

PHYSICALISM AND CONSCIOUSNESS

BY

JONATHAN LEE GARETH NORTON

M. PHIL PHILOSOPHY THESIS

(UNIVERSITY COLLEGE)

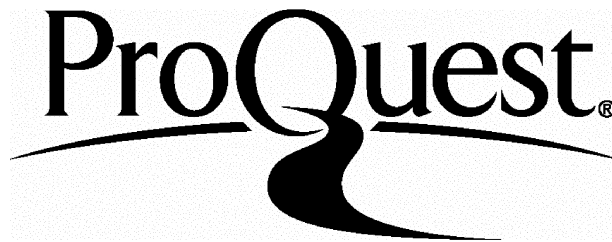
ProQuest Number: U642351

All rights reserved

INFORMATION TO ALL USERS

The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if material had to be removed, a note will indicate the deletion.



ProQuest U642351

Published by ProQuest LLC(2015). Copyright of the Dissertation is held by the Author.

All rights reserved.

This work is protected against unauthorized copying under Title 17, United States Code.
Microform Edition © ProQuest LLC.

ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

Abstract of Thesis

Physicalism is the philosophical theory that identifies mental states with physical states of the brain.

Recently it has come under attack, both from critics who claim that it cannot account for the nature of our conscious experience, and from others who claim that it is a vacuous doctrine. In this thesis I elaborate and defend a version of physicalism against both charges. In the course of this I will make an extensive study of recent work in the philosophy and psychology of perception. I argue that we may accept the critics' claim that we need to acknowledge a special category of experiential fact ('qualia'), but physicalism can accommodate this. The implications of accepting qualia are metaphysical, and I will argue further that the consequences are that we should accept anti-realism.

CONTENTS

Introduction	4
I. The Nature of Physicalism	13
1. What is physicalism?	13
2. What is an adequate philosophical theory of mind?	21
3. Physicalist theories of mind	28
II. Understanding Perception	39
1. The Nature of Perceptual Content	40
2. The Nature of Sensory Concepts	54
3. How we understand qualia	71
4. The projective theory of perception	75
III. The Metaphysics of Qualia	80
Bibliography	86

Introduction

My purpose in this essay is to examine several issues relating to the possibility of giving a physicalist account of the mind. Broadly speaking, physicalism is the claim that we can account for the existence of minds and mental properties in terms of the physical sciences. Of course, this definition encompasses a number of distinct approaches, that address distinct issues in the philosophy of mind in their own particular ways. I believe that physicalism is a coherent answer to what must be the primary question about the mind: what is its ontological status? Furthermore, I believe that if we are clear about the specific questions that physicalist theories have been proposed as solutions for, then we can achieve an overall synthesis which makes good the promise of accounting for mind in physical terms. My own particular concern is with the prospects for a physicalist understanding of our conscious experience, of ourselves and of the world. My conclusion is that such understanding can be had, but it comes with **metaphysical** implications: the doctrine of realism should be rejected. I shall not be arguing instead for idealism or instrumentalism, but I will argue that anti-realism is required to make sense of physicalism.

Following Dummett, I regard the issue between realism and anti-realism to turn on whether we accept that statements may have determinate truth values irrespective of whether it is possible for us to determine that truth value (Dummett 1978 ch.1). How this distinction applies to particular issues depends on how we spell out the requirement for truth values to be open to determination. What matters for the moment is that the realist does not accept such a requirement, whilst the anti-realist does. I shall give my own account of it when I describe my anti-realist physicalism in part III.

I believe that such a metaphysics arises from a problem within physicalist philosophy of mind, namely the need to account for the subjective qualities of experience, or **qualia**. Qualia are the phenomenal aspects of our secondary quality experience, and sensational experience: colours, pains, sounds and so on. I think qualia are an ineliminable part of our conscious experience, and a physicalist theory of mind can accommodate them. But although the existence of qualia may be naturalized and explained in physical terms, their subjective aspect entails that there cannot be determinate truth values for statements about qualia that go beyond our capacity to determine them. This is not to claim that we have infallible knowledge about, e.g. the extensions of colour predicates. We may be wrong in particular colour judgements. But for questions about

qualia themselves, their phenomenal nature could not possibly be other than that which we have knowledge of. It makes no sense to wonder if colours are 'really' quite different from the way in which they appear to us. Or rather: the notion makes sense, but we cannot determine its truth value. If we are realists, then we accept that truth values are determinate even if we never be able to determine them. So for statements about physical properties, there are determinate truths which we may not be capable of knowing - e.g. there are truths about the simultaneous position and momenta of particles, even though we can never ascertain both of them. Qualia are physical properties (so I claim), so equally there are determinate truths about them which we may not be able to ascertain. But qualia are part of our subjective experience, and they have no existence outside it. Either there are no qualia, or they are not physical, or we cannot take a realist metaphysical attitude towards them. I believe that we should accept qualia, regard them as physical, and abandon realism.

I chose the familiar problem of position/momentum measurement above to bring out a consequence of rejecting realism. Metaphysics is our way of summarising the assumptions we make when interpreting our knowledge and theories about the world. We adjust these assumptions to clarify or improve the interpretations. Certain observations suggest that the presence of the observing subject influences the nature of quantum states. Realism holds that these states are determinate in the absence of

the observing subject. So perhaps realism is in error. Abandoning it gives us a clearer, more coherent understanding of quantum phenomena. I am agnostic as to whether this is the appropriate response. However, I think analogous considerations apply in the philosophy of mind. We need to posit subjective facts to complete our theory, but we cannot take the realist line of regarding these as determinate beyond our experience of them. So realism is in error.

I should stress that it is no part of my view to hold that qualia are in any way foundational to our knowledge of the world, in the way that sense data are on some traditional empiricist or phenomenalist views. They are physical properties of physical events in the brain, which is a physical object. The point is that their subjective nature rules out the thought that they may objectively differ from the way they appear to be, and this is a thought we should accept if we take a realist approach to physical facts.

Why should we believe in qualia, and what are they properties of? We should believe in them because they offer some way of understanding our experience of the world. Those of us with unimpaired vision perceive a world of coloured objects, and our other sense modalities present sounds and tastes and smells, and we feel touches and pains. If we concentrate on vision, we must explain what feature of the world gives rise to our colour perceptions. Different objects appear to resemble each other to varying degrees. It turns out that similarly

coloured objects need not have any particular chemical or physical properties in common. Although certain combinations of light at specific wavelengths reliably produce similar colour effects, different combinations may produce the same effect. Thus it seems that colour experience is grounded in the peculiarities of the observing subject's visual system, for if the only property red objects have in common is that we see them as red, then redness must be a property of the perceptual states we are in when we see them. As I think we should accept physicalism, I say we should say that perceptual states are (physical) brain states (or rather, states of the brain and the optical system) and that qualia are the properties of these states that give them the subjective phenomenal nature that they have in common. I give my complete argument for the existence of qualia in part II.

