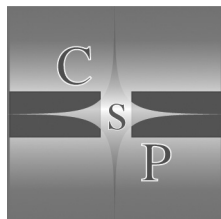


Content, Consciousness, and Perception

Content, Consciousness, and Perception
Essays in Contemporary Philosophy of Mind

Edited by

Ezio Di Nucci and Conor McHugh



CAMBRIDGE SCHOLARS PRESS

Content, Consciousness, and Perception: Essays in Contemporary Philosophy of Mind,
edited by Ezio Di Nucci and Conor McHugh

This book first published 2006 by

Cambridge Scholars Press

15 Angerton Gardens, Newcastle, NE5 2JA, UK

British Library Cataloguing in Publication Data
A catalogue record for this book is available from the British Library

Copyright © 2006 by Ezio Di Nucci and Conor McHugh and contributors

All rights for this book reserved. No part of this book may be reproduced, stored in a retrieval system,
or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording or
otherwise, without the prior permission of the copyright owner.

ISBN 1-84718-017-5

TABLE OF CONTENTS

Editors' Prefaceix

Chapter One: Introduction

The State of Mind

Conor McHugh and Ezio Di Nucci 1

 1 The mental and the physical 3

 2 Content, perception and concepts 11

 Bibliography 31

Part I: Concepts and Mental Content35

Chapter Two

Concepts: Representations or Abilities?

Hans-Johann Glock..... 36

 1 Are concepts representations? 38

 2 Are concepts signified by computational/neural signs?..... 45

 3 Are concepts abilities?..... 48

 4 Concepts or concept-possession? 49

 5 Is concept-possession a disposition or ability? 51

 6 Is the pragmatist explanation of concept-possession circular? 53

 7 Are concepts more finely individuated than abilities?..... 57

 Bibliography 59

Chapter Three

Human Philosophy: Hume on Natural Instincts and Belief Formation

Treasa Campbell62

 1 The traditional use of the term “natural belief” 63

 2 Wolff's propensity/disposition distinction 66

 3 Natural instincts 67

 4 Conclusion 72

 Bibliography 72

Chapter Four**Can Conscious Attention Restore Uncritical Realism?
An Examination of a Neo-Russellian View of Reference**

Gloria Ayob	74
1 Identifying Wittgenstein's worry about the Russellian view of reference	76
2 Campbell's response: Conscious attention to an object does not require sortal classification.....	79
3 Refining the Wittgensteinian worry about the Russellian pre-reflective view	81
4 A problem for Campbell's neo-Russellian semantics.....	85
Bibliography	87

Chapter Five**On Peacocke's Theory of Concepts**

Asunción Álvarez	89
1 The theory.....	90
2 Reference.....	92
3 Implicit conceptions	98
Bibliography	101

Part II: Consciousness and the Metaphysics of Mind 103**Chapter Six****Deflationism about Qualia**

Dave Ward	104
1 The explanatory gap	104
2 The obligations of a deflationist theory	107
3 Intrinsicity.....	111
4 Ineffability	114
5 Privacy.....	118
6 Simplicity	118
7 Direct availability	120
8 Conclusion.....	121
Bibliography.....	123

Chapter Seven**Heidegger and AI: Technological Metaphor and Self-Understanding**

Richard Hamilton	125
1 Introducing the subject	125
2 Reflective self-understanding	128

3 Heidegger	130
Bibliography	137
Chapter Eight	
Causal Exclusion and Overdetermination	
Markus E. Schlosser	139
1 Non-reductive physicalism	139
2 The causal exclusion argument	141
3 Events and property instantiations	143
4 Conclusion	155
Bibliography	155
Part III: Perception	157
Chapter Nine	
Metaphysical Disjunctivism and the Intentional Theory of Perception	
Francisco Pereira	158
1 Introduction	158
2 Considering the argument from hallucination	163
3 Perceptual experience and distal objects	166
4 Phenomenal indiscriminability	168
5 Fundamental ontological asymmetry	172
Bibliography	177
Chapter Ten	
Change Blindness and Counterfactual Dependence	
Nick Jones	179
Bibliography	189
Chapter Eleven	
An Enactive Theory of Phenomenal Intentionality	
Julian Kiverstein	190
1 What is phenomenal intentionality?	192
2 Against reductive representationalism	194
3 Perceptual content and demonstrative thought	199
4 Towards an enactive theory of phenomenal intentionality	201
Bibliography	210
Contributors	212
Index	214

EDITORS' PREFACE

This book is about the mind. The introductory essay is intended to provide, for non-specialists, an overview of the concerns and themes of contemporary philosophy of mind, and to locate each chapter of the volume within the context of that manifold of overlapping concerns and themes. The essays themselves, on the other hand, constitute academic contributions to their respective specialties within the philosophy of mind. We hope, therefore, that the volume will be of interest both to specialists and to readers interested in getting to grips with what philosophers of mind do.

All the essays that compose this collection were presented on June 30th and July 1st 2005, in Edinburgh, at the Mind2005 conference. The book could never have happened without the conference, so we want to take this chance to thank everyone involved with Mind2005.

Our gratitude for making Mind2005 a success goes in the first place to our distinguished keynote speakers, Prof. Hans-Johann Glock, Prof. Christopher Peacocke, and Prof. Timothy Williamson. Their feedback at the conference improved all of the papers in this collection. Also, we want to thank Dr. Peter Milne, Head of Philosophy at Edinburgh University for all kinds of support. Thanks are also due to the conference's other sponsors, the Scots Philosophical Club, the Scottish Postgraduate Philosophy Association, the Mind Association, and the Analysis Trust. Thanks also to Zoë Payne for helping us organise the conference and for setting up its website; to Astrid Johnston for a beautiful conference poster; to Lisa McKeown for helping us with submissions. There are numerous others, including speakers, commentators, chairpersons and helpful friends, who contributed to the conference, and to whom we are grateful.

We want to thank our supervisors, Dr. Matthew Nudds (CM) and Dr. Bill Pollard (EDN). Conor McHugh is generously supported by AHRC doctoral award 2004/107894, and by a scholarship from the College of Humanities and Social Sciences at Edinburgh University. Ezio Di Nucci is supported by a scholarship from the University of Salerno. Finally, thanks to Amanda Millar and Andy Nercessian at Cambridge Scholars Press for making this book possible.

