before defending the claim, let me introduce some terminology. I will call those concepts which are composed of other concepts “complex”, and those which are not complex “simple”. The concept *the tasteless and colorless liquid which actually fills the lakes* is, for instance, a complex concept which refers to water in this world. Moreover, that one word is used to express a concept does not necessarily show that that concept is simple. Insofar as the concept *knowledge*, which is expressed by ‘knowledge’, is analyzable in terms of other concepts like *justification*, *truth*, and *belief*, it is a complex concept. Similarly, insofar as the concept *cause* is analyzable in terms of concepts *necessary* and *connection*, it is a complex concept. The same also goes for the ordinary concept *water*, which roughly corresponds to the complex concept mentioned above. What is the relation between a given complex concept and the referents of the concepts that compose that complex concept? I will introduce a technical notion to name that relation, and say that a given concept *uses* the referents of the concepts that compose it to represent what it refers to. For instance, the complex concept *water* represents the property of being water by using the properties, not concepts, of being tasteless and being colorless (I

assume that the referents of general concepts are properties). I will also sometimes say that not only a given complex concept but also the subject who possesses it uses the referents of the concepts that constitute that concept to denote to its referent. Composition (or constitution) is a relation among concepts, and using is a relation among a given concept (or a subject) and the referents of other concepts.

My claim above that phenomenal concepts represent their referents as they are in themselves can be explicated in terms of the terminology just introduced. The claim is meant to express the idea that a phenomenal concept is a simple concept that represents a particular phenomenal property without using any property other than that phenomenal property. The phenomenal concept *Q* that I form and deploy when I come to think of the particular phenomenal property (phenomenal redness) *Q* instantiated in my current experience does not use any property other than *Q* to represent it. There is no property other than *Q* in play when I attentively experience and think about it.

Phenomenological evidence for the claim that phenomenal concepts are simple is, I hold, indisputable. I contend that if you now attend to, for instance, the particular phenomenal property instantiated in your current visual experience, you will realize that you do not use any property other than that property to form a concept of it. Every subject who is capable of forming concepts of his own experiences on the basis of attending to them can normally recognize that his relation to those experiences is epistemically intimate enough not to require the use of any properties other than the ones instantiated by them.

Kripke's challenge for physicalists rests on a point that supports the claim for the simplicity of phenomenal concepts. The challenge is how to explain away the appearance of contingency of mind-body identities given that identities are necessary. Why does it appear that

pain might not have been the firing of C-fibers given that as some physicalists claim, it is the firing of C-fibers? In cases of other inter-theoretic reductions such as “water is H₂O” and “heat is mean kinetic energy”, Kripke tells us that we have an explanation for the appearance of contingency. The reason why it appears, say, that heat might not have been mean kinetic energy while it is, is that when we appear to think of the possibility of heat's not being mean kinetic energy, what we really think is a different possibility in the vicinity: the possibility we really have in mind is that the cause of heat-sensation is not mean kinetic energy or something to that effect. The idea is that heat is only contingently the cause of heat-sensation, and what gives the appearance of contingency to the identity claim “heat is mean kinetic energy” is the real contingency of the claim “the cause of heat-sensation is mean kinetic energy”. We confuse the contingent reference-fixer of the concept *heat*, which is the cause of heat-sensation, with its actual reference. Kripke argues that a parallel explanation for the apparent contingency of mind-body identities does not work. There is no real possibility in the vicinity that we mistake for an impossibility, which would have explained away why it appears that pain might not have been the firing of C-fiber, because there is no contingent reference-fixer of the concept *pain*. What we can say is that the reference of our concept *pain* is fixed by our painful sensations; but this is only another way of saying that the reference-fixer of pain is its referent, given that pain is identical to painful sensation.

Kripke's reasoning above is in good company with the claim for the simplicity of phenomenal concepts. The claim that phenomenal concepts are simple in that they represent their referents without using any property other than their referents can be naturally translated into Kripkean terminology as the claim that there are no properties other than their referents in play

which are used to fix the referents of phenomenal concepts. I think the underlying point of Kripke's observation and my claim are essentially the same.

Imagine what happens when our brilliant scientist Mary has her first experience of red, and forms a phenomenal concept of it. Which property does she use to think about phenomenal redness she gets acquainted with? Before her release, she already had some concepts of phenomenal redness like *the quality that is caused by objects with such-and-such reflectance properties*, and *the quality that is nomically correlated with such-and-such brain states*, and etc. These concepts of phenomenal redness are complex concepts which use properties other than phenomenal redness to represent it. However, when she visually experiences phenomenal redness, she is in a position to represent it in a new way. She can directly attend to the quality itself, and use that property to form a representation of it, which she can then deploy in entertaining some new thoughts and raising some new questions about it. The new concept she forms of phenomenal redness is simple. In fact, the entire argumentative force of the knowledge argument rests on the simplicity of phenomenal concepts. If phenomenal concepts were composed of Mary's previously acquired physical/functional concepts or were to use previously referred to physical/functional properties, then she would not come to learn something new when she first experiences red. The widespread intuition that Mary learns something new after her release, irrespective of what it ontologically demonstrates, implicitly assumes that the new phenomenal concept she acquires is simple.

Phenomenal concepts are unique in that they are the only simple concepts we human beings possess. Consider our concepts of “macroscopic” objects and their qualities. My general concept *tomato* that refers to the property of being a tomato, for instance, involves certain visual, tactile, and gustatory experiences I had before and kept in my memory; and by making use of

those experiences, I recognize that what I now see or touch is a tomato or falls under my concept *tomato*. Or, similarly, the singular concept I have of this particular tomato I am now holding involve my present experiences that are caused by it. I apply my singular concept in question to this particular object on subsequent occasions in virtue of having certain kinds of experiences that are given rise to by having interactions with it. If the particular point here sounds reasonable, it holds for all concepts we have of macroscopic objects (chairs and tables, cats, and rocks) and their qualities. Concepts of what is sometimes referred as imperceptibles or theoretical entities (molecules, electrons, and so on) are no exception to this. Their complexity is a matter of being constituted of concepts of “macroscopic” scientific instruments by which we detect their presence and hypothesize about their nature. And this general thesis about the complexity of concepts of external things, macroscopic or not, is to be expected, given that concepts are a species of mental representations we form on the basis of the kind of epistemic access we have to things in the world and the kind of epistemic access we have to external objects and their qualities is mediated by our phenomenal experiences. I would like to state it as plainly as possible: the way we represent the outer world depends on the kinds of experiences that are triggered as a result of interacting with it. Change the relevant experiences, you thereby change the corresponding concepts. To use a famous example of Nagel’s, if bats had concepts of the external world, we can safely assume that those concepts would be very different from the ones we now have given the drastic differences between our and their experiences. A conceptually-sophisticated bat’s concepts of external objects would be inescapably affected by the kind of experiences that are given rise to by its echo-location sensory apparatus, which we human beings lack. My concept and the bat’s would-be concept can be of the same thing, but they differ in how they represent it. The general point is that concepts of external things are always complex

because they are mental representations which use phenomenal experiences to represent those things.

One may think that the idea of a concept's using only its referent to represent it is mysterious because it may seem that the idea ignores the distinction between what is represented (a quality) and what does the representing (a concept): if a simple concept uses a quality to represent that quality, and the property a simple concept uses individuates it, then how can we possibly draw the distinction between a simple concept and the quality it represents? I do not wish to deny the distinction between a concept and a quality. I do not deny it, and do not have to deny it to hold that phenomenal concepts are simple. There are two points I want to make in response. First, that something is mysterious does not necessarily count against its existence. Whatever mysteriousness the idea of simple concepts is supposed to have, I think that the idea of the simplicity of phenomenal concepts is well supported by phenomenological and theoretical evidence. Second, the mystery at hand is mostly merely apparent. We already use phenomenal properties to represent things other than those properties, so why cannot we use them for the job of representing themselves? My concept of my dentist (e.g., *the man who caused me that pain*) uses the particular pain I had in my teeth to represent him, so why cannot we use that pain for the job of representing itself? In fact, it would be mysterious if we could not do it. This does not commit us to denying the distinction between a simple concept and the quality it represents. Consider Mark again. He feels pain, and his experience instantiates that hurty, nasty phenomenal quality; but this is consistent with his not having a phenomenal concept of that quality. Generally put, the idea that the only properties simple concepts use to represent their referents are their referents themselves is consistent with simple concepts' being different from their referents given that the conditions of acquaintance with a phenomenal quality are different from the

conditions of acquisition of a simple concept on the basis of acquaintance with it. A conceptually-underdeveloped human being can satisfy the former without satisfying the latter.

5.4 THE CONCEPTUAL/NON-CONCEPTUAL DISTINCTION(S)

Taking phenomenal concepts as simple has certain implications about the conceptual/non-conceptual distinction. There are basically two ways in which the distinction can be drawn, and as Heck (2000) correctly observes, the two have not always been sufficiently distinguished. Firstly, the distinction can be drawn in terms of mental states. It can be said that a mental state is non-conceptual if and only if being in that state does not require the possession and application of any concepts. Being in the phenomenal state of visually experiencing red, for instance, does not require the exercise of any concepts: one can experience red without exercising a concept of redness. Believing that one experiences red, on the other hand, is generally thought to be a mental state that requires the possession and application of certain concepts: one cannot believe that one experiences red without having a concept of redness. A baby can experience red but, lacking a concept of redness, he cannot believe that he experiences red. The reality of the distinction between conceptual and non-conceptual mental states can, I believe, be hardly a matter of serious dispute. And, secondly, the distinction can be drawn in terms of mental contents. Phenomenal states that we are in when we perceive external objects have contents in that they represent the world in certain ways. The visual experience that I have when I look through my window, for instance, has a certain representational content. I visually experience certain objects as having certain features: some appear to be red, some appear to be round, and

etc. The question is now whether the contents of my perceptual phenomenal states can be fully captured by concepts. Moreover, notice that a state's being non-conceptual does not by itself show that its content is non-conceptual. A conceptually-underdeveloped subject may visually experience a specific shade of red, but this does not show that there is no concept that has exactly the content of that experience: it may be the case that while there is such a concept, our subject simply lacks it.