I believe we should posit qualia as properties of brain states in order to account for the facts of sensory perception. But in addition to this ontological claim I require a theory of how qualia play a role in thought and experience, and how these conscious episodes are grounded in neural activity. This general issue is dealt with by an overall physicalist theory of mind. Such a theory accounts for the wider phenomenon of mental **content**, of which perceptual and sensory content are subsets. Contents are representational states of the mind, directed at the world or the mind itself. Such states may represent their objects incorrectly, e.g. by

hallucinations, or sensory illusions, such as when we are deceived about relative temperature or textural properties. A theory of content must show how content can arise and be sustained in a physical system like the brain, and how it can have the structural, conceptual relations that we consider in thought. Whether perceptual content has conceptual structure, like non-perceptual content, is an issue that must be dealt with. Some writers have believed that we should characterise perceptual content as non-conceptual, to capture an apparently significant difference between perceptual and non-perceptual states: perceptions arise in us passively, and are not subject to revision in the way that our belief states are (e.g. Crane 1992). Furthermore, we may consciously and coherently hold beliefs that represent the world differently from our perception of it, as in cases of optical illusion, yet we cannot hold conflicting beliefs in such a way. So perceptual states are different in kind from non-perceptual belief states.

I reject this view. I argue in part II that we should regard perceptual content as conceptual, as non-perceptual content is, for our understanding of perceptual states does involve the use of concepts that apply to the particular structure of perceptual states. The qualitative nature of perceptual and sensory states is complex, and different aspects may be abstracted from it. For example, with colour perceptions, concepts of hue and brightness distinguish different comparative relations between distinct perceptions. The relation between perceptual and

non-perceptual contents may be explained without drawing a distinction in the nature of the contents involved, but rather in a difference in the particular concepts applied in understanding them.

Perceptual contents represent how the world seems to be, and perceptual concepts are the possible aspects of possible objects we may be presented with. My account also requires a theory of how perceptual contents relate to the world, and of how it is that we can discuss and describe our perceptions in terms of a shared range of concepts. If colour experience is subjective, how is it that we can compare our colour judgements, and have a developed science of colour vision? These are further issues I will address in part II. But they connect with a line of criticism that has been given against the view of qualia I am arguing for, and which would disarm my wider conclusions.

When I say that qualia have a subjective aspect, I mean that to be in a state that involves a particular quale necessarily involves a particular phenomenal character - the **redness** of red, the painlike quality of pains, and so on. I hold that having such qualitative contents is to be in conceptually structured states, so the corresponding phenomenal concepts involve these phenomenal characteristics. What these characteristics happen to be forms a category of subjective facts. These are physical facts, but a different kind of physical fact compared to facts about the structural and functional nature of the physical states that their tokens are

instantiated by. Being able to apprehend and grasp these facts is an essential part of being able to exercise the relevant perceptual concepts, e.g. colour concepts.

However, if it can be consistently denied that full possession of colour concepts requires the ability to identify instances of colour by means of their phenomenal characters, then this would cast doubt on the idea that a particular category of subjective facts exists and plays a role in our perceptual concepts. Such a line is taken by Jane Levin (Lycan 1991 pgs.478-489). The motivation is to rebut a particular objection to functionalist accounts of sensory and perceptual experience: that they fail to accommodate the phenomenal nature of such experience. I will explore this issue in part I. My view is that we can acknowledge subjective facts **and** have a theory of the mind that is broadly functionalist. I will show how such a theory can be developed in part I, while defending the claim that subjective facts exist and play a role in our grasp of perceptual concepts in part II.

There is one other issue I will be addressing, and it will run throughout this essay: the relation between content (characterized at the sub-personal, cognitive science level) and conscious thought. Conscious thought involves concepts, and on my view perceptual and sensory experience does as well. Conceptual relations exist between conscious thoughts because of inferential relations between the belief states that they are, and this must be grounded in some way in the existence of sub-personal content and the relations and behaviour of

its states. There are several possible theories that propose how this grounding relation can be understood. As my purpose in this essay is to argue that physicalism **can** give a theory of conscious experience, and that this theory requires an anti-realist metaphysics, I will not pursue the specific issue of **how** conceptual thought is based in (non-conceptual) sub-personal content. It is sufficient to show that physicalism can account for the existence of the latter category of content, and to describe the ways in which such content can give rise to conscious thought, irrespective of how it is realised. This is quite a substantial claim in itself.

In this Introduction, I have described the issues I will be dealing with and sketched the lines of argument I will put forward. The first set of problems are to make clear what physicalism is, what specific physicalist approaches are attempting to achieve, and how I think a unified physicalist approach can be created out of the different elements already in the field.

I. The Nature of Physicalism

In this section I will try to define what I think the key commitments of the physicalist approach are, and judge the various proposals made for a physicalist theory of the mind. In the course of this, I will review the sense data theory of perception. Sense data theory is rightly regarded as a poor theory of perception and its relation to consciousness. By re-examining it, I hope to become clear about what an adequate philosophical theory of perception and consciousness must achieve.

1. What is Physicalism?

In the Introduction, I briefly described physicalism as the claim that the existence of minds and mental properties can be accounted for in terms of physical science. Such a definition leaves its content under-determined, to say the least. 'Physical science' can mean many things. If we stick with this definition, we could say that Aristotle was a physicalist - after all, he had a conception of the position of humans in the natural order. On his terms, the existence of mind is explicable as other aspects of the natural order are explicable: as exemplifying essence or purpose. But this will not do. When we use 'physicalism' we mean something like our present day physics is involved.

However, we cannot just define physicalism as 'explanation in terms of present day physics'. The reason that philosophy of mind is a live subject is that

the material to be explained is so wide and varied that it is open to debate how explanation may proceed, and indeed whether our current science has the resources to complete it. Our knowledge of the brain's structure and functioning advances all the time. The philosophical question is how we relate our knowledge of brain and nervous structure to mental phenomena. The physicalist claim is that we identify mental phenomena with (some particular level of) physical phenomena.

Yet this still takes us no further forward unless we know what the 'physical' is. The apparent failure of attempts to specify the physical/non-physical **divide** has lead some to doubt that a non-vacuous definition of physicalism can be given (see Mellor & Crane, in Mellor 1991 ch.5). I disagree. Physicalism is at the very least the denial of dualism or epiphenomenalism. The dualist posits mental substance to distinguish mind as a different category of thing from the body, and one whose existence and nature cannot be **explained** in terms of the latter's properties. Claiming that mental properties arise **in virtue of** certain physical facts obtaining is substantially different from the straight denial of it, which is what I take dualism to be. Equally, it is different from the epiphenomenal claim that mental properties are of events with physical causes, but are not identical with physical events. I think a case can be made from these distinctions.