Edinburgh, 24.4.2006

CHAPTER ONE: INTRODUCTION

THE STATE OF MIND

CONOR MCHUGH AND EZIO DI NUCCI

“Mental processes just are queer,” Wittgenstein says in the *Investigations* (Wittgenstein 1953, §363).¹ The bulk of the peculiarity of the mental—mental processes, events and states—is that it is both *in* the world and *about* the world. The mind has a unique *perspective from within*. The mental is in the world because it is part of the natural universe; and it is about the world because it represents it. Remarks like Wittgenstein’s reflect the widespread assumption that we should treat all the processes, events and states we call “mental” as constituting a unified set of phenomena or natural kind—or at least as exhibiting the same sort of “queerness”, and therefore as posing a unified philosophical problem. One might wonder if there is any justification for approaching matters in this way. There is, after all, an impressive diversity among the states and episodes that we label “mental”: why should we suppose that a momentary pain in the leg and a longstanding belief that the earth is flat ought to be understood as the same sort of element of reality (c.f. Rorty 1979)? What sort of property do they have in common? Candidates for unifying properties have included that of being composed of an irreducible metaphysical substance (most famously by Descartes; see, more recently, Robinson 2003); that of standing in certain causal-functional relations (see Putnam, e.g. 1967; Dennett 1991); a biological property (Searle 1992); and that of having a conscious subject (Kant 1929; and see more recently, e.g., Eilan 1995).

One influential view has it that intentionality should be seen as the feature that unifies the mental. That is, the peculiar fact that (say) beliefs and pains can *represent the world* is what makes them mental. Intentionality will be the subject of

¹ Thanks to Dave Ward and to Tom Roberts, whose comments on an earlier draft of this paper improved it considerably. The authors are responsible for any remaining errors. Thanks also to Nick Jones for suggesting the title.

the second part of this introductory essay. Whatever one's view of what unifies the mental, there is no doubt that the notion of the mental is closely tied to the notions of consciousness and intentionality. Furthermore, minds are natural; their origin is different to that of other things that exhibit intentionality (words, for example, are human artefacts). Philosophy does not need to explain the origin of the mind (again, see Wittgenstein: we, as philosophers, are not doing natural history²), but that minds naturally emerged from the non-mental universe contributes to their being so queer.³ The queer nature of mind is what philosophy must explain; it must do so, says the fashion of the time, in a naturalistic framework; that is, approximately, philosophy mustn't use in its explanation anything that would make a scientist uncomfortable.

It is widely believed that this explanation will depend heavily on facts about another element of the natural order: *brains*. But how do mental phenomena relate to brain phenomena? (This is another version of the metaphysical question that bugs the philosophy of mind: what sort of thing *is* the mind?) Ask people to point their index finger at their mind; some people will promptly point in the direction of their head; some others will stare back at you, because you have not fooled them. There is a relationship between mind and brain, or at least between mental phenomena and certain brain phenomena; some have thought it a relationship of identity (Smart 1959), others of supervenience (Kim 1966).⁴ But it seems peculiar to try to point to the mind; the grammar of pointing belongs to body-talk rather than mind-talk.⁵

In this introductory essay we aim to put the articles of this collection into context, by describing the state of philosophy of mind in the 21st century—the philosophical problems which the articles herein are addressing, and some of the interrelations between those problems. In the first part of the essay we will discuss the metaphysical question of whether the mental can be accommodated within a naturalistic worldview, without sacrificing what we typically think of as essential features of our conscious mental lives. The second part of the essay will discuss the notion of intentionality, drawing out its connections with the philosophy of perception and the theory of content and concepts.

² See Wittgenstein (1953, §II, xii).

³ So much so that some deny that the mental emerged from the non-mental, because there is no such thing as the non-mental; see Chalmers's panpsychism (Chalmers 1996).

⁴ There are philosophers who think that the mind is *in* the brain; or at least who talk as if the mind is *in* the brain. Recently, at the 2005 APA Eastern Meeting in New York on Dec. 28th, John Searle repeatedly referred to the mind as being *in* the brain. In such circumstances, one is tempted to volunteer to cut someone's head open and look for it.

⁵ For another version of the idea that the mind is not *inside* the brain, see recent proposals that the mind lies in the interaction between brain, body, and environment – see Clark (2001), also Hurley and Noë (2003).

1 The mental and the physical

In this section we address three different but interrelated questions that bear on the metaphysics of mind:

- (1) whether there can be psychophysical laws—deterministic laws that rule upon the workings of the mind by linking it to the physical; the existence of such laws would entail a reduction of the mental to the physical;
- (2) whether the subjective quality of our conscious experience—our qualia—can be reconciled with a physicalist world-view;
- (3) whether mind can be artificially re-created by man.

All three questions have implications for what view we should take of the metaphysical status of the mind within nature. Perhaps they can all be seen as ways of asking the same question—“what sort of thing is the mind?”—but if one does not agree that all three questions address the same general problem, then one can just see (1)-(3) as giving a rhetorical arrangement for the discussion of the mental and the physical.

We will mention, but not refute, one common objection to the claim that these three questions are formulations of the same problem. The objection is based on the distinction (Chalmers 1996) between easy and hard problems of consciousness. Easy problems are those that can be solved by the explanation of some mental phenomenon in terms of functional or computational facts about the brain (how the brain processes information). A hard problem, on the other hand, cannot be solved in functional or computational terms. Chalmers thinks that there is a part of the philosophy of mind which can be conceded to reductionists: problems such as, for example, explaining the nature of propositional attitudes and their role in causing behaviour, are easy problems, and thus materialists are right in thinking that they can be understood in physical terms. The materialist accounts will be massively complicated (and are in many cases still missing), but such an account can be given for these problems.⁶ But explaining the existence and nature of qualia—the subjective properties of conscious experience in virtue of which there is something it is like to enjoy such experience—constitutes a hard problem, according to Chalmers. He believes there is no physicalist solution to the hard problem. If one, with Chalmers, thinks that there is not one unified problem of the metaphysical status of the mind, but two different sorts of problems demanding different solutions, then one will not agree that formulation (1) addresses the same problem as formulation (2), because there will be laws for some features of the mind, namely those to do with how the brain receives and processes information, but not

⁶ One might be tempted to add, “and one day we shall have such a massively complicated physicalist story”; but it must be emphasized that for physicalism to be true, in its application to easy problems or to any problems, it is not required that man actually discovers the physicalist story.

for others, namely those to do with the subjective character of experience; and one will think that formulation (3) must be interpreted in such a way that only the former features can be artificially reproduced.