Distinguishing mental contents in this way has been a matter of serious dispute in recent philosophy. The question is whether there are non-conceptual contents thus defined. A famous argument for the presence of non-conceptual contents which is, in its genetic form, propounded in Evans' *The Varieties of Reference* is from "the fineness of grain" of perceptual experiences, which I will call "the Richness Argument". Chris Peacocke presents a version of the argument:

If you are looking at a range of mountains, it may be correct to say that you see some of them as rounded, some as jagged. But the content of your visual experience in respect of the shape of the mountains is far more specific than that description indicates. The description involving the concepts round and jagged would cover many different fine-grained contents which your experience could have, contents which are discriminably different from one another. (1992, p. 111)

Peacocke's point is that the concepts like *round* or *jagged*, which I would use to give a description of my present experience of the shape of mountains, fail to exactly capture the content of that experience because those concepts can also be used to describe (or to "cover") some other experiences which are "discriminably different" from it. Consider the contents of my experiences of the shapes of the Ozarks Mountains in Missouri and the Agri Mountain (also known as Mount Ararat) in Turkey. Peacocke's point is that I will use the same concepts to

describe the relevant contents of my experiences even though they differ in their specificity. I will say, for instance, that both shapes are round while the roundedness of the Agri Mountain as it is presented in my experience is perceptually distinguishable from the roundedness of the Ozarks Mountains as they are presented in my experience. The concept *round* fails to characterize the relevant contents in such a way as to describe their differences in full detail, and the non-conceptualist thesis is that this is only indicative of a more general failure: there are no concepts which can indeed fully capture the richness of content (or fineness of specificity) that is presented in my perceptual experiences of the relevant shapes.

I argue that the richness of contents of perceptual experiences does not support the non-conceptualist thesis. The falsity of the non-conceptualist thesis directly follows from the simplicity of phenomenal concepts. Phenomenal concepts use the qualities they denote to represent them, and hence they are as fine-grained (or rich in detail) as those qualities themselves. This means that the contents of our perceptual experiences can be captured in their exact specificity by means of the phenomenal concepts we form on the basis of attentively experiencing them. I think the gist of McDowell's idea is right when he argues against the non-conceptualist thesis in the following way:

In the throes of an experience of the kind that putatively transcends one's conceptual powers - an experience that *ex hypothesi* affords a suitable sample - one can give linguistic expression to a concept that is exactly as fine-grained as the experience, by uttering a phrase like 'that shade,' in which the demonstrative exploits the presence of the sample. (1996, pp. 56-7)

The content of a given experience can be fully accommodated within “the space of concepts” by exploiting it to form a concept of it. Concepts thus formed are as determinate in content as the qualities they represent.

The thesis that experiences have contents the having of which does not require the exercise of any concepts (call this thesis *A*) is to be distinguished from the non-conceptualist thesis that experiences have contents which cannot be fully conceptualized (call this thesis *B*). I defend the former but argue against the latter. A little child can have experiences with certain representational contents, which he can not even begin to characterize in adequate terms. But this does not mean that the contents of those experiences cannot be conceptualized and subsequently be a part of a belief about that experience. The non-conceptualist wishes to move from the plausible *A* to the contentious *B*. But from the mere fact that the subject of a given experience does not have the concepts required for a full characterization of the content of that experience, it does not follow that there cannot be any such concepts. And, indeed, if there are phenomenal concepts, and if they are simple, then *B* is false. Though I do not think McDowell would agree with the account of phenomenal concepts that I attempt to develop here (I suspect that he would say that placing the idea of acquaintance at the roots of an account of phenomenal concepts carries the traces of the infamous Myth of the Given - more on this later), the point we want to make against the non-conceptualist is in effect the same: the content of a perceptual experience can, so to speak, be taken up as the content of a concept.

However, McDowell argues not only against *B* but also against *A*. Here, I side with the non-conceptualist who thinks that *A* is true. I will make three points about rejecting *A*. First, just as the truth of *A* does not imply the truth of *B*, the falsity of *B* does not imply the falsity of *A*. So, pointing at the falsity of *B* is not enough to show the falsity of *A*. We need a separate

argument. Second, the idea that the content of an experience can be conceptualized by “exploiting the presence of the sample [experience]” (McDowell) appears to assume that the content of an experience is prior to conceptualization: first, we have the content; and then, we have the conceptualization (or “exploiting”). And, if the content of an experience is prior to conceptualization, and if conceptualization is, by definition, a process which transforms the non-conceptual to the conceptual, then A is true. Conversely, if A were false, that is, if the contents of experiences required the exercise of concepts in order to be enjoyed, then the whole idea of conceptualization would evaporate simply because those contents would already be conceptual. Fodor makes a similar point when he writes:

It can't be both that you learn the concept RED from your impressions [or experiences - E.D.] of red, and that you need to have that concept in order to have such impressions. The form of the argument appears to be perfectly general; if you can learn concept X from your impressions of Xs, then it must be possible to have impressions of Xs without having the concept of X. (2005, pp. 44-5)

If acquiring a certain concept is a result of having a certain experience, then having that concept cannot be required for having that experience. In other words, the falsity of B, which McDowell rightly argues for, implies the truth of A, which he wants to argue against.

Thirdly, let us consider McDowell's argument for the falsity of A. His idea is that the only plausible way to account for the justificatory relations between perceptual experiences and perceptual beliefs is to hold that perceptual experiences present us with conceptual contents. Briefly, the argument goes as follows. It starts with a premise which is almost universally accepted: perceptual experiences provide reasons for forming certain beliefs. Seeing an X does not just cause one to believe that there is an X, but it is one's reason for believing that there is.

The relations between the contents of experiences and the contents of corresponding perceptual beliefs are rational or “reason-constituting” (McDowell). But, and here comes McDowell's original contention, the relations between the contents of experiences and the contents of beliefs can be reason-constituting only if those contents are of the same semantic kind: if the contents of experiences are of a kind different from those of beliefs, then how can the former be reasons for the latter? McDowell's response is that they simply cannot: only if we take perceptual experiences to have belief-like contents, we can make sense of the reason-constituting relations in question. He writes:

[W]e can coherently credit experiences with rational relations to judgment and belief, but only if we take it that spontaneity is already implicated in receptivity; that is, only if we take it that experiences have conceptual content. (1996, p. 162)

The argument here can be stated thus:

1. Experiences provide reasons for perceptual beliefs.
2. Experiences provide reasons for perceptual beliefs only if they have conceptual content.

Hence,

3. Experiences have conceptual content.

Now, I assume that (1) is, as some say, “not negotiable.” In any case, the original idea McDowell puts forth is stated by (2), which I will focus on. What is the relation between an experience's providing a reason for a perceptual belief and its having a certain kind of content? What exactly makes non-conceptual contents putatively disqualified from being reasons for conceptual contents?

Before proceeding further, let me say a few things as to how McDowell understands the notion of providing (or being) a reason. There is a sense in which that it is raining is a reason for taking an umbrella before leaving the apartment, but it may not be the subject's own reason if the subject is not aware that it is raining. McDowell is clear that his talk of reasons should be understood from the point of view of the subject in question but not from a third-person perspective which can be adopted, say, by a theorist trying to account for the behavior of the subject under consideration. Reasons are what are recognized or recognizable as such by the subject. The procedure for bringing out the subject's own reasons is roughly this. Take a subject who holds certain beliefs, for instance, the belief that Sully is sad; ask him what makes him believe that Sully is sad, the (sincere) answers he gives ("because she is crying") state his own reasons for his beliefs: that Sully is crying is the subject's reason for believing that she is sad. McDowell writes:

[S]uppose one asks an ordinary subject why she holds some observational belief, say that an object within her field of view is square. An unsurprising reply might be "Because it looks that way". That is easily recognized as giving a reason for holding the belief... [S]he gives expression to it in discourse (1996, p. 165).

The idea here is that reasons must be, as McDowell puts it, "articulable" by the subject. This does not of course mean that reasons for a certain belief must already be explicitly articulated before the subject comes to hold that belief. The subject may believe that the object before her is square without thinking or saying to himself or others that it looks square. The point is that the

reasons she has for her beliefs must be articulable at the moment of, what Sellars calls, “the game of asking and giving reasons”, e.g., when asked why she believes that it is square.⁴⁶

What argument does McDowell present for the premise (2)? What exactly is it, on McDowell's view, that bars non-conceptual contents from being the subject's own reasons for his perceptual beliefs? McDowell writes:

The routine point is really no more than that there can be rational relations between its being the case that P and its being the case that Q (in a limiting case what replaces “Q” can simply be what replaces “P”). It does not follow that something whose content is given by the fact that it has correctness condition that P can *eo ipso* be someone's reason for, say, judging that Q, independently of whether the content is conceptual or not. We can bring into view the rational relations between the contents...only by comprehending the putatively grounding content in conceptual terms, even if our theory is that the item that has that content does not do its representing in a conceptual way. A [non-conceptualist] theory like Peacocke's does not credit ordinary subjects with this comprehensive view of the two contents, and I think that leaves it unintelligible how an item with the non-conceptual content that P can be someone's reason for judging that Q. (1994, p. 166, note omitted)

⁴⁶ Notice that the requirement that reasons must be articulable by the subject is stronger than the requirement that reasons must be recognizable as such by the subject. It is arguably conceivable that our subject can recognize his reasons as such without having the linguistic ability of “giving expression to [them] in discourse”. The two requirements are necessarily co-extensional only if there is a necessary connection between the linguistic ability and the recognitional capacity; and, an argument for the putative necessary connection appears to require some substantial, and controversial, theses about the relations between thought and language or, as McDowell says, the relations “between reason and discourse” (p. 165). I will bypass such considerations here and go with the weaker recognition requirement. First, nothing in McDowell's argument above essentially hangs on them. Second, McDowell himself sets the standards for the articulation requirement low enough when he says, “I do not mean to suggest any special degree of articulateness” (p. 165); and hence, there is a sense in which in the present context, he can be plausibly taken as having only the recognition requirement in mind.

McDowell makes several points in this passage. First, he makes “the routine point” that rational relations are between contents (e.g., its being the case that P and its being the case that Q). Second, he makes it clear that he does not intend to deny that there are non-conceptual contents. Our theory may be that “the item that has [some] content does not do its representing in a conceptual way”. Third, he argues that we can bring into view the rational relations between the contents only if we “comprehend” those contents in conceptual terms. There may be rational relations between some contents including non-conceptual ones, but those rational relations cannot be the subject’s ground for holding some content as a reason for some other content unless he comprehends those contents in conceptual terms.