I take science to be an essentially **explanatory** exercise, but I do not intend to suggest that the theoretical vocabulary has purely heuristic or instrumental significance. Scientists posit structures and properties because the evidence justifies the belief that reality contains such entities, and that these entities play a role in bringing about the studied phenomena. Our improved theoretical descriptions are improvements on the old ones because they explain things better.

I believe that the key element in our current physical science is not the particular ontology it currently posits, but the pattern of explanation in which this ontology is employed. Different areas of science describe the world in differing levels of detail: biology (as distinct from biochemistry) deals with the level of cells and organisms, chemistry deals with the level of atoms and molecules, quantum physics and particle physics deal with the sub-atomic level. What gives science its unity is that, at each level, the existence of what are, comparatively, macro-properties, may be explained in terms of micro-properties that are of their components, and the structural relations between these components. The properties of these components are then explained at the lower level in terms of some finer degree of resolution at a level below that.

The relation of chemistry to physics illustrates this point. Given the existence of atomic components such as electrons and nuclei, with their properties of charge, energy, momentum and spin, we may explain

why a particular number of electrons with a particular nucleus and energy state will adopt a particular atomic structure: the structure that matches a wave mechanical solution to the quantum physical laws that govern these particles. Given the existence of atoms and ions and their structures, we can account for the bonding and reactive properties out of which molecules form, what structure they assume, and their tendency to react or decompose. Thus the existence of chemical properties and laws may be entirely explained in physical terms; this is how physicalism unifies the two sciences.

It could now be objected that physics itself is not unified: there is the famous conflict between relativity and quantum theory. All I can say on this is that, for my version of physicalism to be upheld, both theories must be reconciled in terms of a lower level, higher resolution theory that explains both of them, and I take it that it is an open question whether this is possible. Many physicists seem to think it can be done. If it should happen that we decide it **can't** be done, then I accept that my model of physicalist explanation will have failed, for reasons of physics. But I am putting the model forward as an account of the **principles** at work in the physical sciences, and if the sciences reject the possibility of unification my model will be redundant because physics itself will have changed course.

The chief problem with formulating physicalism is to avoid the vacuous relativity of tying it to 'present or future (completed) science'. The content of scientific

theory may change, and it may never reach a state of completion - or if it did, it is not clear to me that we should be able to identify that it had. I think some kinds of theory change are benign, however, and I take it to be an advantage of my view that it can accommodate them. I do not claim there is a basic ontology in which physical theories must be stated; we may discover ever finer levels of structure to the universe, and there may not be any foundational level of basic objects and properties. What counts for physicalism is unification in terms of explanation. We can see that this process has occurred in 20th Century science with the rise of biochemistry, allowing a physicalist account of biological processes. Pure biology - the subject that Darwin and Mendel worked in - may be explained in terms of the genetic basis of the characteristics of organisms. Transmission and manifestation of genetic material is explained in terms of the biochemistry of the cell, which may be explained in terms of the chemical properties of DNA and other molecules.

For a definition of the physical itself, I think we can do no better than: concrete objects and their properties. This excludes numbers, as well as epiphenomena - at least, the sort that mental epiphenomenalists have supposed to be involved in consciousness. If we accept nominalism we would reject numbers and do without properties, but this is a separate issue. The relativity of this definition of the physical comes from the relativity present in the notion of a concrete object:

we take the human body to be one, but it is highly diffuse compared to the atomic nucleus. A physicalist theory of mind places mind in the explanatory hierarchy of physical theory, explaining its nature in terms of the physical and structural properties of the nervous system.

The divide between the physical and human sciences lies in their different styles of explanation. Psychology itself, on my view, appears to straddle the divide. If physicalism is true then the existence of psychological states and properties is a physical matter and can at least be influenced to some extent by physical interventions; but this is something that we already know: anaesthesia and clinical psychiatry depend upon it. Non-clinical areas of psychology, such as developmental (child) psychology and mass psychology do not appeal to physical properties primarily, but environmental or even historical factors which may themselves be psychological. Yet the fact that psychological properties have a physical basis does not impugn the validity of this approach, any more than its lack of attention to specifically physical facts refutes physicalism. Their relation to physical science is rather like that of ecology to biology and meteorology. Biological and climatic facts certainly cause the effects that the ecologist studies. But the processes and facts she is concerned with are of a different, more general order, and specific physical predictions are not practically available (due to the complexity and open-ended nature of the physical situations under study) and would not be

to the point even if they were. If we want to understand how climate change will affect how the population of a particular forest will alter over the course of a decade (a perfectly worthwhile scientific question) then an analysis of temperature and humidity changes tells us nothing unless we have some idea of the likelihood that certain organisms will die out, will migrate, will arrive, will discover new food sources. We could not hope to give detailed mathematical analysis and predictions of such changes, yet we can still see ourselves in making serious predictions of how they will occur overall and in general.

Ecology is not a human science, but its problems are analogous to those investigated in economics, sociology, and other human sciences concerned with large scale human behaviour. We can still regard them as worthwhile, scientific endeavours despite lacking the predictive value of the physical sciences. So now we can also dismiss one bad argument used for physicalism: that the lack of well-established laws in the human sciences (in supposed contrast to the physical sciences) shows that psychology should give way to physics in the explanation of human behaviour. Firstly, it does not follow from the truth of physicalism that we should reject the theory of intentional psychology (it could be that we can explain the latter in terms of the former). If we did choose to reject psychology, as many physicalists happen to think that we should, it would have to be on the basis of separate arguments (as in e.g. Stich 1983).

Secondly, we can appreciate the lack of firm laws in the human sciences in terms of the particular systems - large and diverse human groups and activities, generally - that they attempt to study. People change their beliefs and plans in response to material factors and the behaviour of others, and they cause others to do so. Such actions are influenced to varying degrees by the amount of information available to participants in the system, and their interpretation of it - two variables difficult to quantify and assess. Behaviour may also be partially governed by expectations, tradition, past experience to potentially unlimited levels of generality. None of this precludes the possibility that behaviour is governed by **some** strict or at least probabilistic laws, but it does explain the limited scope and precision of such models as we can construct.

The structure of human situations place similar limitations on the application of social scientific or economic theory. Following MacIntyre (MacIntyre 1981 ch.8), I believe that human social life is subject to systematic unpredictability that precludes the human sciences offering predictive laws with accuracy and universality of the physical sciences. But equally, this is only a reason to reject human sciences if we expect them to be in the business of giving strict, quantitative predictive laws - like MacIntyre, I think that the explanatory study of human behaviour can be conducted without trying to fit this particular model. Thus I accept Crane and Mellor's view that we should not repudiate the human

sciences for failing to meet this model (Mellor 1991 ch.5), but I reject their wider conclusion that physicalism is vacuous as a philosophical theory about the mind.