1.1 The anomaly of the mental

Donald Davidson, in “Mental Events” (Davidson 1970), states *mental anomalism* (sometimes, as for example in Schlosser's contribution to this volume, called “psychological anomalism”) as the principle that “there are no strict deterministic laws on the basis of which mental events can be predicted and explained”⁷; thus mental events are contrasted with non-mental physical events, such as one’s pen falling to the ground, which fall under deterministic laws. If laws could be established which explained and predicted our mental events, on the basis of, say, brain-states, then there would be a reduction of the mental to the physical via psychophysical laws. The problem of the mental would dissolve; we would have reconciled the mental with the natural, and the philosophy of mind could have peace. But the anomaly of the mental is, according to Davidson, an “undeniable fact” (ibid., p. 207), and one that makes us human. It is not that philosophy must remove the anomaly, but rather that philosophy must show how mental anomalism is consistent with the dependence of the mental upon the physical.

Mental anomalism can be seen as an expression of the intuitive thought that our mental lives constitute a realm of rational freedom, a realm incompatible with the determinism of natural law.⁸ Davidson aims to capture that thought, while maintaining that our mental lives are in causal interaction with the physical, and without adopting an implausible metaphysical dualism. According to Davidson, there is only one metaphysical sort of stuff, but it can have two different sorts of properties—physical and mental. His claim is that there can be no reduction of one sort of property to the other: mental properties, which are in the realm of reason, cannot be assimilated within the realm of law.

Davidson points to *normative* constraints on the correct attribution of mental properties, that have no parallel at the physical level. That is, in order for a mental event, enjoyed by a particular subject, to have the property of, say, having a certain intentional content, that event must play a certain *rational* role in the subject's cognitive and behavioural economy.⁹ This role is constitutive of the event's having

⁷ Davidson 1970, p. 208 (page numbers for Davidson refer to the 2nd edition (2000) of his *Essays on Actions and Events*).

⁸ Davidson says he feels “in sympathy with Kant” (ibid., p. 207), where the latter says that the apparent contradiction between freedom and natural necessity must be “eradicated” (quoted in Davidson 1970, p. 207; from *Grundrisse*).

⁹ We say more about this below, §2.1.

that mental property. No such constraints are in force at the level of physical properties: at the physical level, there are no normative or rational relations, only causal and descriptive ones. Thus, there is a mismatch between the constraints on correct attribution of mental properties and the constraints on correct attribution of physical properties.

Davidson sets out to show that there is no contradiction between the following commitments, each of which is individually plausible:

- (a) the principle that “mental events interact causally with physical events” (ibid., p. 208);
- (b) the principle of the nomological character of necessity: “where there is causality, there must be a law: events related as cause and effect fall under strict deterministic laws” (ibid., p. 208);
- (c) mental anomalism.

The contradiction seems to arise from the fact that the first two principles imply that mental events are subject to psychophysical laws, which is what mental anomalism denies.

There are at least four possibilities: one can side with Davidson and attempt to reconcile the three principles, or one can drop any of the three principles in such a way as to avoid the apparent contradiction. Epiphenomenalism, for example, would deny principle (a) (see Jackson 1982 and 1986), and claim that the mental does not interact with the physical. Unlike interactionist dualism, epiphenomenalism can maintain that mental properties are non-physical, without denying that the realm of the physical is causally closed. On the other hand, epiphenomenalism leaves the mental as something explanatorily redundant: it denies that, for example, you cry out *because* you feel pain. This is, surely, highly counterintuitive. On top of this, epiphenomenalism is false to the phenomenology of our introspection, since it is surely part of our phenomenology that our thoughts sometimes bring about our behaviour.¹⁰

Abandoning principle (b) does not seem a promising move: (b) merely expresses an understanding of causality that is common to almost all philosophical views of causality, both realist and anti-realist.

What about (c)? Dropping mental anomalism amounts to claiming that there are psychophysical laws and that therefore the mental can be reduced to the physical. There is a fundamental disagreement between Davidson and those who propose to drop mental anomalism: Davidson thinks that mental anomalism is part of what the philosophy of mind must explain; his opponents think that it is a red herring. In an

¹⁰ This is not supposed to be a grand claim about mental causation, nor a claim about action being necessarily brought about by thoughts. We happen to think (see Di Nucci, unpublished MS) that some actions come about automatically, without any previous thinking on the part of the agent.

important sense, the defender of (c) and its opponent disagree upon the subject-matter rather than its explanation.

Davidson's own solution is a version of the identity theory, "a theory that identifies at least some mental events with physical events" (ibid., p. 209). Mental events that are identical to physical events are such that they have both a mental description (what Brentano 1973 called an *intentional* description, such that the event is *about* the world)¹¹ and a physical description—that is, they have both mental and physical properties. Davidson calls his position "anomalous monism", because it holds that there are no non-physical events, but it denies the existence of psychophysical laws linking mental and physical properties. This is because Davidson accepts identity of particular mental and physical events (event-tokens), but not identity of the mental and physical properties (event-types) possessed by those events; and only the latter would imply the existence of psychophysical laws. There are no laws because laws link properties, and, as we saw, there are normative constraints on correct attribution of mental properties (application of mental descriptions), which have no parallel at the physical level.

Markus E. Schlosser, in his contribution to this volume (Chapter Eight), defends non-reductive physicalism from its best-known challenge, the causal exclusion argument. The exclusion argument is supposed to show that non-reductive physicalism is implausible because it holds on to both mental causation (the thesis that mental events cause physical events) and to the causal closure of the physical (the thesis that every physical event has a sufficient physical cause). According to the argument, the conjunction of these claims is inconsistent with the principle that events are not typically *overdetermined* by their causes—they do not typically have more than one sufficient cause. Schlosser sets out to show that at least three versions of the argument fail. The version according to which mental events are token-identical with physical events fails because it cannot rule out the causal relevance of mental properties. The version according to which mental events are metaphysically included and determined by physical events fails because, if physical events include mental events, then the two cannot be rival causes. Finally, the version according to which mental events are entirely distinct from physical events fails because, mental events and physical events not being independent causes, they cannot overdetermine their effect.

1.2 Consciousness

Even if our experience of the world and ourselves could be predicted and explained through laws, just as the rest of the physical universe is understood through laws, it would still *appear* to us as if our experience were a unique realm.