McDowell’s concern at this and some other passages is not properly appreciated by some philosophers. For instance, Fodor writes:

I’m aware, of course, that some philosophers hold that only a belief can warrant a belief; a fortiori, that no impression can. I guess the intended argument is that only something with content can bestow warrant, and (rhetorically) what except a propositional attitude could have content? Well, it appears that impressions do, so why shouldn’t beliefs be warranted by impressions? (2005, p. 88)

To the (rhetorical) question Fodor asks at the end of this passage, McDowell has a clear reply. McDowell can admit that impressions are among those items with content “that [do] not do [their] representing in a conceptual way.” But he will ask how impressions with non-conceptual content can be *reasons for the subject* for holding some beliefs. How do impressions construed as having non-conceptual content serve as reasons for the subject unless the subject conceptualizes them or, as McDowell puts it, “comprehends [them] in conceptual terms”? Having some content is a necessary condition for being a reason for the subject, but it is not

sufficient. There may be rational relations among items with contents that are not recognized as such by the subject. Only when the subject conceptualizes those contents, he can recognize those relations as rational relations, and only after that, he can be in a position to present them as *his* reasons for holding a belief.

McDowell is right to insist that contents can be counted as reasons for the subject only if they are comprehended by him. This is a very plausible point. Moreover, it strongly appears that concepts are necessary for comprehension. I can comprehend the thought that the sun shines in summer only if I have the concepts *sun*, *shine*, and *summer*. Thoughts are constructed out of concepts, and they can be comprehended only when the subject possesses the relevant concepts. The first point is about the connection between being a reason and comprehension, and the second point is about the connection between comprehension and having concepts. Taken together, they imply that contents that are not conceptualized by the subject cannot be reasons for him because they cannot be comprehended by him.

I agree with McDowell that items with non-conceptual contents cannot be reasons for the subject unless they are conceptualized by him. Call this thesis *C*. But, put in this way, this thought is not a point against the thesis *A* spelled out above, the thesis that experiences have contents the having of which does not require the exercise of any concepts (simply put, the thesis that experiences have non-conceptual contents). There is no inconsistency in holding both *A* and *C*. Experiences may have non-conceptual contents, *and* they may be reasons for the subject for holding certain beliefs only if their contents are conceptualized by the subject.

What about the argument McDowell proposes for the thesis that experiences have conceptual contents? Reflection shows that the second premise (that is, experiences provide reasons for perceptual beliefs only if they have conceptual contents) is not supported by the

connections McDowell points out among being a reason, comprehension, and conceptualizing. What those connections support is something considerably weaker, namely, that experiences provide reasons for perceptual beliefs only if their contents are conceptualized by the subject. In the absence of any further support, it is reasonable to conclude that while McDowell manages to cast serious doubts on the idea that non-*conceptualized* contents can be reasons for perceptual beliefs, his reasoning falls short of warranting the second premise of his argument against the thesis that experiences have contents the having of which does not require the possession of any concepts or, in McDowell's terminology, "the passive exercise of conceptual capacities".

5.5 THE "MYTH OF THE GIVEN"

Is the account of phenomenal concepts that I develop here "one of the forms" of what Sellars famously called the "Myth of the Given"? It is hard to tell. This is partly because Sellars never explicitly specifies what the Myth of the Given is. Rather, what the reader finds in his seminal work "Empiricism and the Philosophy of Mind" (1956) is a shift of the targets, ranging from sense-data theories to a certain form of epistemic foundationalism to logical/semantic atomism. It is clear, however, that there are no apparent necessary entailments between these views (e.g., one can consistently be a logical atomist without being an epistemic foundationalist). Sellars' famous assault on "the entire framework of givenness" contains some suggestive ideas but the broad range of the targets renders a focused and sustained discussion difficult.

Let me begin by noting that it is clear that my account of phenomenal concepts does not presuppose or entail anything like sense-data. Sense-data are traditionally conceived as mental

particulars with which subjects are supposed to be in a direct epistemic contact. In this work, my focus has been on phenomenal *properties* and our concepts of those properties. A commitment to phenomenal properties, and in general to phenomenal realism, does not entail a commitment to sense-data. This is because one can hold that phenomenal properties are instantiated by states or *experiencings* instead of phenomenal particulars. Moreover, even if one accepts the existence of phenomenal particulars, one might still reject a sense-data model of perception, on which one perceives the world by virtue of perceiving those particulars.

Moreover, my account of phenomenal concepts does not fall prey to Sellars' "inconsistent triad" objection. Sellars writes:

It is clear from the above analysis, therefore, that classical sense-datum theories ... are confronted by an inconsistent triad made up of the following three propositions:

A. *x senses red sense content s entails x non-inferentially knows that s is red.*

B. The ability to sense sense contents is unacquired.

C. The ability to know facts of the form *x is phi* is acquired.

A and B together entail not-C; B and C entail not-A; A and C entail not B. (Sellars 1956, section 6)

Surely, Sellars is right to think that this is an inconsistent triad. This "inconsistency" objection is originally raised against sense-datum theories, but it is easy to see that it applies to a wider class of views because nothing in the formulation of theses A-C assumes the existence of sense-data or something like indirect realism. Thesis A assumes that sensing is a relation that connects subjects to certain contents; but, literally taken, it does not take for granted a certain conception of what those contents are. The inconsistency objection would hold even if those contents were not mental but physical.

The most problematic thesis in Sellars' triad is A, and my account is well-suited to reject it. I agree that sensing (or, as I call it, acquaintance) is a relation that connects the subject to certain contents but I deny that this relation automatically enables one to non-inferentially know those contents. There is a distinction between acquaintance and knowledge by acquaintance, and there is no direct entailment from the former to the latter. Infants may be acquainted with some of their phenomenal states and the qualities instantiated by those states without knowing that those states have those qualities. It is true that on my account, knowledge by acquaintance is non-inferential in the sense that its justification does not come from having *other* beliefs but from being acquainted with the experience itself. But its being non-inferential does not mean that just by being acquainted with something the subject knows what he is acquainted with. In order for a subject to know the contents of his acquaintance, he should *conceptualize* those contents. And, the possibility of the failure of the satisfaction of this condition blocks an entailment from acquaintance to knowledge by acquaintance.

The thesis I accept instead of A is A*:

A*. That x senses red sense content s puts x in a position to non-inferentially know that s is red.

Clearly, the triad A*-B-C is not inconsistent.

In considering the possibility of rejecting A, Sellars writes:

He [the sense-datum theorist – ED] can abandon A, in which case the sensing of sense contents becomes a non-cognitive fact – a non-cognitive fact, to be sure which may be a necessary condition, even a logically necessary condition, of non-inferential knowledge, but a fact, nevertheless, which cannot constitute knowledge. (1956, section 5)

Sellars claims that abandoning A implies that sensing is a non-cognitive fact because he appears to assume that “constituting knowledge” is the only form of being a cognitive fact. If there is no

entailment from sensing to knowing, then, Sellars thinks, sensing is non-cognitive. But why does Sellars think so? It is possible that a relation can be cognitive without guaranteeing knowledge. A subject may hold certain phenomenal beliefs; and, when it comes to providing *justifications* for those beliefs, his sensings of the contents of those beliefs may, and indeed do, play an indispensable role even though they do not entail the knowledge of those contents. Sensing may do some crucial justificatory work, and this appears to be a good enough reason to call it “cognitive”.

Brandom gives the following definition of the Myth of the Given:

The Myth of the Given is the claim that there is some kind of experience the having of which does not presuppose grasp of concepts, such that merely *having* the experience counts as *knowing* something, or can serve as *evidence* for beliefs, judgments, claims, and so on. (Brandom, “No Experience Necessary” p. 3)

The Myth of the Given is, for Brandom, then, that non-conceptual experiences play some cognitive role, either by entailing knowledge or by providing evidence for beliefs. Brandom’s definition of the Myth is thus wider than the definition that is suggested by Sellars’ inconsistent triad, which roughly corresponds to the acceptance of A. If we take Brandom’s definition of the Myth of the Given, then my account is indeed a form of the Myth of the Given. But, then, it is hard for me to see what is “mythical” about the Myth. Having experiences does not by itself count as knowing – this is something we can all agree. But what is wrong with the idea that having non-conceptual experiences can serve as evidence for beliefs? I have the experience of intense pain, and having this experience is non-conceptual in that its existence does not presuppose grasp of concepts. I also have the belief that I have the experience of intense pain. Now, what is wrong with the idea that my belief is justified or warranted by my experience of pain?

The main worry Brandom has in mind is, if I get it right, how something with non-conceptual content (experience) can be evidence for something with conceptual content (belief): evidential relations are between the same kinds of contents; and, if non-conceptual and conceptual contents are of two different kinds, how can the former be evidentially related to the latter? But, when I say “non-conceptual experience”, what I mean is merely “experience the having of which does not presuppose grasp of concepts”, and that does *not* imply any difference in kind between the contents of experiences and the contents of beliefs. In fact, the idea that the contents of experiences are *fully conceptualizable*, an idea that I and McDowell share, implies the falsity of the thesis that the contents of experiences are different in kind from the contents of beliefs, which are constituted by the contents of corresponding concepts. One of my claims above was that the contents of phenomenal beliefs or concepts are derived or inherited from the contents of experiences. And, if the contents of certain beliefs are not different in kind from the contents of corresponding experiences, then it is hard to see why there cannot be evidential relations between experiences and phenomenal beliefs.

A distinction needs to be drawn:

1. A given content is non-conceptual only if it cannot be the content of a concept.
2. A given content is non-conceptual only if having a state with that content does not require grasp of concepts.

On the one hand, if 1 is what we have in mind, then it is true that something with non-conceptual content cannot serve as evidence for something with conceptual content. But, in this case, I do not think that experiences have non-conceptual content. On the other, if 2 is what we have in mind, then experiences have non-conceptual content. But, then, experiences’ having non-conceptual content is not incompatible with their serving as evidence for something with

conceptual content because 2 does not imply that the contents of experiences and beliefs cannot be identical.