2. What is an adequate philosophical theory of mind?

Before considering the options available in developing a specifically physicalist theory of mind, it is worth getting clear about our objectives in constructing a philosophical theory. Aside from the problems of psychology and neurophysiology, what are the questions that philosophers need to answer, and why do they need answering? There are two obvious duties for a theory of mind: to tell us how the mind relates to the world of matter, and how it relates to itself. Rather than discussing these requirements in the abstract, it would be more interesting to look first at a **bad** theory of mind, one that has fallen by the philosophical wayside, and see what morals we can draw from it. The theory in question will be the Sense-Data Theory (hereafter SDT).

I realise that SDT as such, being a theory of perception, is more a common element in a range of theories rather than a particular philosophy of mind itself. Sense data occur in G.E.Moore's Indirect Realism as well as Ayer's Phenomenalism - two theories with quite different views of the external world, and consequently of the overall status of mind within it. However, I think it is right to gloss these differences, since what is at stake is the mind **itself**, and it is in the nature of these theories that the mind and the world can be bracketed

and treated separately from each other. In any case, I would contend that adopting the SDT as a model of perception commits one to a particular view of mental events, namely one that regards them in terms of mental **objects** as distinct from physical objects. The problems involved in adopting this viewpoint can be brought out by examining SDT itself.

SDT holds that the mind is related to the world by receiving sensory items - sense data - that represent it. These are automatically apprehended by the mind on account of their nature as mental items, and they have a similar ontological status as other mental contents such as memories, beliefs, and so on. The mind is the totality of such contents, and we understand them because they possess their representational properties intrinsically. So understanding a mental item is an act of 'grasping' it - making mental contact with it as a meaningful, representational item.

Two obvious points can be made about the theory's inadequacy. Firstly, the introduction of sense data explains nothing about the mind's perceptual link to the world. Saying that sense data are apprehended when they are present to the mind says no more than that perception is achieved by an act of perception: we have no more of a hold on the process than we did before. So SDT works by an implicit appeal to the mental processes of perception and understanding that it purports to explain. Secondly, it fails to say anything about what it is for representational items to exist **at all** - the

whole business of mental representation remains as opaque as ever.

There are two more subtle problems lurking. If our awareness of the world is mediated by mental items, how is it that we achieve agreement in our judgements about the objective external world that we cannot apprehend directly? Also, how can we be certain that we are continuing to register the same patterns of similarity in our sense-data, if there is no way either to compare them to some standard or to appeal to someone else's authority? These are the familiar difficulties addressed in discussions of the Private Language Argument (see e.g. Pears 1987, 1988 chs. 11-15). This essay is not concerned with the possibility of private language. However, the problems connected with conceiving of mental facts as private affect any attempt to understand perception.

Any theory that posits private facts must explain how they can come to play any role in our vocabulary of mental states. For a fact to be 'private' in any interesting sense, we must mean that knowledge of it is only possible for the subject possessing it. SDT falls into this category: I am aware of the world through my sense-data, and I am aware that others seem to be aware of the world, but I am not aware of **their** sense-data. So I have no grounds for comparing our sense-data. Yet it is certainly the case that we compare our perceptual judgements, and we can also ascertain whether or not someone's perception is systematically deficient, due

to colour blindness or astigmatism. Either such claims are illegitimate, or sense-data do not play the role that the theorist gives them. If they do not play that role, then we have no use for them : 'If you say he sees a private picture before him...you have still made an assumption about what he has before him...If you admit that you haven't any notion what kind of thing it might be that he has before him - then what leads you into saying...that he has something before him?' (Wittgenstein 1953, Remark 294).

I will return to the issue of privacy in part II, when I offer my own account of perception. I think we can use the notion of subjective properties whilst avoiding the difficulties that SDT falls foul of.

Even if we accept SDT on its own terms, it gives us no clue as to how we should relate it to the physical facts that are obviously involved in perceptual and sensory processes - the transmission of signals through the optic nerves, events in the sensory cortex - or how to understand **where**, in what sort of **medium** we are to posit sense-data as existing. Having traced the causal story of perception from its object to the brain, to halt now at a sense-datum is to abandon the trail rather than completing it.

One apparent advantage of sense-data is that they at least preserve the phenomenology of perception. We experience various coloured objects in spatial relations (usually); the sense-datum maintains this structure at the intermediate level by its nature as a mental image

composed of sensory qualities in a mosaic-like pattern. This structure can accommodate hallucinations and after-images: they are sensory fields that fail to have a genuine object at their cause. But this advantage does not survive inspection. We do **not** infer the existence of external objects from a field of colour, not even sub-consciously. In any case, clinical evidence fatally damages the model. The phenomenon of 'blindsight' may be studied behaviourally: a subject suffering from neurological damage avows partial or total lack of visual sensations, yet seems capable of making 'guesses' about her environment which have greater accuracy than we would expect from chance. I will look at the topic of neurological disorders, and their significance for models of perception, in the next section. What matters for the moment is that blindsight cases show that a subject may have some perceptual awareness of her environment in the absence of sensory experience; yet SDT makes the latter into the basis of the former.

The failure of SDT yields three lessons for any theory of perception, and consequently for the theory of mind underwriting it. Firstly and obviously, the existence of blindsight and related disorders challenge the approach of basing perceptual judgements on awareness of a particular aspect of the world. The SDT bases our awareness of objects **qua** objects on our sensory awareness of their secondary properties, in the visual field. Yet the existence of perceptual disorders show that perception may occur, admittedly in a degraded

way, when sensory awareness fails. So we should regard our normal perceptions as the result as several capacities working **together** rather than in sequence.

Secondly, a theory of perception must explain how it is possible for us to have the capacities to discuss and compare perceptual judgements in the way that we do. Perceptual judgements cannot rest on what is ultimately private knowledge, there must be some way in which we can justifiably account for it in publically accessible ways. But this does not preclude there being **some** personal idiosyncrasies in perceptual experience. We mark the existence of these by distinguishing partial or complete colour blindness, and other disorders. Yet the nature of our perceptions must be in some way available for investigation by others, for otherwise we could have no grounds for making such distinctions. SDT leaves it a complete mystery how we obtain such grounds.

Thirdly, a theory must not only settle the ontological issue of mind, but give a plausible account of how it relates to physical objects. An SDT-supporter might think she could avoid the problem of interaction by conceding: 'Of course, sense-data are **physically realised** by states of the visual cortex'. Would this save the theory? It would at least make it an empirical question whether it were true. For now the contention is that there are physical, neural states which in some **structural** way embody the nature of sense-data. For the sense-datum model depicts representation in terms of a

field of sensory qualities from which judgements about objects are deduced. So a physicalist SDT theorist would have to hold that the visual cortex contained some structure realising a pattern of sensory values, causally mediating the optic nerves from the 'interpretive mechanism' (and parallel structures for the other sensory modalities). Note that we could not take the pattern of light across the retinas as physically realising the sensory pattern, for the former is discontinuous across the retinal blind spots, yet the apparent sensory field is (in ordinary perceivers) continuous. Although our knowledge of neurophysiology is incomplete, we know enough to reject this proposed model. Without some sort of structural isomorphism between our physical model of the brain and the SDT model, I think it makes no sense to claim that the former physically realises the latter. Thus sense-data have no meaningful role in a physicalist model of the brain.