¹¹ See below, §2.1; and see Searle (1983).

That is, there remains something inherently peculiar (queer) about our experience, even if the strictest version of physicalism turns out to be true (and, as we stand, we don't know whether physicalism is true; let alone which version of it is true). Introspection tells us that our experience is private in a way that nothing else is; and it possesses intrinsic sensory qualities that elude characterisation in physical terms. The colour of your favourite jumper; the smell of a woman's hair; the salty taste of the sea; the feeling of being back home; the sight of the fire on in your grandfather's house.¹² This aspect of experience has been called "what it is like" (Nagel 1974), phenomenal consciousness (Block 1995), qualia (Jackson 1982). The physicalist must understand qualia as physical properties, but every physical description seems to miss something about what an experience feels like to its subject.

This puzzle about experience is illustrated by Jackson (1986) with a famous thought-experiment. Mary, a world-leading colour-vision scientist, not only knows everything that science knows about colour-vision, but knows *all* the physical facts about colour-vision. But Mary has never *seen* colours, because she has spent all her life in a black and white cell. Jackson asks whether, when Mary leaves the cell and sees colours for the first time, she learns a new fact—whether she acquires some information that she didn't have before, information about what it is like to see colours. It seems that, since Mary already knew all the physical facts, if she learns a new fact in leaving the cell, then that must be a non-physical fact. This is the position that Jackson himself endorses in his paper: Mary does learn a new fact in seeing colours for the first time; and since this new fact cannot be physical, Mary must learn a non-physical fact.¹³ Therefore, according to Jackson, there are non-physical facts about phenomenal experience (qualia).

Jackson's paper seems to illustrate that our conception of consciousness allows for room between facts about consciousness and physical facts. The widely shared intuition that Mary learns something new when she sees colours supports the claim that facts about conscious experience are somehow *different*; and, perhaps, the claim that the phenomenal character of experience cannot be captured by anything that the physical sciences could tell us. Philosophers have deployed a plethora of other imaginary scenarios to support such a claim. For example, Chalmers (1996) asks us to imagine a zombie: a molecule-for-molecule replica of you, who lacks conscious experience. Although the zombie is physically just like you, and

¹² Note that these examples have been deliberately chosen to show that the problem of phenomenal experience does not only apply to what are, after Locke, usually called secondary qualities, like colours or smells. It applies to any object of experience; and therefore also to what Locke called primary qualities. The literature on qualia is full of secondary qualities examples, like in Jackson (1986), only because they make the problem more apparent.

¹³ See Jackson (2003) for his recantation of this view.

therefore behaves just like you, it has no qualia. There is nothing it is like to be a zombie. The apparent ease with which we can entertain the supposition of zombies seems to suggest a gap between consciousness and the physical—just as Jackson urges with his case of Mary, the colour-vision scientist.

What is interesting about Mary's thought-experiment is its extremism: it seems to force one either to concede that Mary does discover something new, namely a non-physical fact, or to suppose that, even though Mary has a new experience, she learns nothing new from it. It is not within the scope of this chapter to answer Jackson's challenge; we have presented his thought-experiment only as a very appealing way of introducing the puzzle about qualia. But we will mention David Lewis's response. Lewis (1988) distinguishes between propositional knowledge—knowledge of facts, or knowledge *that* something is the case—and non-propositional knowledge—which includes knowledge of *how* to do things, and knowledge *of* objects, also called “acquaintance”. Lewis claims that Mary, before leaving the cell, has all the propositional knowledge one can possibly have about colours and colour-vision. What she learns in coming out is some non-propositional knowledge, or *know-how*, about colours; she acquires the ability to recognize and distinguish colours. Note that Lewis here is very close to contemporary accounts of phenomenal experience according to which qualia just are, or are the base for (depending on the particular account) the subject's *abilities* to sift, sort, and track objects.¹⁴

Why should Mary learn something new? Because, according to Lewis, the know-how she obtains can be acquired only through direct experience of the object of knowledge (colour); this is unlike propositional knowledge, which can be acquired through, for example, testimony. That does not violate the example's stipulations, because Lewis isn't limiting Mary's knowledge of all the physical facts, which is purely propositional; and at the same time it does account for Mary's learning something new. Lewis thinks that his account vindicates materialism. Whether that's true isn't the point here. What is interesting is that even Lewis has to concede some irreducibility for the character of phenomenal experience: namely, that it can be discovered through experience alone. One cannot explain what it is like to see the sun setting on Edinburgh; or, rather, anyone's attempt to explain what it is like to see the sun setting on Edinburgh will fall short of finding out for oneself what it is like. There is only one way to find out: looking out of the window to see the sun setting behind Appleton Tower.

There is an important lesson to be learned from this. Experiences are, necessarily, someone's experiences. It is their nature to belong to a subject and to be inseparable from that subject in so far as they are that subject's experiences. This is what is often expressed by calling them subjective. Experiences are subjective in

¹⁴ See, in this volume, Ward's discussion of Petit (2003).

the sense that they have to be some subject's experiences. The important point is that it is not required, for some form of physicalism to be true, that experiences can be somehow detached from their subjects. The challenge for physicalism is to provide an account of the character of phenomenal experience that is consistent with our naturalistic framework. A plausible physicalist account will naturally concede that experiences are subjective in the sense, outlined above, that they are essentially some subject's. The sense of subjectivity that physicalism will not and cannot concede is that the character of phenomenal experience is explained by mysterious non-physical entities.

Dave Ward, in his contribution (Chapter Six), argues against the supposed mysterious character of qualia. He distinguishes five intuitively plausible features of qualia which appear problematic for the physicalist, and shows how each of them can be accounted for within a naturalistic framework. The features are intrinsicality, ineffability, privacy, simplicity, and direct availability to consciousness. Ward takes what he calls a "deflationist" approach: although nothing actually has all of these five features, an explanation is available for why it appears to us as if our experience does have these features. In brief, his position is that ineffability, privacy and direct availability are genuine features of experience, and that intrinsicality and simplicity are not. However, both the genuine features and the introspective evidence for intrinsicality and simplicity can be explained by appealing to the fact that the computational mechanisms giving rise to experience are massively complex, and that we have access to the *deliverances* of these mechanisms rather than the mechanisms themselves, since the mechanisms lie at the subpersonal level.