5.6 THE INFALLIBILITY OF PHENOMENAL BELIEF

Suppose I am having an experience with a certain phenomenal quality, *Q*. I introspectively attend to the quality exemplified by my experience and form the concept *Q*, the content of which is the quality *Q* itself. Moreover, I deploy the concept *Q* by predicating it of the very experience the introspection of which results in my forming the concept, and accordingly come to possess the belief that this (experience) is *Q*. What is the epistemic status of this belief?

It is, or should be, clear that the belief in question is infallible in the sense that it cannot be false. The alleged belief attributes to my experience a certain property the instantiation of which by that experience is the very ground for the formation and my possession of that belief. The property my belief predicates of my experience is the same property instantiated by that experience to which I attend to form the corresponding concept of my belief. In order to *have* the belief, I *must* have the experience instantiating the alleged property in the first place. The truth of my belief that this (experience) is *Q* is guaranteed by the very way through which it is formed.

The infallibility of phenomenal beliefs directly follows from the account of phenomenal concepts I propose above. Since I think that my account is, at least with respect to its basic claims, correct, I also think that the infallibility of phenomenal beliefs, which is a consequence of the basic claims of my account, is correct. I will not here deal with the alleged counter-examples

present in the literature to the infallibility thesis, but only point out that *if* my account is correct, there *must* be something wrong with those counter-examples.

It is important to note that the infallibility of phenomenal beliefs does not imply that subjects that hold those beliefs think that they are infallible or even true. Obviously, subjects may be confused about the epistemic features of their beliefs, and there is no reason that precludes a proponent of the infallibility thesis from endorsing this trivial point. Moreover, the infallibility thesis does not imply that subjects cannot fail to have phenomenal beliefs about their phenomenal states. A proponent of the infallibility thesis can consistently grant that subjects do not have phenomenal beliefs about all of their phenomenal states simply because the conditions for having phenomenal states are different from the conditions for having beliefs about those states: the satisfaction of the former does not guarantee that of the latter.

Now let us combine the infallibility thesis with the “justification” thesis I explored in the previous sections that having experiences justifies beliefs about having experiences. These two theses imply that the mere possession of phenomenal beliefs guarantees the possession of phenomenal knowledge. If a subject has a phenomenal belief to the effect that he is having a certain type of experience, then the infallibility thesis implies that his belief is true (that he *is* having that experience). And, the justification thesis implies that his having that experience justifies his belief. So, from the belief we get the truth; and from the truth, we get the justification. Here we have a beautiful harmony of truth and justification, which is absent in the case of our beliefs about external objects. Any belief about an external object, perceptual or not, can be justified and false, or be true and unjustified. This is because what makes a belief about an external object justified (the justifier) is different from what makes it true (the truth-maker). I may now be as much justified as conceivably possible in believing that there is a computer on

the table I am now seeing; and, despite this justification, my belief may still be false. However, in the case of phenomenal beliefs, truth-makers and justifiers are the very same thing. Having this very experience is both the truth-maker and the justifier of my belief to the effect that I am having it.

The “implication” thesis, i.e., that the mere possession of phenomenal beliefs implies phenomenal knowledge, needs to be qualified. There may be a subject with a certain phenomenal belief who is so confused that when asked, he fails to quote his experience as evidence for his belief. This subject’s epistemic access to his own experience is the ground that makes the formation of his phenomenal belief possible; but, he may fail to appreciate the justificatory relation between his experience and his belief. This possibility proves the falsity of the implication thesis. But, still, something very like the implication thesis is true. The character of the subject’s access to his own experiences enables him to “notice” the truth of his belief; and, this act of noticing the truth of the belief provides the subject with a very (indeed, maximally) good justification for holding it. What renders a given justification for a certain belief *epistemic* is its being truth-conducive, that is, its being or providing some evidence to the effect that the belief is *true*. And, if the subject can notice the truth of a certain belief, then he has all he can possibly ask for to justify his belief. A point I stressed before is that the character of the subject’s access to external objects does not enable him to notice the truth of his beliefs about those objects. If it had enabled him, then he would not need to fall back to his beliefs about his experiences to justify his beliefs about external objects but he would be in a position to quote the truth of his external beliefs as a justification for them.

What precludes the mere possession of a phenomenal belief from implying phenomenal knowledge is the possibility that the subject may fail to notice the truth of his belief even though

the belief is infallible. But we can reasonably claim that noticing the truth of a phenomenal belief is a minimal rationality requirement for a subject that holds it: a minimally rational subject cannot have a phenomenal belief without noticing its truth. How can a minimally rational subject fail to notice the truth of his phenomenal belief when what is required for noticing the truth of his belief is the same as what is required for the formation of his phenomenal belief, that is, his attending to his own experience? We can thus attain a qualified version of the implication thesis: the mere possession of a phenomenal belief by a rational subject implies that the subject has phenomenal knowledge.

6.0 AN ARGUMENT FOR PROPERTY DUALISM

In this chapter, I assume two theses. First, phenomenal concepts are cognitively independent from all physical/functional concepts in the sense that there are not any a priori relations among phenomenal and physical/functional concepts. A concept *X* a priori implies a concept *Y* if a rational being that possesses both concepts is in a position to know that whatever *X* is *Y* solely on the basis of possessing those concepts. The concept *kill* a priori implies the concept *die*, the concept *knowledge* a priori implies the concept *belief*, and the concept *human* a priori implies the concept *animal*. A rational person who has one of these pairs of concepts does not need to make any empirical investigation to know that whatever the first concept applies to falls under the second concept. Under this light, the first thesis amounts to this: a rational being that possesses a phenomenal concept and all physical/functional concepts is not in a position to know, merely on the basis of possessing them, that whatever that phenomenal concept applies to falls under any of those physical/functional concepts.

Here are a number of amplificatory remarks about the first thesis. 1) It does not claim that if a concept *X* implies another concept *Y* then a necessary condition for the possession of *X* is the temporally prior possession of *Y*. For all it claims, it may be possible to possess *X* without possessing *Y*. It is a claim that assumes the *joint* possession of certain concepts but not a claim about the necessary conditions for their possession. 2) It rests on the idea of a priori relations

between pairs of concepts and hence on the distinction between a priori and a posteriori truths. The thesis takes for granted that there are some pairs of concepts the mere possession of which enables a rational subject to know certain truths about what fall under them. If there are no a priori relations between any concepts whatsoever, as some Quineians claim, then the first thesis is empty. It is significant only if what it claims is understood as the idea that there *are* a priori relations among physical/functional concepts and there are *not* any a priori relations among phenomenal and physical/functional concepts. 3) How are we to understand “a rational being”? There is a problem here. A necessary condition for rationality is avoiding inconsistencies. A rational being should reject a statement like “some pieces of knowledge are not beliefs” on grounds of its inconsistency. But the inconsistency of this statement is due to the fact that *knowledge* implies *belief*. The apparent problem is circularity: a sufficient condition for the implication relation among concepts is originally given in terms of a rational being, and rationality in turn is defined in terms of the implication relation. It is not very clear how serious this problem of circularity is or whether it is really a problem at all. The notions of consistency/implication and rationality may belong to a “logical space” in which each term is defined in terms of others. An understanding of any of the notions that belong to that logical space may be holistic in the sense that it is not possible to have that understanding without having understood a good deal of other notions in that space. I am not very sympathetic to holistic theories of meaning, but I do not think that they are obviously wrong.

The second thesis I will assume is this: the mode of presentation of a phenomenal concept is the phenomenal property to which it refers. This idea is, as you will remember, endorsed by the thick account of PCS discussed before. Suppose that you are having a certain type of pain experience. It does not go unnoticed but captures your introspective attention. You

focus on the particular phenomenal quality it has, and form a thought about it. How is this quality represented in your thought? The second thesis says that the quality itself is a part of your thought. No mode of representation that is different from the quality you attend to enters into your thought as a constituent.

It is instructive to contrast phenomenal concepts with perceptual concepts of ordinary physical objects. There is a table in my room. I carefully look at it from a certain angle under certain lighting conditions, and form a visual concept. Name that concept 'C'. An important point is that it is the concept *C* but not the table itself that enters into my thoughts about the table. There are a number of arguments that support this point (this is of course what leads Frege to introduce the notion of sense in his famous essay "On Sense and Meaning"), and here I prefer the one which seems to be the simplest: there is only one table but there is a multiplicity of thoughts, which are not a priori related to each other, I can form of the table. The thought I would express by using *C* would be different from the thought I would express by using another concept *C'* of the table I form from another angle under different conditions. I may not know that these two concepts are of the same table, and hence that *C* is 'C' may be informative for me. This means that the uniqueness of the table cannot account for the multiplicity of my thoughts and the informativeness of some thoughts about it. We need *representations* of the table to do justice to the extent of my thoughts. The table is not the same thing as its representation, and the latter is a constituent of my thought. Now consider my phenomenal concepts. I feel this pain in my arm now, attend to its distinctive quality, and form a thought about it. How is this quality represented in my thought? Obviously, there are no angles of view which I can adopt when I attend to my pain – at least not in the sense in which I can adopt when I attend to the table. There is no

multiplicity of representations with respect to phenomenal states in the sense we have with respect to external objects.

In the fourth chapter, I have argued that these two theses (the conceptual independence of phenomenal concepts and the constitution of phenomenal concepts by their referents) are inconsistent with physicalism. This is clear once we ask the question what phenomenal concepts refer to. If physicalism is true, then they refer to physical properties. But, given the constitution thesis, this means that phenomenal concepts are constituted by physical properties. However, this implies the falsity of the thesis of conceptual independence. Hence, the truth of the two theses at hand implies the falsity of physicalism. Physicalism tells us that there are no non-physical properties. Its falsity tells us that there are non-physical properties. This is a simple argument for property dualism. In this chapter, I will try to make clear why it works in more detail by providing an assessment of the physicalist strategy that attempts to incorporate the two theses in question.

6.1 ASCENDING LEVELS

Phenomenal concepts are employed in our introspective thoughts about the qualitative characters of our phenomenal states. There is something it is like to be in a certain phenomenal state, let us grant this and put it aside for now. Let us instead try to understand the state one is in when one thinks about the phenomenal state she is actually in. Consider a subject feeling pain that thinks about the feeling he is actually having. In his thought about the feeling, he employs a phenomenal concept of the feeling. A question that is not sufficiently addressed in the literature

is the ontological status of this sort of higher level state. The subject employs a phenomenal concept, which is granted to be conceptually independent of all physical/functional concepts. In order for physicalism to succeed, it should account for this higher level state as well as the lower level state of feeling. That is, a criterion of adequacy for physicalism is that it tells us how there can be a fact such as the following in an entirely physical world:

A. Susan thinks that the state she is in is *P*.

Susan thinks about the phenomenal state she is in by employing a phenomenal concept *P*. This is a fact about the actual world, and the statement A states this fact. What can physicalism say about this?