This conclusion does not rule out the notion of sensory states per se. We have rejected the possibility that physicalism could accept sensory states or items **of the sort that SDT trades in**, but this does not mean that we cannot say that (some) neural states are identical with mental states, or at least aspects of some mental states, that possess a sensory aspect - an aspect of our conscious life distinguished by its **qualitative** nature, such as our awareness of a region of colour. Indeed, the possibility of identifying qualitative states with brain states was one of the original attractions of physicalist

Identity Theory, as originally proposed. I will examine the different forms of physicalist theory in the next section. For the moment, I note that the rejection of SDT is not the rejection of sensory states or qualities, though some writers seem to regard the positions as identical, and argue as though the notion of qualitative properties must fail for the same reasons as SDT (see e.g. Dennett in Lycan 1990 pgs.519-545).

3. Physicalist Theories of Mind

Before we begin to offer any theory of mind, there are two further issues that must be considered. They affect the scope and shape of our putative subject matter. Firstly, we should ask whether we are offering a theory of mind **in general** - as it may be realised in different species and structures, including the human nervous system - or the human mind **in particular**. Our attitude on this point determines how we judge the issue between functionalism and type-identity theories, as I will argue shortly. I think this conflict can in fact be dissolved, and we may have a unified physicalist model combining the advantages of both approaches.

The second and more vital issue is our attitude to commonsense psychology (or 'folk psychology'). In our everyday use of mental language, we attribute beliefs and desires and sensations, and cite them in explanations of behaviour. We also talk of concepts and inclinations when we wish to generalise over these states. In short, it seems we already have a theory of mental states and

their link to overt behaviour, employing its own categories and vocabulary. But what is the status of this theory? We could regard it as a true account of mental activity, and leave it as a task for physicalism to show how these states may be physically grounded in the nervous system, and how they manifest their typical effects. Alternatively, we could deny that the model truly represents the mind: there are no states with the properties that beliefs and intentions and so on are claimed to have. In which case we should decline the task of relating it to our physical theory of the brain.

Such a position is adopted by many theorists who expect cognitive science to supercede previous, psychological approaches to understanding the mind. It is best referred to as Intentional Anti-Realism: the rejection of the intentional idioms in which everyday psychology expresses itself, and which depend on a notion of mental states as **representational**. This is not to deny that mental activity involves states that vary in ways dependent on external input, and in some way cause information about that input to be embodied (e.g. as in some computational theories of the structure of visual perception - see Dennett 1991 ch.11). But it does deny that there are mental states in representational relations to states of affairs in the world (or, in the case of desires, to possible states of affairs).

Different writers offer differing reasons for rejecting intentional states (for a selection, see Stich 1978, Dennett 1978 part I), and ~~give~~ differing

conclusions about the status of intentional language. Dennett argues that we should continue to use belief and desire ascriptions because of their predictive utility: they are generally true. However he rejects the ontological claim that there are distinct brain states corresponding to true ascriptions, and thus recommends an instrumentalist interpretation of intentional theory. Stich, in contrast, attacks the view that belief ascriptions correctly represent the causal behaviour of the brain (viewed as a cognitive system), and further doubts if any place can be found for intentional states within the models of brain structure and behaviour currently emerging in neural and cognitive science.

This is a large issue. My concern in this essay is with the relation between physicalism and conscious experience. I think the two questions are independent: at the level of **experience**, what counts is that intentional states certainly seem to exist, or at any rate we can proceed as though they do (as Dennett acknowledges). So we are justified in taking their existence as part of the datum to be understood. I will argue later that even if we were to abandon hope of grounding intentional psychology in cognitive science, it would make no great structural difference to the view of the mind I am drawing in this essay; my main argument is for the existence of states with **qualitative** content which may be present in the sort of functional framework that cognitive science posits. However I think there are

two reasons for believing that intentional states exist and can be accommodated by cognitive science.

Firstly, the success of intentional psychology in predicting behaviour is good reason for realism about its theoretical ontology. This is of course a general argument for preferring realism to instrumentalism, but it is as valid in this case as any other. If we are concerned to make all our theories compatible with science, why not accept a **global** instrumentalism about all theories? That would exorcise the problem of incompatibility. But we would not accept it if we believe in the realist model of science as a descriptive exercise, cataloguing the constituents of reality (which I accept as well - anti-realism only rejects **undecidable** questions). So the burden of proof is on those who claim that intentional states cannot be placed in the structure of cognitive science.

Stich has offered (Stich 1978) an argument on grounds of **modularity** to the effect that the two ontologies are incommensurable. The failure of these considerations constitutes my second reason for intentional realism. The claim is that, if states such as beliefs exist, they should be identical with 'naturally isolable' parts of the neural structure of the brain - hence the claim that they must be **modular**. Yet neuroscience suggests that the brain is functionally structured such that no particular regions can be identified with particular beliefs; thus modularity fails in practice.

Following Horgan and Woodward (1985), I can see no

principled reason for accepting the modularity constraint. Certainly, the physical states that belief-states are identical with may be highly complex, diffuse structures, but this is hardly a good scientific reason for denying their reality: neurons and other biological structures are equally complex and diffuse from the standpoint of atomic physics. As I argued in my discussion of physicalism, what counts is the possibility of explaining the behaviour and properties of higher-level states in terms of lower-level ones, and there is no reason for supposing that this should not apply to belief-states as well.

So I believe we are justified in accepting intentional states and expecting physicalism to locate them in the structure of the brain. Exactly **how** they are realised is not an issue I will pursue in this essay. It may be in terms of syntactic items belonging to a 'language of thought' (Fodor 1975,1981), or the working of Connectionist networks (Churchland 1986), or perhaps even some hybrid or alternative theory. In this essay I am mainly concerned with phenomenological issues. What I am interested in is whether the physicalist framework can offer an explanation for all the facts of experience. Henceforward, I shall take it for granted that intentional psychology may be construed realistically, and that its use will be under-written by some physicalist theory of content being true. The question now is whether we can fit the notion of **qualitative** states into physicalism.

By a 'qualitative state', I mean some mental state exemplifying a phenomenal characteristic - a sensation of pain, a perception of a green patch, a sound, and so on. Such qualitative properties are **qualia**, and they are properties of states with qualitative content - some content which represents the existence of a phenomenal property such as a colour being instantiated in the subject's environment. A qualitative state may be present even if the quality it presents is absent from the environment - in hallucination, we may perceive that a green patch is present when there is no corresponding green object. For the moment, I want to argue that our best physicalist models **can** accommodate qualitative states. In part II I will argue that we need to posit qualia, and thus we **should** posit qualitative states.