1.3 Mind as artefact

Arguments like Jackson's try to show that the mind cannot be captured fully in physical terms. Proponents of Artificial Intelligence disagree. If the mind can be artificially recreated by man, then it must not be so mysterious. Thus there is a fundamental relationship between AI and the philosophy of mind. If it were possible to recreate the mind artificially, then reductionists would have won the day. It is worth clarifying at the outset that the converse is not immediately true. It might be that reductionists are right even if the mind cannot be artificially recreated by man: there might be some nomological limitations upon man such that we cannot hope to re-create the mind, and this would still be compatible with the mind being part of the physical world. So AI has a particular place in the debate: its success could, theoretically, provide an argument for reductionism, though its failure cannot provide an argument against reductionism.

There is a general objection to the connection between AI and the mind: if the mind's evolutionary history is essential to the mind being what it is—to the mind's

essence—then AI is necessarily cut off from many philosophical questions about the mental. Man cannot replace nature, and that is independent of how similar the artificial mind can be to the natural mind. The artificial mind will always have a different history, and therefore it will never duplicate the natural mind. Searle (1992) seems to hold a similar position when he says that consciousness is a biological phenomenon, and that the mind is essentially organic. So, according to this objection, there is at least one thing that will always distinguish AI and the mind.

There are at least two replies to this way of diminishing the importance of AI. First of all, one could take issue with the idea that the history of something is part of its essence. One could hold that the essence of the mind, and what philosophy of mind is interested in, is what the mind *does*: its functions and phenomenology. And even if that is causally related to the way in which the mind evolved, the history of the mind is not constitutive of what it is and does. Thus, to duplicate the mind it would not be necessary to duplicate its history; it would be sufficient to replicate what it does.

Secondly, one could attack the motivation for the objection. The motivation seems to be the thought that the human mind is part of the natural, biological world, in which we understand the nature of things in terms of their evolutionary history, whereas an artificial mind could not be natural in this sense. But this thought involves an assumption about what counts as natural and what doesn't—about, so to speak, what is the will (or work) of God and what isn't. One could take a Humean line (see Hume 1777), and question this conception of “natural”: why should human artefacts, such as a mind created in a lab, not count as natural? One needs an argument for distinguishing between natural and non-natural things, when they both happen in the natural world.

We can distinguish two different theses: Strong AI and Weak AI. Strong AI is the thesis that the mind *is* a computational system, and that therefore the mind could be artificially recreated. Weak AI is the thesis that even though a computational system could simulate the functions of the mind, a computational system could never be a mind.¹⁵ Therefore, according to Weak AI, the mind cannot be artificially recreated. The distinction between Strong AI and Weak AI gives rise to three different philosophical positions with respect to AI: one can accept Strong AI; one can reject Strong AI but accept Weak AI; or one can reject them both.

The previous historical objection to AI can be mapped onto this distinction between Strong AI and Weak AI. One could think that if, as the objection goes, a mind's history is essential to something being a mind, then the objection applies to

¹⁵ The term “Weak AI” is sometimes used to refer to a weaker view, that is agnostic about Strong AI, rather than a view that rejects Strong AI, as we have characterised it. The stronger characterisation suits for our purposes.

Strong AI, because nothing artificially reproduced by man can share the history of a *natural mind*, and therefore AI cannot create minds simply because it cannot (re)create *natural minds*. On the other hand, though, the historical objection could be powerless with regard to Weak AI, as long as the functional profile of the mind did not depend on its history. Then, the possibility of artificially reproducing something that was functionally identical with the mind would be independent of the mind's history.

But is it coherent to distinguish between Strong AI and Weak AI? Such a distinction seems to hide an assumption about explanation, namely that a model can explain something, by simulating its functions, without being or becoming that thing. If one endorsed the above-mentioned view that the essence of the mind is its functions, then one would think that the only way in which a computational system could properly and completely explain the mind would be by deserving to be called a mind, and so the distinction between Strong AI and Weak AI would not hold. It would not hold because a simulation that did not deserve to be called a mind could, at best, only provide an approximation or metaphor for the mind. But, as a thesis about explanation, Weak AI would not be a coherent alternative to Strong AI.

Richard Hamilton's article (Chapter Seven) proposes, similarly, that Weak AI is no alternative to Strong AI, but he rejects them both. Hamilton presents us with a reading of Heidegger according to which it is a mistake to think that the mind can be artificially reproduced. The reason that it is a mistake, according to Heidegger, is to do with what he calls "reflective self-understanding". Man cannot help understanding itself in light of the world to which it belongs. And, ours being an essentially technological world, man cannot help understanding itself in technological terms; and, in the case of AI, it cannot help understanding the mind in computational terms. Because this reflective self-understanding imposes itself upon the human being, one cannot have any proper theory of the mind—a theory that could be independently addressed and chosen, as required for the assessment of any theory. A technological view of the mind imposes itself at a pre-intellectual level, and prevents us from doing philosophy of mind: thus, no AI thesis can be assessed for truth.

2 Content, perception and concepts

2.1 Mental states and intentionality

We have been discussing the prospects of naturalising the mental. As noted above, the mental bears a peculiarly intimate relation to intentionality. Intentionality is, roughly, the property of being *about* or *directed on* something. The suggestion that intentionality is the "mark of the mental" was explicitly formulated by Brentano (1973; it is defended by Fodor 1987, 2003); although the

notion of intentionality, and the view that it has a fundamental role to play with respect to the mental, go back at least as far as scholastic metaphysics (e.g. Augustine's *On the Trinity*; Augustine 1873).

An intentional state (or event, or item) is a state (event, or item) with intentional content; it is in virtue of having this content that it is about something. Thus, the belief that the earth is flat is an intentional state whose content is the proposition *that the earth is flat*; thus it is *about* the earth (*inter alia*). It is not about the words “the earth”, or some representation of the earth; it concerns that very concrete thing, the planet earth itself. The wish that the earth be flat, and the supposition that it is flat, are states of different types with the same intentional content, and directed on the same object or state of affairs. A pain in the leg, on some views, also has an intentional content: *that my leg is injured*, perhaps. It is directed on one's leg. That is why it is a pain in the leg, not the arm (though the view that pains are intentional is highly controversial; see below, §2.2).