Now, let us assume, for the sake of the argument, that there is a physical concept *P'* that is co-referential with the phenomenal concept *P*. Physicalism can claim that the independence of *P* from physical concepts is not a problem if *P* refers to a physical property. The strategy here is one of ascending levels. At first instance, it seems that phenomenal properties are different from physical properties. The physicalist of a certain sort tells us that this appearance is misleading: phenomenal properties are indeed physical properties and the reason why it seems otherwise is that the concepts by which we grasp these properties are fundamentally different from ordinary physical concepts under which we grasp ordinary physical properties. But what can this strategy say about those facts in which phenomenal concepts figure? Take the fact that Susan is thinking about *P* and the fact that Susan is thinking about *P'*. These two facts appear to be different facts. For one thing, one may occur without the other. How can there be a fact like Susan's thinking about *P* in an entirely physical world? The physicalist can ascend levels with respect to the problem of phenomenal properties but he does not address but only displaces the problem if he does not provide an answer to this question.

The point I wish to emphasize is this. Susan's thinking about *P* is *a fact of our world*. This is a different fact from Susan's feeling *P*. And, what works for the latter need not necessarily work for the former. To the question "how can Susan's feeling *P* and Susan's being in such-and-such brain state be identical?" the physicalist can say "they are the same but they do not appear to be so because the concepts *feeling P* and *being in such-and-such brain state* are cognitively independent". But to the question "how can Susan's thinking about *P* and Susan's thinking about *P'* be the same?" the physicalist does not seem to be in a position to provide an analogous reply. These are different facts. But the former fact is one in which a phenomenal concept figures. It is a fact that cannot be captured by a physical concept if phenomenal concepts are conceptually independent from all physical concepts. Susan's thinking about *P* is different from all the facts that we can have by replacing *P* by any physical concept if the conceptual independence claim holds. But then Susan's thinking about *P* is not a physical fact because all physical facts can presumably be captured by physical concepts.

It is important to note here that what I say above does not trade on a sort of ambiguity of what "a fact" is. That Phosphorus is bright and that Hesperus is bright are two different facts if we take "fact" in a fine-grained sense but they are the same fact if we take it in a coarse-grained ("states-of-affairs") sense. It is the latter sense in which I claim that Susan's thinking about *P* and Susan's thinking about *P'* are different facts, and this is the sense which matters for the debate on physicalism. They are different facts even if *P* and *P'* refer to the same property. This is because they are facts in which the modes of presentation associated with *P* and *P'*, which we assume to be different, figure. Thinking about Plato as the brightest student of Socrates and thinking about him as the teacher of Aristotle are different facts, in a coarse grained sense of the term, even if the brightest student of Socrates is the teacher of Aristotle. In this coarse-grained, states-of-

affairs sense of the term, Susan's thinking about P is a fact that is different from all the facts that we can have by replacing P with physical concepts. It is a further fact, an addition. The physicalist can claim that the fact that Susan has P and the fact that Susan has P' are the same fact, but he cannot make the same claim about the facts of Susan's thinking about P and Susan's thinking about P'.

6.2 A POSSIBLE REJOINDER

The physicalist of a certain sort invites us to ascend levels and to appreciate the difference between physical and phenomenal concepts as a reason why we are generally inclined to think that phenomenal properties are not physical. There seems to be a sort of what-it-is-like aspect of phenomenal states that resists a physicalist account. The physicalist of a kind tells us that this aspect needs to be assigned away from the phenomenal state to a mode of presentation. This move is plausible so far as it goes, but it does not go far enough. Now, the crucial issue becomes whether a state that features this mode of presentation is physical or not. Let me consider what a physicalist can say in response to the argument above to the effect that a state in which phenomenal modes of presentation figures cannot be physical because given the claim as to the independence of phenomenal and physical modes of presentation, there can be no physical mode of presentation that can be replaced by a phenomenal mode of presentation that gives us the same state. A physicalist can argue that the state featuring a phenomenal mode of presentation appears to us different from what it really is in the way that a phenomenal state that is grasped through a phenomenal mode of presentation appears to us different from what it really is. What we have

here is a hierarchy of modes of presentation of modes of presentation. Our access to the fact of Susan's thinking about *P* is, the physicalist may say, mediated by a mode of presentation of *P* that represents it in a way different from it really is. That is, Susan's thinking about *P* and Susan's thinking about *P'* are indeed the same fact, but they appear to be different because the modes of presentation of *P* and *P'* are different.

The main problem with this line of response is that it is dubious whether there are modes of presentation of modes of presentation in the sense required by the response. A mode of presentation is, by definition, "a mode" (whatever that is) through which something is presented to a subject. The subject's access to the thing is mediated by the mode of presentation, but there is no compelling reason to think that the subject's access to the mode of presentation is mediated by another mode of presentation. The mode of presentation through which I think about Aristotle is the concept *the author of Metaphysics*, and in that sense my relation to Aristotle is mediated. But my access to the concept *the author of Metaphysics* is not mediated. In fact, what explains the queerness of the question "how do you know that the concept you employ when you think about Aristotle is *the author of Metaphysics*?" is the immediate nature of my access to my concepts. Notice that there is no analogous queerness in asking the question "how do you know that Aristotle falls under the concept *the author of Metaphysics*?"

Moreover, if my access to a certain first-level mode of presentation were always through a certain higher-level mode of presentation, then that would give rise to a question about my access to that higher-level mode of presentation. We would then have a problem of regress. But the regress can be stopped before it starts by admitting an independently plausible idea, that is, that our access to modes of presentation is direct.

6.3 WHERE IS WHAT-IT-IS-LIKE?

Some physicalists are mistaken in thinking that assigning away the what-it-is-like aspect from the phenomenal state to a mode of presentation solves the problem of phenomenal consciousness. Now let me try to put forward some reasons why it is an implausible move. If the what-it-is-like aspect occurs at the level of modes of presentation, then there is nothing it is like to be in pain if being in pain is not represented to the subject by a mode of presentation. The physicalist that makes use of the idea of phenomenal concepts makes even a stronger claim: there is nothing it is like to be in pain if being in pain is not represented by the subject by a *conceptual* mode of presentation. An implication of this claim is that there is nothing it is like to be in pain for subjects that do not clearly conceptually represent the world (e.g., new-born infants). This implication clearly appears to be false. Phenomenal consciousness does not require any conceptual capacities to be present in a subject. In the temporal order, if not in the order of explanation, phenomenal consciousness comes first, and the ability to have certain types of conceptual representations is not one of its necessary conditions. So, assigning the what-it-is-like aspect to conceptual modes of presentation runs counter to a hypothesis that is so clear that we are inclined to take it as a datum: feeling what-it-is-like is independent of having concepts.

Another reason why I think that the move of assigning away the what-it-is-like aspect to a mode of presentation is implausible is that it deprives us of a very compelling account of how we acquire modes of presentation of phenomenal states. If the what-it-is-like aspect is something we find at the level of states, then a compelling account, a version of which we can find in Hume's *Treatise*, is that the modes of presentation of those states we have are acquired through a mechanism of "copying" that aspect. The what-it-is-like aspect is already there, in the world, as an aspect of a certain state; and, that aspect is reflected or mirrored by the mode of presentation

of that state through a direct relation the subject bears to it. The details needs to be added, but the rough picture is clear. Now, if the what-it-is-like aspect is not an aspect of the state but is only featured by the mode of presentation, then the question of how the mode of presentation is acquired cannot be answered in this way. But it is also very unclear how it can be answered at all. How do we happen to have these modes of presentation? Phenomenologically speaking, a mechanism of inwardly attending to a certain sensory state of mine is responsible for my having the modes of presentation of that state. And, if this is not the beginning of the correct answer, then I do not see what the correct answer can possibly be.

6.4 CONCEPTS, PROPERTIES, AND CARTESIANISM

Concepts represent the world to be in certain ways to the subject who possesses them. My concept of Aristotle may be different from that of a professor of classics: the way Aristotle is represented in his mind is different from the way Aristotle is represented in mine. Moreover, Aristotle is different from my concept of Aristotle: Aristotle is (or was) a man, but my concept is not. My concept *represents* Aristotle. I may have different concepts that represent Aristotle, the very same man, without being aware of it. I may have the concept *the author of Metaphysics* and *the best pupil of Plato*; and, even if my grasp of these concepts is not deficient, I may lack the knowledge that the author of Metaphysics is the best pupil of Plato. The concept *the author of Metaphysics* stands for Aristotle but it is different from another concept that stands for the same man. Frege dubbed this difference between concepts as a difference in their modes of presentation. An important point is that modes of presentation are *contents*. They are ingredients

of thoughts that bear truth-values. The thought that the author of *Metaphysics* is the best pupil of Plato is different from the thought that the author of *Metaphysics* is the author of *Metaphysics*. A subject may reasonably affirm the latter while denying the former. What explains the difference between the two thoughts is a difference in the *contents* of the concepts *the author of Metaphysics* and *the best pupil of Plato*. Let us call the referent or extension of a concept its representational content, and its mode of presentation its cognitive content. What is the relation between these two types of contents? If what I have claimed in the previous chapter is correct, representational contents of phenomenal concepts are their cognitive contents. More accurately, cognitive contents of phenomenal concepts are inherited from their representational contents.

According to the picture I have been alluding to, there are phenomenal properties, the instantiation of which constitutes phenomenal states, and there are subjects that stand in a sort of direct epistemic relation (namely, acquaintance) to those properties and states. Subjects acquire concepts of those properties in virtue of the relation in question, and the existence and cognitive contents of concepts are entirely accounted for by an appeal to properties, subjects, and the relation between the two. In this picture, the subject is in a position to directly demonstrate phenomenal properties, and the cognitive contents of concepts he forms of phenomenal properties via demonstration are these properties themselves. The virtues of this picture I would like to stress are that it is phenomenologically adequate (our relation to our phenomenal states, from the subject's point of view, *appears* to be direct) and that it is in a position to answer the question how cognitive contents are formed. Moreover, this picture enables us to make sense of the ambivalence that is commonly attributed to the empiricist tradition, and especially to Locke, about the usage of the term 'idea'. A common worry is that traditional empiricists such as Locke were not very careful in their deployment of the term 'idea' and it is not clear whether it is

intended by them to pick out the representational content or what does the representing. But, Locke and others were at least partly justified in their carelessness if, in the case of properties with which we are acquainted in our experiences, what does the representing is something inherited from the representational content.