In holding that there are such states, I am not claiming that they are structured in any meaningful way as 'images', in the way that sense-data are clearly image-like structures of phenomenal qualities. Maybe there are neural structures that can be best understood as functionally realising capacities for mental imagery (Block 1983); I have nothing to say about this. What I contend is that there are states with qualitative content, and that they play a role in perception. **How** they contribute their qualitative content is a separate, though related question.

I said earlier that we should ask whether a theory addresses mind in general, or the human mind in particular. The question matters because we often attribute mental states to other animals; probably not snails or spiders

but certainly monkeys and other primates, as well as cats and dogs. Perhaps if we came to regard whales and dolphins as having languages, we would regard them as thinking creatures also. We attribute minds - and, consequently, beliefs and desires - because these creatures seem to have the capacity to reflect and judge aspects of their surroundings (e.g. whether a predator is present, whether to fight or flee, whether something out of the ordinary is going on). The capacity to notice regularities and to act on them, in various ways, suggests that they possess belief states, and also desire states that vary with time.

The existence of these other minds poses a problem for our approach to physicalism. There are three ways in which we could state the identity of mental and physical states: (i) token-token identity, (ii) type-type identity, and (iii) functionalist identity theory. Of these, I think that (i) can be rejected as immediately inadequate. We should accept at least **some** psychophysical laws as applying: those that relate the effect of anaesthetics and painkillers on pain sensations. Yet if we accept token identity theory, we hold that pain sensations are identical with some neural state tokens, but not a particular state **type**. But this would make it entirely fortuitous that use of anaesthetics was lawlike: it can cause only a specific range of chemical effects in the brain, involving certain neural types, and if these types are not either identical with sensation types (or in another lawlike relation to those that are), then the pain-relieving effect is purely

accidental. On the reasonable assumption that there are at least some true psychophysical laws, we should therefore reject token identity theory as being explanatorially inadequate. It is worth noting that the main current version of the theory - 'Anomalous Monism' - is premised on the **denial** of any strict psychophysical laws obtaining (Davidson 1970).

So the options seem to be: either identify mental state types with physical (neural) state types, or identify mental states by functional descriptions, and identify these with physical state types meeting similar descriptions. Functional descriptions characterise states in terms of their causal relations to other states and to environmental influences. A functionalist model of perception would show how states are affected by sensory input, and how perceptual content is determined by this input and internal cognitive functions.

If we adopt a type identity theory, we have a model that can accommodate human mental behaviour, including qualitative states. Qualitative states will happen to be certain neural state types, and so their nature will be a matter of structure. But this precludes the possibility of generalising to other mental systems. We attribute pain or visual sensations to other creatures, regardless of whether they are neurologically identical with us. Lower animals such as cats and dogs may have pain-realising states of a quite different nature to ours.

So we may reply that what matters in the identity of mental states are the **functional** relations. Lower animals

do not embody the whole range of human psychology, but we may apply some human psychological terms to them (belief, sensation..) because there are structural similarities between parts of the functional theories realised in the different species.

Accepting functionalism now confronts us with the many objections cited in relation to 'qualia' (see e.g. Block 1978, Hill 1991 ch.3). Intuitively, we understand the notion of a qualitative state as a state exemplifying a certain phenomenal character. Functionalism individuates mental states according to their functional descriptions. But it is conceivable that there could be beings obeying identical functional descriptions as ourselves, but either lacking or having different qualia. So functionalism cannot be the whole story.

We now seem to have reached a dilemma. If we choose type identity theory, then we may have to **deny** true mental state attributions to different creatures on grounds of neurological difference. If we choose functionalism, we have to **accept** false attributions to creatures on grounds of functional identity.

The dilemma is false. What we want is some way to preserve the specificity of type identity theory, whilst also obtaining the pluralism allowed by functionalism. Such an approach is available in the theory of **homuncular functionalism** (Lycan 1988 ch.5).

Homuncular functionalism takes the basic functional model but adapts it by introducing the notion of a hierarchy of functional complexity. If we restrict

ourselves to organisms, we can regard physico-chemical processes as the lowest level of functional organisation; basic biological processes supervene on them and thus instantiate biological functions. As we refine the complexity of biological functions, we reach the level of consciousness - at the top of the hierarchy - which supervenes on the overall brain and nervous structure and is instantiated in the different neural systems that realise mental functions and items.

Functional properties increase in generality as we ascend the hierarchy. At the lower end we have purely biological support systems that maintain the organism in a viable, life-supporting state; at the higher end are the mental systems that support representational and goal-directed activity. If we characterise each level in teleological terms, then the existence of specifically mental layers can be naturalised by explaining their emergence by appeal to evolutionary adaptation. The theory is 'homuncular' because, with respect to a certain level, the structures that realise its functions may be regarded as **homunculi** - little workers that perform various tasks. These workers can be regarded as 'empty-headed' (hence homunculi), for whatever functional organisation **they** have does not play any defining role in giving the functional structure of the level that supervenes on them.

The theory reconciles our intuitions in two ways. Firstly, we may use belief/desire terms to describe different organisms because they generally share the

functional descriptions that pick out particular beliefs and desires, irrespective of the details of physical realisation. We may also extend psychological terms to non-living things such as computers, for we can interpret the teleological aspect of the theory in terms of design. Secondly, we may acknowledge differences in qualia. A perceptual system may have the same functional structure in different organisms, but present the world in terms of different qualia. This can be explained in the two systems differing at the lower, homuncular level, either in terms of functional structure at **that** level, or in terms of physical difference - the presence of different neural state types, realising different qualitative state types.

I believe that homuncular functionalism offers the best unified physicalist approach to understanding mind. It offers a naturalistic basis for the existence of content-bearing and qualitative states, one that generalises across different species and physical systems. I will now examine the nature of perceptual experience, to argue that we need to posit qualia to account for it. Given homuncular functionalism, there are no problems in supposing a physical realisation of qualia-bearing states - or at least none internal to physicalism.

II. Understanding Perception

The nature of our perceptual link to the world is the first aspect of conscious experience I want to consider. In part I, I argued against the sense-data theory. In this section I shall try to do better. Like many contemporary philosophers, I believe we must posit qualia to account for secondary quality experience, and thus qualitative mental states that instantiate qualia tokens. Unlike some writers (Jackson 1977, also Block 1978), I think the existence of qualia pose no special problems either for functionalism, or physicalism in general, as I argued in part I. I do think that the existence of qualia has metaphysical significance, but that is the topic of part III. For the moment, I want to put the case for them, and this belongs in a study of perception. I will concentrate on visual perception, and the existence of colour, but my conclusions will generalise to other modalities and qualia.