For a state to have intentional content is for it to determine a set of conditions under which it *would* correctly represent the world. Having intentional content does not require that these conditions are satisfied. It was an error of the scholastics and, on some readings, of Brentano, to suppose that the *object* of an intentional state must be internal (“immanent”) to the state, and therefore existent. This is suggested by the notorious phrase “intentional inexistence” (see Brentano 1874, p. 88); the idea is that, since intentionality is directedness, an intentional state embraces the object of this directedness, and so the mere occurrence of the state ensures that there is something on which it is directed.¹⁶ On the contrary, intentional content ought to be understood as a *semantic* notion, related to the notion of *meaning*. A state with intentional content *represents* something, in a sense roughly like that in which an expression in language represents something (which is *not* to say that an intentional state *comes* to represent something in the same way as linguistic items do). The possibility of misrepresentation ensures the following peculiar fact: semantic representation is such that what is represented need not exist. I might believe that the fountain of youth is around the corner, despite the fact that there is no fountain of youth. This modifies the sense in which intentionality is relational, for a state can be intentional even when there is no object or state of affairs to which it is related intentionally—as is the case with my belief about the fountain of youth.

Brentano's claim is that the realm of the mental is the realm of the intentional; the task of accommodating the mind within nature then becomes the task of naturalising intentionality. While this is an important step, it does not obviously

¹⁶ According to the direct realist view of perception, the object of a veridical perceptual experience *is* internal to the experiential state—it is constitutive of the state. Thus, this view denies that perception is intentional. See below, §2.3.

move us any closer to success, for the task of naturalising intentionality is itself a notorious philosophical quagmire. Before going on to discuss the plausibility of Brentano's claim, we will set out very briefly the problems surrounding intentionality. These problems don't depend on one's acceptance of Brentano's claim: even if intentionality is not the mark of the mental, it is still a crucial feature of many mental phenomena—one that does important explanatory work, and thus demands assimilation into our worldview.

The first of these problems, which we can call “the problem of content”, is this: how can there be mental states with intentional content at all? The problem of content is given force by what we take to be involved in having a naturalistic worldview—and by the excellent reasons we have to adopt such a worldview. Mental states are natural states. The problem of content is to explain how a natural state—one composed of the brute materials of an apparently meaningless natural realm—could have the property of being *directed* on, of *referring* to, of *concerning*, some other element of the world.

The difficulty of the problem of content lies in the fact that intentionality is a normative notion. A state's having an intentional content places it squarely in what Sellars (1956) calls “the space of reasons”. Part of what individuates an intentional content is the rational relations in which it stands; being in a state (a belief, say) with a certain content makes it appropriate, to some extent, for one to enter certain states, and avoid certain other ones, individuated by *their* intentional contents. My belief that the earth is flat makes it *pro tanto* rational for me to believe, for example, that there is at least one flat planet, and puts pressure on me to avoid believing that the earth is round. A state's having an intentional content *is* (in part) its standing in normative, rational relations to other such states, episodes and actions. But it is hard to see how the notion that something *ought* to occur, and the notion of something's being a *reason* for something else, can get any grip in the context of naturalism.

The problem of content is given concise expression by Haugeland (1990, pp. 384-5):

“given the vapid materialism now generally conceded (roughly: without any matter, there wouldn't be anything else contingent either), how can it be that any part or feature of the universe represents or is a reason for another? How can there be norms among the atoms in the void? How, in short, is intentionality compatible with materialism?”

What the problem of content amounts to depends on the answer to another question: what *is it* for a state, episode or item to have intentionality? This need not be a conceptual or definitional question; it need not be understood as requiring an analysis of the concept of intentionality. Rather, it can be construed as a question

about a putative natural kind: what are the intrinsic and relational properties in virtue of which something falls under the natural kind *intentional*?

The question of what it is for a state (say) to be intentional breaks into two distinct but interdependent questions, corresponding to two ways of picking out the intentionality of the state. The first is the state's property of representing *anything*—of having an intentional content. This feature is contrasted against the property of being blank, meaningless, or empty of content. The second feature is the state's having the *particular* intentional content that it has. This is contrasted against the property of having some other intentional content. Thus, there arise the following related problems. What is it for state M to have an intentional content at all? And what is it for state M to have the content C, rather than the content C'?

Providing an account of what it is to have an intentional content—in the general or particular sense—may make the problem of content easier. For example, if one holds that a state M has the content C by virtue of an evolutionary function linking states of that type to situations of the type specified in C (Dretske 1995; Millikan 1989), then the problem of content is answered by reference to evolutionary functions. On the other hand, one might hold that intentionality is an intrinsic feature of certain biological systems (Searle 1992); this seems to leave unaddressed the problem of content. A third route is to deny that there is any substantive philosophical problem (e.g. Rorty 1979): intentionality is *not* a natural property, nothing beyond our attributions constitutes M's having the content C, and so the problem of content does not arise. We cannot here argue for a particular view of content, but only show how these problems bear upon the question of how intentionality and mind fit together.

2.2 The role of intentionality in the account of the nature of mind

The prospects of solutions to these general problems determine how much progress we make in our overall task—understanding the mind as a natural phenomenon—by conceiving of the mental in terms of intentionality. Nevertheless, we can consider independently of these problems whether we ought to conceive of the mental in this way. That is, we can ask what the role of intentionality in the account of the mental should be. Someone uncorrupted by philosophy might be tempted to think that it is only once we have done our metaphysics, and explained how mental phenomena arise in a physical world, that we face the task of accounting for the fact that all, or many, phenomena of this type exhibit intentionality. On this view, the mental is a peculiar natural kind, perhaps associated with intentionality, but whose nature is explicable independently of intentionality. Brentano's suggestion may be read as the denial of this view.

Now, one can read Brentano's suggestion in stronger or weaker ways (we leave aside here the question of how Brentano himself intended it). It could be read as a

mere claim about demarcation: all and only mental states exhibit intentionality of a certain sort. Read this way, the claim has no implications for explanatory priority. More strongly, Brentano could be read as making a constitutive claim about mentality: having intentional content is part of what constitutes a state's being mental. Even more strongly, it could be read as a reductive thesis: the concept of the mental is reducible to the concept of intentionality. Other readings are possible: for example, the thesis could be read as proposing a metaphysical, rather than conceptual, reduction.