There are some accounts of concepts that attempt to understand all concepts according to a certain view of logical constants like conjunction. The idea is that the cognitive content of the concept *and* is exhaustively accounted for by the appropriate “introduction” and “elimination” rules like the following:

$$\frac{P \text{ and } Q}{P}$$

$$\frac{P \text{ and } Q}{Q}$$

$$\frac{P}{Q}$$

$$P \text{ and } Q$$

According to this picture, what the subject knows when he possesses the logical concept *and* is nothing but an implicit or explicit knowledge of these rules. And, as Lewis Carroll’s Achilles and the Tortoise forcefully makes us aware, rules are not statements. On this picture, concepts enable the subject to do certain things like inferring, but there is nothing more to the concept *and* than what the subject is capable of doing thanks to having it. The subject’s possession of the concept *and* is not a matter of grasping a certain content but a matter of ability to make certain inferences given certain premises.

Whatever the merits this view of logical constants has for other kinds of concepts, I do not think that this view of logical constants is the right model to understand phenomenal concepts. Imagine a human being, Susan, who is incapable of feeling pain. Being incapable of feeling pain

is no hinderance against making inferences about pain. Susan may know what is sometimes called “folk psychology” and all the truths incorporated in it. This knowledge as well as information about certain conditions may enable Susan to make inferences about pain. For instance, Susan may draw the inference from Mark’s behaving in certain ways that Mark is in pain, and further draw the inference from Mark’s being in pain that Mark has a certain feeling. With respect to her capacities of making inferences about pain feeling given certain premises, Susan may be as perfect as any other human being. Despite this, Susan does not grasp the concept *pain* or, if you wish, her grasp of the concept pain is deficient, because of her incapacity to feel pain. What Susan does not fully grasp is the cognitive content of the concept *pain*. She is not in a position to fully understand the contents of some thoughts such as the content of the thought *that Mark has pain*. So, there is more to the concept *pain* than being in a position to make inferences about pain.

Phenomenal concepts have cognitive contents that are inherited from their representational contents (i.e., phenomenal properties); and having a phenomenal concept requires grasp of those cognitive contents. If this is so, then a Cartesian view of concepts is correct at least for phenomenal concepts. According to Cartesianism, as Fodor (2005) defines it, having a concept of X is being able to think about X as such. And, being able to think about X as such requires grasp of a mental content that represents X. Now, on the picture I have drawn, having a certain phenomenal concept requires grasp of a certain cognitive content that represents a certain phenomenal experience. One cannot be said to have a phenomenal concept without having a grasp of its cognitive content, even if one’s abilities to draw inferences about its representational content is as perfect as possible.

7.0 CONCLUSIONS (STILL MORE PROMISSORY NOTES)

There are, as I am and you are well aware, many loose ends in this dissertation. In this concluding chapter, I would like to say a couple of things about those points that I wish to further explore in my future research.

7.1 THE MYSTERIES OF DIRECT EPISTEMIC RELATION

This dissertation started as an investigation of an argument against physicalism but it quickly transformed itself into an assessment of the nature of the epistemic relationship subjects bear to their own phenomenal states. The transition was smooth because of two reasons. One reason is that the argument in question (the Knowledge Argument) is an epistemic one that draws upon the impossibility of a priori derivability of a piece of knowledge from a set of other pieces of knowledge. The argument takes for granted that knowledge of phenomenal states gained on the basis of experiencing cannot be a priori derived from knowledge of brain states gained through scientific lessons, and one cannot but wonder what it is about our epistemic relation to our own states that makes such a derivation impossible. Another reason is that a dominant strategy (the Phenomenal Concept Strategy) developed in response to the said argument accounts for the

impossibility of a priori derivation in question by an appeal to the special nature of the concepts under which we grasp phenomenal states. The question that naturally arises is what the epistemic conditions are under which we form and possess those special concepts. As I see it, by directing the attention to the nature of phenomenal concepts, the strategy at hand captures the heart of the matter, though it may not, and as I believe it will not, be happy in the end about the account that is plausibly forthcoming. The allure of the strategy as an attempt to save physicalism disappears as soon as we ask the question where the contents or, as they say, modes of presentation of phenomenal concepts come from.

I have argued that the epistemic relation a subject has to his own phenomenal states is direct in that his beliefs formed on the basis of having those phenomenal states are non-inferentially justified. The notion of direct epistemic relation (call it acquaintance as I do, or apprehension, or what have you) is found mysterious by some philosophers for different reasons. One reason derives from the writings of Sellars, some of which give the impression that the mystery in question resides in what direct epistemic relation implies. Some of Sellars' criticism of the Myth of the Given (especially his "inconsistent triad" objection against classical sense-data theories) can be plausibly construed as suggesting that direct epistemic relation entails or guarantees or is by itself sufficient for knowledge. The suggestion is that if a subject S is in a direct epistemic relation to Y, then S knows Y. In chapter five, I argued that this suggestion is mistaken. A subject can be in direct epistemic relation to, say, his current pain states without knowing that he is in pain if he does not have the concept *pain*. The sense in which a certain subject is in direct epistemic relation to his phenomenal states is that *if* the subject has phenomenal beliefs about those states, *then* his beliefs are immediately justified by those states; but this is clearly consistent with the possibility that the subject lacks those beliefs.

Another reason why direct epistemic relation is found mysterious is that a positive characterization of it is missing. Direct epistemic relation is, as I take it, a relation that makes non-inferential justification possible. Clearly, this characterization is negative simply because “non-inferential justification” means “justification that is *not* inferential.” Aside from some schematic gestures, I have not given a positive account of what non-inferential justification is. But, notice that even when we have a positive characterization of justification that is not inferential (e.g., something like “a justification that is not inferential is a justification that has the features F, G, and H”), this would not by itself be sufficient for a positive characterization of direct epistemic relation. Because, in that case, although we would know that direct epistemic relation is a relation that makes a justification with the features F, G, and H possible, we would not know what it is about that relation that makes such a justification possible. Compare: from our observations of the rational behaviors of a given organism, we can plausibly infer that the organism has a certain inner mechanism that makes those behaviors possible. But this inference would not enable us to know what it is about that mechanism that makes those behaviors possible. The same goes for the inference from the existence of non-inferential justification to the existence of direct epistemic relation. Some philosophers (Fumerton 1995) argue that the relation we have to the direct epistemic relation that connects us to our experiences is also epistemically direct. If this view is correct, then we are in a position to answer the question what it is about our relation to our own experiences that makes non-inferential justification for phenomenal beliefs possible. But, I have not argued for such a view.

So, it is true that my account of direct epistemic relation is negative in both respects. I have not given a positive characterization of non-inferential justification, and I have not provided a satisfactory answer to the question what it is that makes our epistemic relation to our

experiences direct. However, I take these points as serious challenges rather than decisive objections. There are various accounts (BonJour and Sosa 2003) one can find in the literature that address these and nearby issues. A more focused attention to the nature of non-inferential justification and direct epistemic relation is *sine qua non* for a serious development of the project to which I committed myself by this dissertation.

7.2 LOGICAL ATOMISM

Sellars combined his assault on the Myth of the Given with an attack on logical atomism. Let us try to find an answer to the question whether my account of phenomenal concepts is logical atomist in character. Sellars writes:

I have arrived at a stage in my argument which is, at least *prima facie*, out of step with the basic presuppositions of logical atomism. Thus, as long as looking green is taken to be the notion to which being green is reducible, it could be claimed with considerable plausibility that fundamental concepts pertaining to observable fact have that logical independence of one another which is characteristic of the empiricist tradition. (Sellars 1956, p. 147)

What is of interest to me at this moment is what Sellars takes logical atomism to be. As I understand from this passage, logical atomism is a conjunction of two theses. (1) There are fundamental concepts. And, (2) fundamental concepts are logically independent of one another. Talking in Sellars' terminology, if *looking green*, *looking red*, and *looking yellow* are fundamental observational concepts, then logical atomism holds that these concepts are logically

independent of each other. Let me now try to figure out what the notions of “fundamental concepts” and “logical independence” come to in Sellars’ account respectively, and then check whether my view of phenomenal concepts is logical atomist.

What is a fundamental concept? There is no formal definition one can find in Sellars (1956) but its talk of reducibility (p. 147) is suggestive. On the empiricist tradition, Sellars says, *being green* is taken to be reducible to *looking green*, and hence *looking green* is taken to be more fundamental than *being green*. In general, then, we can say that a concept *X* is more fundamental than another concept *Y* given that *Y* is reducible to *X*. Of course, the empiricist tradition Sellars has in mind takes look-concepts not only to be more fundamental than is-concepts but also to be *the* fundamental concepts that cannot be reduced to other concepts. In this sense, a concept *X* is fundamental *simpliciter* given that *X* cannot be reduced to any other concepts. It is clear that Sellars thinks that the concept *being green* is more fundamental than the concept *looking green* (the story of the necktie seller John is, I take it, intended to establish that idea (pp.142-4)); but, as far as I can tell, it is not very clear whether he thinks that the concept *being green* is fundamental *simpliciter*. Now, on my account, phenomenal concepts are simple in that they are not constituted by any other concepts: their contents are completely inherited from their referents. If this is so, it is clear that phenomenal concepts are not reducible to any other concepts, and hence that they are fundamental *simpliciter*.