My argument divides into four sections. Firstly, I examine the nature of perceptual content - the representational states that perception gives rise to. Contrary to many recent writers on qualia, I do not think that we need a novel conception of content for a theory of perception - a category of content distinct from that involved in our ordinary belief states (Peacocke 1983 ch.1, Luntley 1988 ch.7, Crane 1992). Perceptual contents are conceptually structured belief states, like our other belief states. The distinction lies in the nature of the

sensory concepts employed. Secondly, if we accept this as the relevant distinction, then much criticism of qualia misses its target. The notion of qualia can be upheld as required to make sense of our use of sensory concepts. Thirdly, I shall clarify the relation between qualia and our sensory experience, to show that they are not susceptible to the worries about privacy and communicability that can be raised against the notion of sense-data. Fourthly, I shall show how qualitative states can be regarded as playing their role in perception. Following Baldwin (Baldwin 1992), I call such a theory **Projective**, and I will show how it meets empirical requirements and fits the methodological requirements of my overall physicalism.

1. The Nature of Perceptual Content

Perceptual experience gives rise to perceptual states. It also gives rise to beliefs about the perceived world. Sometimes, our perception-based beliefs may differ with the literal **appearance** of the world. When we view a Muller-Lyer diagram, for example, it appears to us **that** there are two lines of differing length, whilst we nevertheless believe that they are the same length (Crane 1992. Jackson 1977 uses this example to argue for the existence of sense-data, but I reject his position from the arguments of part I, section 2). Equally, in our ordinary, non-illusory perception of objects receding into the distance, we may believe them to be of roughly the same size, yet the laws of perspective cause nearer

objects to appear larger (Peacocke 1983 ch.1).

So perceptual content would seem to be distinct from the content of our beliefs. It seems that acquiring perceptual content cannot just be a matter of acquiring beliefs in a perceptual manner. It cannot be a matter of acquiring **beliefs** at all - otherwise the Muller-Lyer diagram would cause us to be in a state of having contradictory beliefs. But when we receive contradictory information in a non-perceptual manner - e.g. being told 'P and also Not-P' - the effect is to fail to impress **any** beliefs in us. Equally, it cannot be the case that we have a 'suppressed inclination to believe' that the lines are a different length. If the inclination is genuinely suppressed, it should cease to apply, as in non-perceptual cases.

Such is the argument that Crane uses to argue that perceptual content is non-conceptual. I shall now consider the options available.

Perceptual content may be:

- (i) representational and conceptual (i.e. belief states);
- (ii) representational but non-conceptual (Crane 1992);
- (iii) sensational and non-conceptual (Peacocke 1983).

I shall argue for (i) by first dissolving the attractions of (ii) and (iii), and showing that they are unnecessary and cannot do the work their sponsors want them to, and secondly showing how (i) can capture the intuitions that drive the others.

Let us first consider Peacocke's notion of sensational content. The thought is that our visual impressions, like our sensations, persist in spite of our belief judgements, and have a nature of their own that is non-representational. As he puts it, sensational properties are 'properties an experience has in virtue of some aspect - other than its representational content - of what it is like to have that experience.' (Peacocke 1983 ch.1) He does not deny that there is representational content to experience, but he does deny that the representational properties are **all** the properties. On this view, we might say that blindsight patients (part I, section 2) differ from normal perceivers in lacking any awareness of sensational properties.

I think this model is inadequate for two reasons, though I should add that Peacocke seems to have subsequently changed his position to something closer to (ii) (see Peacocke 1992). Firstly, it is simply false to say that our awareness of the properties of our visual field - which are Peacocke's putative sensational properties - is of **non-representational** properties. Our awareness of colour is **of coloured objects**; even after-images have extension and homogeneity, which are the characteristics of our object-perception. There is also an ambiguity in whether these properties are meant to be of specific regions of the visual field, or of relations between regions. The latter case seems to be suggested by his earlier cited example, that trees of the same size have different apparent sizes when viewed at differing

distances. Yet it may happen that, for example when viewing some kind of *trompe l'oeil* painting, we initially suppose that we are viewing two similar-sized objects at different distances, but then realise that we have been misled by a trick of perspective and in fact the scene is of two **differently** sized objects at the **same** distance (or vice versa). In which case we would undergo an aspect shift with respect to our perception, whilst the overall visual field - the arrangement of colour patches - would remain the same. In another example, Peacocke characterises sensational properties as those properties of the visual field that remain unchanged during an aspect shift. So perhaps the properties of structural relations are not sensational properties. Yet they seem to be of the image in the same way that its distribution of colour is. I think what Peacocke is trying to say is that the having of a visual experience is having an experience that there is a particular array of sensational properties present to the mind, in the same way that sense-data are constituted by an array of sensory properties. But at the same time he realises that our awareness of perspective effects (as in the 'two trees' example) is governed by our awareness that the experience is of external objects, with spatial relations to each other. On the sense-data view, such awareness is inferential; Peacocke correctly locates it as an integral part of our awareness of the sensory field as a field of objects. The problem is that he wants to say that our awareness of the sensory field is, as it were, 'uninterpreted' - no beliefs automatically

spring from it. He realises that the relational properties of the field are, phenomenologically, of a piece with our awareness of it. But to be aware of those relational properties is to have already 'interpreted' the field, and structured its content.

A second and more general objection is against classing such properties as 'sensational'. We understand sensations as modes of awareness of our bodies; they have qualia, but they do not represent except in a very primitive sense. Having a sensation of pain in my foot indicates to me that something is wrong with my foot (or perhaps I have a trapped nerve somewhere else), but this is not a **structured** signal. Awareness of coloured objects does present a structured signal: the colour defines the shape of the object and it also indicates a pattern of similarity across its surface, which may be compared with that observed on other objects. What the significance of this signal is, is a subject I will turn to later. So colour functions as a medium for presenting information in a way that pain does not (begging no questions as to whether this information is of a conceptual nature). I think it is perfectly right to describe colour awareness as **sensory**, as this draws the important similarity to sensations in that both exemplify qualia. But to describe it as **sensational** is incorrect. It conflates two quite different modes of experience. I think the difficulties in Peacocke's conception that I mentioned earlier stem from this.

Peacocke has now developed a model in terms of what he calls 'scenario content' (Peacocke 1992), which he

describes as 'non-conceptual representational content'. Such is the option (ii) favoured by Crane, so I shall now concentrate on this alone.

Unlike the Peacocke of Sense and Content, Crane does not hold that the visual field has any non-representational content. Perceptual content is representational, but it does not have conceptual structure. Conceptual structure is a matter of having certain inferential relations to other contents. We talk of conceptual content because these inferential relations hold in virtue of the constituents of the relevant contents: if I have the contents that **Fa** and **Fb**, they both share the concept of F-ness; if I also know that $(x)(Fx \supset Gx)$, then I may deduce **Ga** and **Gb**.