There are indeed considerations that, *prima facie*, count in favour of accepting a strong version of Brentano's claim: at least the constitutive version. The intentionality of a particular mental state seems to be an essential feature of that state. Take my belief that the earth is flat. It is difficult to make sense of the possibility of the very same belief having some other content. The belief that the sun is flat, and the belief that the earth is round, would be different beliefs. We individuate beliefs by their contents. And it is even more difficult to make sense of the very same state failing to have any content at all—failing to be a representational state. These points hold not only of beliefs, but of other intentional states: of all the propositional attitudes, of occurrent thoughts, of imaginings, and so on.

Now, the belief's being *mental* and its being *intentional* both seem to be essential features of that belief; and that fact does not entail that one of the features is constitutive of the other. But the essential role of intentionality implied by the above remarks entails that we cannot strip intentionality away from belief and be left with something recognisable as mental. There is thus some attraction in the idea that part of what it is for the belief to be a mental state is for it to have the normative, relational property of intentionality. It is not clear that the nature of belief as a mental state could be explained without reference to intentionality.

Someone unsympathetic to the stronger versions of Brentano's thesis might try to point to mental states or events that are not wholly characterised by their intentional contents; if there are such states or events, and if the mental is a natural kind, then the property of being mental cannot be constituted by the property of being intentional. For example, the qualia of certain experiences¹⁷ have been argued to be essentially independent of those experiences' intentional contents. On some views, pains are an example of mental phenomena that involve more than intentionality (e.g. Block 1996; but see Tye 1997, for a rejection of this view). What matters to pains, one might think, is not primarily a semantic content, but their *phenomenology*: it is essential to pains that they *feel* a certain way, and the way they feel (they *hurt*) is not fully characterised by specifying an intentional content such as "my leg is injured" (compare feeling a pain in one's leg to merely

¹⁷ See above, §1.2; and Ward (this volume).

coming to believe of one's anaesthetised leg that it is injured). Another oft-cited example is orgasm: although orgasm is perhaps partly constituted by a representation of ejaculation occurring, the experience seems to have sensational properties that go beyond any such representation.

If this is right, then we cannot use intentionality to establish the unity of the mental as a kind. Pains and orgasms seem to have phenomenology that outruns their intentionality; on some views, they lack intentionality entirely. On the other hand, propositional attitudes have intentionality but, on most views, lack phenomenology: believing that the earth is flat does not seem to have a subjective phenomenal feel that distinguishes it from believing that the earth is round (though Searle 1992 can be read as claiming that propositional attitudes have phenomenology; and see Campbell's chapter, this volume, for remarks on Hume's "belief-feeling"; see, e.g., Dennett 1991, for the rejection of this). What, if anything, might tie these two sorts of phenomena together?

2.3 Perception: intentional content and phenomenology

The task for the defender of the strongest version of Brentano's view is to reduce phenomenology to, or identify it with, intentionality: they must show that the phenomenal character of a mental state or episode is nothing more than its having this or that intentional content. Perhaps this can be done for pains or orgasms (see, e.g., Tye 1997); or perhaps pains and orgasms comprise an isolated category of phenomena that cannot be understood in the same way as other mental states and episodes. The same cannot be said, however, for perceptual experiences—the exemplary case of the interplay of intentionality and phenomenology, since such experiences have phenomenal character and intentional content. Perception is the locus of any number of problems in metaphysics, epistemology and the theory of intentional content. Whatever we say about intentionality and phenomenology in this case is likely to have profound philosophical import.

Unlike pains, it is clear that perceptual experiences are world-directed: they represent, or present, the world as being a certain way to their subjects. They have content, in this undemanding sense at least (we do not mean to imply that it is obvious that perceptual experiences have intentional content in the semantic sense). On the other hand, perceptual experiences, unlike beliefs, are uncontroversially and fundamentally phenomenological. There is something it is like to enjoy a particular conscious perceptual experience, and the nature of the experience cannot be understood independently of this phenomenological aspect. There *might* be something it is like to entertain a certain belief, but it was suggested above that what is essential for understanding a belief is its content, rather than any phenomenology it might have.

The subjective character of perception strongly suggests some sort of dependence, or interdependence, between these two features. To put it roughly, in experience the world is presented to one *in* the phenomenology. Consider a conscious experience of seeing red. Such an experience is a representation, or presentation, of redness—it is directed on a property in the world—*only* in so far as it has the peculiar phenomenal character that we associate with visual experiences of red things.¹⁸ The representational aspect does not seem to go beyond the phenomenal aspect. When one makes judgments (such as the judgment “that is red”) on the basis of such experiences, it seems that it is having an experience with that *phenomenology* that justifies one in making the judgment. And yet, on the other hand, the phenomenology is essentially *world-directed*. The phenomenal character of seeing red does not seem to be anything more than the redness in the world revealing itself to one. Phenomenology and intentionality seem, in such cases, to be inseparable; they are not independent features of experience such that we can abstract from one in order to see the other.

What, then, is the role of intentionality in perceptual experience? According to the *intentional theory of perception* (or *intentionalism*), there is no more to the character of a perceptual experience than its intentionality. That is, the phenomenology of experience—its feeling as it does to its subject—is a representational feature; phenomenal properties are identical to, or reducible to, representational properties (c.f. Byrne 2001, for a recent statement and defence of intentionalism; also Harman 1990; Tye 1995) The intentional theory, confronted with the considerations adduced in the foregoing paragraphs, opts for a reductionist version of Brentano’s thesis, at least as it applies to perceptual states. It claims that the nature of perceptual experience should be explained entirely in terms of the notion of intentionality.

The intentional theory of perception understands “intentionality” in the somewhat demanding semantic sense according to which the intentional content of an experience is understood in terms of the conditions under which the experience would be veridical. The view that perceptual experiences are intentional in this sense commits one to more than the uncontroversial claim that experiences are world-directed. Thus, a substantive intentional theory of perception must be distinguished from the mere claim that the subjective character of a perceptual experience can be fully described in terms of its being directed on certain objects or state of affairs. This latter claim, though entailed by most versions of the substantive intentional theory (Peacocke 1983 is an exception), is not equivalent to it. It is the claim that perceptual experiences lack non-representational qualia.

¹⁸ Perhaps there are unconscious or subconscious perceptual representations that have no phenomenal character for the subject. These remarks are intended only to capture what our experience seems like to us, introspectively.