What about “logical independence”? What is it to say that a concept is logically independent from another concept? Again, there is no formal definition one can find in Sellars (1956), however its talk of “presupposition” is suggestive (p. 146). A concept *X* is logically independent of another concept *Y* given that *X* does not presuppose *Y*. The notion of presupposition in turn can be understood in terms of the notion of analysis (p. 148): a concept *X*

presupposes another concept *Y* given that *X* can be analyzed in terms of *Y*. Intuitively, the concept *die* is logically independent of the concept *kill* while the latter is not logically independent of the former, because the concept *die* cannot be analyzed in terms of the concept *kill* while the latter can be analyzed in terms of the former (e.g., killing is, roughly, causing to die). Now, given this, the second thesis of logical atomism is that fundamental concepts cannot be analyzed in terms of each other. It is, again, clear that my account of phenomenal concepts endorses this thesis too; and, again, the reason is the simplicity of phenomenal concepts. So, the conclusion is that my account of phenomenal concepts is logical atomist.

It is worth noting as a side point that one can hold the first thesis of logical atomism without holding the second. Imagine that there is a set of concepts *X*, *Y*, and *Z* to which the concepts *A*, *B*, and *C* are reducible. The concept *A* is reducible to the concept *X and Y*, and the concept *B* is reducible to the concept *X and Z*, and the concept *C* is reducible to the concept *Y and Z*. In this case, the concepts *X*, *Y*, and *Z* are more fundamental than the concepts *A*, *B*, and *C*. But notice that this is consistent with the possibility that *X*, *Y*, and *Z* are not logically independent of each other. In particular, the fact that *X*, *Y* and *Z* are fundamental concepts does not block the possibility that *X* can be analyzed in terms of *Y* and *Z*, *Y* can be analyzed in terms of *X* and *Z*, and *Z* can be analyzed in terms of *X* and *Y*. Of course, in this case, there is a kind of circularity of analysis but it is not clear how serious a problem this is. According to the Newtonian framework, for instance, the concept *force* is analyzable in terms of *mass* and *acceleration*, the concept *mass* is analyzable in terms of *force* and *acceleration*, and the concept *acceleration* is analyzable in terms of *mass* and *force*. The basic terms of a theory may form an interconnected web in which each term in that web is defined in other terms.

My account of phenomenal concepts is, let me repeat, logical atomist in the sense above. But there is another conception of atomism one can find in Sellars, and there is nothing in my account that commits me to that sort of atomism. Sellars writes:

[O]ne can have the concept of green only by having a whole battery of concepts of which it is one element... [W]hile the process of acquiring the concept of green may –indeed does– involve a long history of acquiring *piecemeal* habits of response to various objects in various circumstances, there is an important sense in which one has *no* concept pertaining to the observable properties of physical objects in Space and Time unless one has them all –and, indeed, as we shall see, a great deal more besides. (1956, p. 148, italics original)

I am not concerned about the plausibility of the idea that one cannot have the concept *green* without having *all* other observational concepts presumably including the perceptual visual concepts *magenta*, *saffron*, and *ceil*. I think the idea is clearly false, and *something* must have gone wrong before if this is a conclusion that follows from some of Sellars' earlier assumptions or presuppositions. What I would like to note now is that the notion of atomism employed in this passage is different from logical atomism as it is defined above. Here Sellars talks about the *conditions of acquisition* of concepts rather than their irreducibility or logical independence. It is one thing to claim that there are concepts that cannot be analyzed in terms of other concepts, and it is yet quite another thing to claim that there are concepts that can be acquired by the subject in isolation from all other concepts. I hold that phenomenal concepts cannot be analyzed in terms of other concepts but I would be hesitant to claim that a certain phenomenal concept (e.g., one that is of the experience of seeing green) can be acquired by a subject that does not possess any other concepts. It is, after all, very hard to imagine a subject who possesses just one concept whether

phenomenal or not. The idea that concept *acquisition* is a one-thing-at-a-time affair has an implausible ring whether or not there are fundamental concepts that are logically independent of each other.

Sellars conflates logical atomism with what I may call “acquisitional atomism.” This is an understandable mistake. The falsity of logical atomism appears to imply the falsity of acquisitional atomism, but logical atomism does not imply acquisitional atomism. Take the concept *watch*, and the internet gives the definition *a small portable timepiece*. Now, clearly, then, the concept *watch* is a complex concept that can be analyzed in terms of other concepts. So, a subject who does not possess the concepts in terms of which the concept *watch* is analyzed can be plausibly said not to possess that concept. Now, if all our concepts are like *watch* in this respect, then the conclusion is that acquisitional atomism is false. Now assume that logical atomism is true. There are some fundamental concepts that cannot be analyzed by other concepts. But this is consistent with the falsity of acquisitional atomism. It is perhaps the case that our minds are, as a matter of fact, structured in such a way that they are not capable of acquiring concepts in isolation from all other concepts. Or, perhaps, one can only be said to acquire a certain concept if one is in a position to make certain judgments using that concept; and since making a judgment requires at least two concepts, having only one concept is impossible in principle. My point is that the truth of logical atomism does not block these possibilities.

So far as I know, a distinction between logical and acquisitional atomism is not properly addressed in the literature of concepts. I believe the hot debates between conceptual holists, localists, and atomists may benefit from such a distinction, the exact nature of which needs to be studied more carefully than I have done so far.

7.3 CONCEPTUALIZATION

I define conceptualization as a process that takes non-conceptual content as input and gives conceptual content as output. The question is this: is there a conceptualization process or is this definition of conceptualization empty?

Some philosophers might want to argue that there is nothing that corresponds to the definition at hand because it is inconsistent. Non-conceptual content is, they may add, to be understood along the lines “a given content is non-conceptual only if it cannot be the content of a concept.” It is true that if this is the notion of non-conceptual content that one has in mind, then there cannot be a process of conceptualization. But it is also true that it is highly doubtful whether there is non-conceptual content in this sense of the term. How do we know whether there are non-conceptual contents in this sense if our knowledge of contents is only through our concepts? In any case, this is not the notion of non-conceptual content I adopt. I take non-conceptual content to be a sort of content the having of which does not require grasp of concepts. Phenomenal experiences are paradigm examples of states with non-conceptual content in this sense of the term. Now, the question is whether there is a conceptualization process that takes experiential content as input and gives conceptual content as output.

Intuitively, there is a process of conceptualization. Assume that I have a certain visual experience of a certain shade *S* of red that I have not had before, and assume also that I do not have a visual concept of that shade of red. The content of my experience – the particular redness of that shade – is non-conceptual. Now, assume that while having the experience, someone asks me the question “what kind of experience are you having now?” This question leads me to direct my attention to my current experience and introspect it to figure out what it is like. As a result, I form the occurrent (conceptual) thought *that I am seeing S*, where *S* is a concept that stands for

the particular shade of red present in my experience. It strongly appears that, in this case, there is a transition from non-conceptual content (namely, S) to conceptual content (namely, S). The redness of the shade in my experience is taken as input, and a certain concept of that redness that is an ingredient of my thought is given as an output. The content of the concept S is formed as a result of my attending to the content of my experience S .

At a more theoretical level, it also seems that we need a process of conceptualization to explain the ontogeny of our perceptual concepts of proper sensibles. Where do the contents of these concepts come from? If there is a process of conceptualization, then the answer is straightforward: we have non-conceptual states; and, by virtue of conceptualization, the contents of these states are transformed into conceptual contents. So, if there is a conceptualization process, the contents of our concepts come from the contents of our non-conceptual states. However, if there is no process of conceptualization, then how can we explain the contents of our concepts? Does the reflective mind randomly create, without any input, the contents of the concepts with which it thinks? If the reflective mind is not restricted with respect to the concepts it can form and employ, is there any sense that we can attach to its being justified for its beliefs or to its applying its concepts to outer things? I do not think there are any plausible answers forthcoming.

Admitting that there is a process of conceptualization is a way of capturing the insight that the mind's power to form concepts is structured by the contents of experiences that are caused by the external world. The mind does not act in a vacuum or operate in a space of contents it itself creates out of nothing. The external world causes certain states with certain contents in the subject and thus provides the raw material with which the mind can work. The mind's "touch with the reality" is through the non-conceptual contents the reality causes. The

logical space of reasons and the logical space of causes, if you want to talk in these terms, are not independent of each other; rather, the latter affects the former by endowing it with the materials with which it can operate.

In order to appreciate that the contents of our concepts come from the contents of our non-conceptual empirical states, we need not accept a wholesale conceptual empiricism, i.e., the thesis that we have only those concepts the contents of which are possible objects of experiences. Conceptual empiricism as a global thesis is *very* strong. It certainly seems that our moral, mathematical, and logical concepts (and maybe some of our metaphysical concepts such as *space*, *time*, *substance*, and *causality*) have contents that cannot in principle be given in our perceptual or phenomenal experiences, and the thesis that contents of these concepts are ultimately end products of conceptualizations of experiential contents stands in need of a *very* strong argument. However, my endorsement of conceptualization needs to be understood as an endorsement of conceptual empiricism as a local thesis about the contents of perceptual and phenomenal concepts. My claim is that the contents of our perceptual and phenomenal concepts come from the contents of our experiences. This is a relatively weaker and much less controversial claim than its global version.

Is the thesis that concepts acquire their contents from non-conceptual states a form of what is sometimes called the resemblance theory of mental representation? According to the resemblance theory, what a concept represents is what it resembles most. Now, this theory appears to be clearly false for a good number of concepts. The concept *justice* represents justice but it does not resemble it. The same goes for our mathematical, logical, and metaphysical concepts. Does this theory work for our perceptual and phenomenal concepts? In this case, the theory is not *clearly* false if the contents of these concepts come from what they represent

(proper sensibles and phenomenal qualities, respectively). If the contents of perceptual and phenomenal concepts come from what they represent, then it is plausible that these concepts resemble what they represent. But, still, this need not commit us to a resemblance theory of perceptual and phenomenal concepts. On a resemblance theory, the concept *red* represents redness *because* redness is what it resembles. This because-claim is something a conceptualization account is not necessarily committed to. One may hold both that the concept *red* represents redness and that redness is what *red* resembles without holding that what causes the former is the latter. One may instead attempt to explain why *red* represents redness in terms of a causal-informational or teleological account of mental representation. One may, for instance, claim that the concept *red* refers to the property that normally causes classificatory uses of that concept. This claim does not apparently attribute any role to the fact that *red* resembles redness in the explanation of what *red* represents. I have not committed myself to the because-claim that is definitive of the resemblance theory; and so, my account of the contents of perceptual and phenomenal concepts is not a version of that theory.