In its adherence to this demanding sense of intentionality, the intentional theory assimilates perception to beliefs and desires. This suggests that we should understand perceptual experiences as occurrent states or episodes that represent the environment as being a certain way, and are caused (when veridical) by its being that way. This conception of perception is challenged by the so-called enactive view, which conceives of perceptual experience as being a process of *interaction* between perceiver and environment, in which the perceiver *explores* the environment, rather than something which is simply caused by the environment. According to the enactive view, your experience's being of a tree is not simply a matter of the content "tree" being tokened in a mental state which you passively enjoy; rather, it is a matter of how the tree affects your sensory receptors, and how you expect it would do so, over an extended period, in which you might explore the tree in various ways, for example by walking around it (c.f. Noë 2004, for an influential version of the view).

Now, it is a matter of some controversy just what the enactive view of perception is committed to. For it is not clear that intentionalism must deny, for example, that perception is a temporally extended process involving interaction with objects. Julian Kiverstein's contribution to this volume (Chapter Eleven) presents an understanding of the enactive view, on which it has specific and substantive disagreements with intentionalism. According to Kiverstein, the intentional content of an experience is fixed by the subject's sensorimotor knowledge: her grasp of how a scene's phenomenal appearance to her is contingent on her movements. This entails that the phenomenal properties of conscious experiences not only are independent of their representational properties, but explain those experiences having the intentional contents they have. *Pace* intentionalism, intentionality derives from, rather than being the source of, phenomenology.

The emergence of new views of perception, such as the enactive view, has stimulated reflection on the difficulties facing intentionalism—once installed as a near-orthodoxy. Consideration of these difficulties will help to put into context some other competing accounts. Broadly speaking, concern has focused on whether the intentional theory can do justice to the manifest role of perception in our cognitive, rational lives. We can distinguish two sorts of considerations that have led to such concern.

The first consideration is the need to respect the phenomenology of perception. As we noted already, perception is distinct from the propositional attitudes in being associated with a salient phenomenology. The intentional theory purports to bring perception within our understanding by assimilating it in important respects to propositional attitudes: it holds that enjoying a certain sort of perceptual experience is just entertaining a certain intentional content—albeit in a distinctively perceptual way. This would suggest that phenomenology is to be explained by this distinctive

way in which contents are entertained in perception. The resultant worry, then, is twofold. On the one hand, intentionalism owes us a plausible account of this distinctive way that will explain why perceptual experiences have phenomenology—and it is not clear that such an account is in prospect. On the other hand, some philosophers claim that our introspective evidence tells us that the phenomenology of perception is *not* merely a matter of entertaining a certain content in a certain way or mode: the phenomenology tells us that in experience we grasp objects, or their sensible properties, in a distinctively sensory way that is *inconsistent* with intentionality in the semantic sense. This second worry is expressed by Mike Martin with the thought that experiences are partly constituted by their objects:

“the phenomenological character of all perceptual experience requires us to view the transparency and immediacy of perceptual experience as involving actual relations between the subject and the objects of perception and their features. In just the case of veridical perception, the experience is a matter of certain objects being presented as just so, and in virtue of that, the subject ought to conform their beliefs to how things appear.” (Martin 2002, p. 402.)

Why is this inconsistent with the intentional theory? That theory is committed to a distinction between the *content* of a perceptual experience, and the *object* on which the experience is directed. One influential way to understand this distinction is in terms of sense and reference: on this understanding, intentional content is a phenomenon at the level of sense, while the objects of experience are referents. The intentional theory avoids what it perceives as the error of the scholastics¹⁹ by holding that the content of an experience, rather than the object, is a constituent of the experience. The object is external to the experience; and so the very same experience could have occurred in the absence of the object. The worry is that phenomenology tells us otherwise: objects of experience are immediate in the sense of being constituents of experience. The objects of experience are not merely *represented* (as they are, for example, by concepts); they are directly presented *in* experience.

The second sort of consideration against the intentional theory is epistemological. It is a manifest fact that veridical perception is a source of knowledge. Furthermore, it is a source of knowledge about particular objects: in virtue of experiencing a tree, I can come to know about that very tree. My thoughts latch directly onto that particular; I do not need to be able to pick it out by, say, giving a unique description of it. This requires the tree itself to enter into my experience in some robust sense. But, it is claimed, this is inconsistent with intentionalism, for intentionalism claims that the object itself is not a constituent of

¹⁹ See above, §2.1.

the experience, but extrinsic to it. The flavour of the objection is given by McDowell (1994a, p. 193):

“since there is no rationally satisfactory route from experiences, conceived as, in general, less than encounters with objects, glimpses of objective reality, to the epistemic position we are manifestly in, experiences must be intrinsically encounters with objects. But how could they be that if their aetiology were phenomenologically extrinsic?”

Note that this whole objection could be recast in non-epistemic terms. Perception enables us to *think about* particular objects (never mind knowing about them). We must, therefore, have experiences that involve those objects in all their particularity. The worry, then, would be that intentional content cannot achieve this particularity. It is not that states with intentional contents cannot refer to particulars; of course one can have a belief about a particular tree. It is rather that this particularity must have its source in something other than intentional content; the particularity of the content of that belief must be based on an experience that involves that very tree.²⁰

Considerations along these lines have led some philosophers to conclude that the semantic conception of perceptual content is on entirely the wrong track. Thus we have seen renewed interest in direct realist accounts of perception. Direct realism used to be contrasted with indirect realism: the view that in perceptual experience one’s awareness of the external, mind-independent object of experience is mediated by awareness of mind-dependent sense-data (c.f. Russell 1912; Price 1950). But intentionalism is certainly not an indirect realist theory. It holds that one is aware of the object of one’s experience *in virtue* of one’s experience having a certain content, but not that one achieves awareness of the object by being aware *of* the content; just as one can think of an object by employing a certain concept, without thereby thinking *of* the concept. So direct realism, if it is to be in competition with intentionalism, must be saying more than that perceptual awareness of objects is unmediated. The crucial further claim is that a perceptual experience is a state of phenomenological awareness to which mind-independent objects are internal: one’s experiencing a particular object *entails* that that object is before one. One could not have had that experience (an experience of that type) in the absence of that object,²¹ because its presence is partly constitutive of the

²⁰ But see, e.g., Burge (1991), Peacocke (1993), Soteriou (2000), for the view that intentional content in perception can be particular-object-involving in a robust enough sense to meet this objection.

²¹ This is not to say that one could not have had an indistinguishable experience in the absence of the object. Rather, any experience which one could have had in the absence of the