Notice that conceptualization presupposes a primitive capacity of non-conceptual attendance or awareness. Conceptualization requires taking non-conceptual contents as inputs. In order to take those contents as inputs, the conceptualizing subject needs to attend to or to be aware of those contents. Now, clearly, if this attendance or awareness were possible only if one has the relevant concept, then the whole idea of conceptualization would be discarded simply because there would be no need for it in the presence of that concept. If awareness of a shade of red, which is necessary for conceptualization to get off the ground, were possible only if one has the visual concept of that shade, then we would not need a process of conceptualization to form the concept in the first place simply because we would already have the concept. So, if there is a

process of conceptualization, then there is non-conceptual awareness of contents that serve as inputs to that process. This is a controversial result but it seems to me to be inescapable.

Of course, it is one thing to argue that there is a process of conceptualization, and it is an altogether different thing to provide a detailed account of what it is. Seeing red and thinking red as red are two different things, and what I have said so far is intended to point out only that passing from the former to the latter requires conceptualization. But I have not said anything about how we should understand conceptualization, whether it is an abstractive process or, as some like to put it, a sort of digitalization of analog contents or something else. This is clearly a very important question that needs to be addressed with a more sustained attention.

BIBLIOGRAPHY

- Armstrong, D. 1968. *A Materialist Theory of the Mind*. Routledge.
- Balog, K. "Acquaintance and the Mind-Body Problem." Forthcoming in *Identity Theory*, eds. Christopher Hill and Simone Gozzano. Cambridge University Press.
- , "In Defense of Phenomenal Concept Strategy." Forthcoming in *Philosophy and Phenomenological Research*.
- Block, N. 2002. "The Harder Problem of Consciousness." *The Journal of Philosophy* XCIX, 8, 391-425.
- , 2007. "Max Black's Objection to the Mind-Body Identity Thesis." In *Phenomenal Concepts and Phenomenal Knowledge*, eds. Torin Alter and Sven Walter, 249-307. Oxford University Press.
- BonJour, L and E. Sosa. 2003. *Epistemic Justification*. Australia: Blackwell Publishing.
- Boyd, R. 1980. "Materialism without Reductionism: What Physicalism Does not Entail." In *Readings in Philosophy of Psychology*, Vol. 1, ed. Ned Block: 67-106. Cambridge: Harvard University Press.
- Brandom, R. "No Experience Necessary." Unpublished Manuscript: 1-27.
- Broad, C.D. 1925. *The Mind and its Place in Nature*. London: Routledge and Kegan Paul.
- Carruthers, P. 2000. *Phenomenal Consciousness*. Cambridge: Cambridge University Press.
- Castaneda, Hector-Henri. "Indicators and Quasi-Indicators." *American Philosophical Quarterly*, 4 (1967): 85-100.
- Chalmers, D. 1996. *The Conscious Mind*. Oxford University Press.
- , 2003. "The Content and Epistemology of Phenomenal Belief." In *Consciousness: New Philosophical Perspectives*, eds. Q. Smith and A. Jokic, Oxford University Press.

- , 2007. "Phenomenal Concepts and the Explanatory Gap." In *Phenomenal Concepts and Phenomenal Knowledge*, eds. Torin Alter and Sven Walter, 165-197. Oxford University Press.
- Conee, E. 1994, "Phenomenal Knowledge." *Australasian Journal of Philosophy* 72: 136–150.
- Crane, T. and D. H. Mellor. 1990. "There is no Question of Physicalism." *Mind* 99 (394): 185-206.
- Dennett, D. 1991. *Consciousness Explained*. Boston: Little, Brown.
- , 1998. "Quining Qualia." In *Consciousness in Contemporary Science*, eds. A. Marcel and E. Bisiach, Oxford: Oxford University Press, pp. 42-77.
- , 2007. "What RoboMary Knows," In *Phenomenal Concepts and Phenomenal Knowledge*, eds. Torin Alter and Sven Walter, Oxford University Press, pp. 15-32.
- Dunne, J. W. 1958. *An Experiment with Time*. London: Faber and Faber.
- Farrell, B. A. 1950. "Experience." *Mind* 59: 107-198.
- Feigl, Herbert. 1958. "The 'Mental' and the 'Physical'." In *Concepts, Theories, and the Mind-Body Problem*, ed. Herbert Feigl, Michael Scriven, and Grover Maxwell, 370-497. Minneapolis: University of Minnesota Press.
- , 1967. *The 'Mental' and the 'Physical': The Essay and a Postscript*. Minneapolis: University of Minnesota Press.
- Fodor, Jerry A. 2005. *Humean Variations*. Oxford: Clarendon Press.
- Frege, G. "On Sense and Reference." in A.W. Moore (ed.) *Meaning and Reference*. Oxford: Oxford University Press
- Fumerton, R. 1995. *Metaepistemology and Skepticism*. Lanham: Rowman and Littlefield.
- Gertler, B. 2001. "Introspecting Mental States." *Philosophy and Phenomenological Research* 63 (2): 305-328.
- Heck, R. G. "Nonconceptual Content and the Space of Reasons." *Philosophical Review* 109: 483-523.
- Hill, C. 1991. *Sensations: A Defense of Type Materialism*. Cambridge University Press.
- Hill, C. and B. McLaughlin. 1999. "There Are Fewer Things Than Are Dreamt of in Chalmers's Philosophy." *Philosophy and Phenomenological Research* LIX: 445-454.
- Hintikka, J. 1970. "Knowing that One Knows." *Synthese* 21: 141-62.

- Horgan, T. and J. Tienson. 2001. "Deconstructing New Wave Materialism." In *Physicalism and its Discontents*, eds. Carl Gillett and Barry Loewer. Cambridge University Press.
- Jackson, F. 1982. "Epiphenomenal Qualia." *Philosophical Quarterly* 32: 127-136.
- , 1998. *From Metaphysics to Ethics: A Defense of Conceptual Analysis*. Oxford University Press.
- Jackson, Frank. 1982. "Epiphenomenal Qualia." *Philosophical Quarterly* 32, pp. 127-36.
- , 1986. "What Mary Didn't Know." *Journal of Philosophy* 83: 291-295.
- Kaplan, D. 1989. "Demonstratives." In *Themes from Kaplan*, eds. Joseph Almog, John Perry, and Howard Wettstein. Oxford University Press.
- Kripke, Saul. 1980. *Naming and Necessity*. Cambridge: Harvard University Press.
- Levin, J. 2007. "What is a Phenomenal Concept?" In *Phenomenal Concepts and Phenomenal Knowledge*, ed. Torin Alter and Sven Walter, 87-111. Oxford University Press.
- Levine, Joseph. 1983. "Materialism and Qualia: The Explanatory Gap." *Pacific Philosophical Quarterly* 64: 354-61.
- , 1998. "Conceivability and the Metaphysics of Mind." *Nous* 32: 449-80.
- , 2007. "Phenomenal Concepts and the Materialist Constraint." In *Phenomenal Concepts and Phenomenal Knowledge*, ed. Torin Alter and Sven Walter, 145-167. Oxford University Press.
- , 2010. "Demonstrative Thought." *Mind and Language* 25 (2): 169-195.
- Lewis, D. 1966. "An Argument for Identity Theory." *The Journal of Philosophy* 63 (2): 17-25.
- , 1979. "Attitudes De Dicto and De Se." *The Philosophical Review*, 88, pp. 513-543.
- , 2004. "What Experience Teaches." In *There's Something about Mary*, eds. Peter Ludlow, Yujin Nagasawa, and Daniel Stoljar, Cambridge: Massachusetts, pp. 77-105.
- Lycan, W. 1996. *Consciousness and Experience*. MIT Press.
- Loar, Brian. "Phenomenal States (Revised Version)," (originally published in 1990, revised in 1997) in Ludlow, Nagasawa and Stoljar 2004, 219-241.
- Ludlow, Peter, Yujin Nagasawa, and Daniel Stoljar, eds. 2004. *There's Something About Mary*. Cambridge, Mass.: The MIT Press.
- McDowell, J. 1996. *Mind and World*. Harvard University Press.
- McGinn, C. 1989. "Can We Solve the Mind-Body Problem?" *Mind* 98 (July): 349-66.

- , 2001. "How not to Solve the Mind-Body Problem." In *Physicalism and its Discontents*, eds. Carl Gillett and Barry Loewer. Cambridge University Press.
- McLaughlin, B. 2001. "In Defense of New Wave Materialism" In *Physicalism and its Discontents*, eds. Carl Gillett and Barry Loewer. Cambridge University Press.
- Meehl, Paul E. 1966. "The Complete Autocerebroscopist." In *Mind, Matter, and Method: Essays in Philosophy and Science in Honor of Herbert Feigl*, ed. Paul Feyerabend and Grover Maxwell, 103-180. Minneapolis: University of Minnesota Press.
- Nagel, Thomas. 1974. "What is it Like to Be a Bat?" *Philosophical Review* 83: 435-450.
- Nida-Rumelin, Martine. "What Mary Couldn't Know" (originally in 1996) republished in Ludlow, Nagasawa and Stoljar 2004, pp. 241-269.
- O'Dea, J. 2002. "The Indexical Nature of Sensory Concepts." *Philosophical Papers* 32 (2): 169-181.
- Papineau, D. 2004. *Thinking about Consciousness*. Clarendon Press: Oxford.
- Peacocke, C. 1992. *A Study of Concepts*. Cambridge: MIT Press.
- Perry, J. 1979. "Frege on Demonstratives." *Philosophical Review* 86 (4): 474-97.
- , 1999. *Knowledge, Possibility, and Consciousness*. Cambridge, Mass.: MIT Press.
- Robinson, H. 1982. *Matter and Sense*. Cambridge: Cambridge University Press.
- Robinson, H. "Dennett on the Knowledge Argument." *Analysis* 53, pp. 174-177.
- Russell, B. 1912. *Problems of Philosophy*. Oxford: Oxford University Press.
- Sellars, W. 1956. "Empiricism and the Philosophy of Mind." In *Science, Perception and Reality*, Atascadero: Ridgeview Publishing Company
- Stoljar, D. 2001. "Two Conceptions of the Physical." *Philosophical Quarterly* 34 (1984): 147-152.
- , 2005. "Physicalism and Phenomenal Concepts." *Mind and Language* 20: pp. 469-494.
- Tye, M. 2000. *Color, Consciousness and Content*. MIT Press.