Perceptual content cannot be conceptual because it is not susceptible to such relations and inferences: it cannot consist in beliefs because it does not behave like my other beliefs. The basis of this claim, for Crane, comes from the existence of Muller-Lyer type cases. I now want to attack the argument at its base.

Certainly, if we can consistently accept the claims:

(P) I perceive that **p**.

and:

(B) I believe that **Not-p**.

then this shows that (P) does not mean the same as 'I believe that **p**'. Of course it doesn't. It means:

(A) I believe that it appears that **p**.

That we cannot substitute 'I believe that **p**' for the 'it appears...' clause in (A) is no more problematic

than the general failure of substitutivity in intentional contexts.

Is this a trick? No. I suspect that what does the real work in convincing Crane and Peacocke that perceptual states are not belief states is just their sensory quality: things actually **look** a certain way, which is not the case when we reflect on our other belief states (not that reflection makes any difference). Perfectly true, but this is because our perceptual beliefs are specified using **sensory concepts**; and our awareness of them involves awareness of sensory qualia.

One fact about perceptual content that made a great impression on Peacocke was the unrevisability of the constitution of the visual field itself: thus he held that the unchanging sensory structure apparent across an aspect change was due to its unchanging sensational properties (1983 ch.1). But it is simply not true that the sensory elements of perception are completely unrevisable. We can experience aspect changes in the colour properties of objects, from black to blue (see Broackes 1993). Admittedly such phenomena are rare, but their existence refutes the claim that experiences are **wholly** unrevisable.

One reason why we might doubt that perceptual states are belief states is that we are not immediately aware of their content: it seems that we have to form our beliefs **from** the state. I think that this effect is illusory. Being in a perceptual state means being in a highly complex informational state, and our awareness

diminishes away from the centre of vision. In order to mentally focus on some aspect and bring it to the centre of our attention requires a moment of reflection. But it does not follow from this that we are in any way inferring or gleaning the relevant belief from some other non-belief state. For aspects at the centre of vision, there will be no reflective gap in bringing the content to the centre of awareness: we report the way things seem because we already have a belief about how they appear.

Compare perception to the way in which we acquire beliefs through reading. Suppose I were reading a reference book - which I have good reason to treat as authoritative. As I read, I am aware of the words and sentences passing through my mind; I am not aware of any beliefs crystallizing out of them, so to speak. Yet if I were questioned at any time about something I had read I would have the appropriate belief ready to hand. The same principle applies in perception. Our beliefs about the appearance of the world form without our being aware of their extent and detail (and they are often quickly revised and forgotten). We only become aware when an illusion is noticed, and we realise the incongruency between how we believe the world is and how we believe it **appears**.

When I say that we have beliefs about appearance, I don't mean that we have beliefs about a two-dimensional array of colours, or even a 2D-array into which we have projected some depth properties. Our beliefs are about the presence of **objects**. This is an aspect of perception

at which the non-conceptual view displays its inadequacy. Let us take Peacocke's later 'scenario'-content model, as it is more clearly specified (Peacocke 1992). A scenario is 'a way of locating surfaces, features and the rest in relation to ...a labelled origin and family of axes' (Peacocke 1992 p.107). The axes and origin specify the orientation of the observer with respect to the viewed scene. The point of 'scenarios' is that they are possible visual field presentations; the nature of the visual field under-determines the arrangement of external objects that give rise to it (e.g. as in my own earlier example of differently-sized objects at differing distances: the presented scenario is compatible with both interpretations). Since scenarios have representational content but not **semantic** content (since they are not conceptually structured items) they do not have truth conditions but 'correctness conditions': possible arrangements of objects that would give rise to the same apparent scenario. So perceptual content is veridical when the actual **scene** (the arrangement of objects and their visible properties before the observer) is amongst the correctness conditions of the scenario, for then the observer is possesses an accurate representation of the visible world.

The first thing to say about this is that such a specification can be achieved by conceptual content. Peacocke indeed openly admits that the full description of any scenario may involve highly specific visual concepts. But he will not identify the perceptual state

with a content state actually specified in terms of such concepts, for he thinks that it is a fatal objection to the conceptual model that the subject may not possess the concepts needed to describe her visual field correctly. He assumes that any belief/conceptual model must be one that accounts for perceptual content in terms of a relation between the subject and a Russellian proposition (see his remarks on 'the pure propositional account', 1992 p113-116). But we need not take this as our model of belief states. I realise this is a large subject, so I shall just say that I regard visual belief states as informational, cognitive states of the brain, not individuated by reference to propositions (however we specify them); talk of propositional content helps us to discern the elements of such a state but it does not imply that its actual constitution involves a relation to anything, at least not to any putative abstract object.

Peacocke's objection to conceptual specifications is only sound if we suppose that we must use concepts which we cannot justifiably attribute to ourselves and others in our ordinary perceptual activity, when we are not under the influence of any cognitive scientific theory. In his scenario model, such concepts are needed, but only because the scenario is specified **as an arrangement of surfaces and not as objects**. For such concepts are needed to capture the 'fine-grained' nature of the visual field:

[A]n experience can have a finer-grained content than can be formulated by using concepts possessed by the experiencer. If you are looking at a range of mountains,

it may be correct to say that you see some 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)

Obviously the content of my perception of the mountain is not just that it is round or jagged, but it does not follow from this that I must be in a state that does not have a conceptual structure in order to capture the 'fine-grained' aspect: in this case, the minute detail that reveals surfaces as shear or rolling, smooth or rubble-strewn, that distinguishes shadows from crevasses...

There is an ambiguity here about the subject's location with respect to the mountain, which is relevant to assessing this case. If the mountain were over in my peripheral field of view, I would not be in a clear belief state about it, but this is because I would not be in any definite **perceptual** state with respect to it either: it would not be 'fine-grained' if it were specified by a scenario content either (Dennett 1991 ch.11 makes clear the varying determinacy of the visual field). If I were focussing on the mountain itself, I would notice all this intricate surface detail, but this would be grasping the mountain's structure as a three-dimensional object.

For understanding the mountain's appearance is a question of interpreting its features, and its features are determinate when I am sure what they are : whether that dark patch is the shadow of an obtruded rock, or

a cave opening, or a patch of darker rock. Although I can only speak for myself on this, I find that, phenomenologically, the fine-grained structure of complex scenes (such as mountain views) becomes determinate when the structure of the mountain itself - its geographical layout - becomes clear. Yet in Peacocke's theory, the fine-grained structure is a phenomenologically basic structure, the array of sensory properties filling my field of view.

The point is important because scenario contents are necessarily highly complex structures of sensory qualities in order to fit Peacocke's model: they would have to be arcs and skeins of colour and texture, diffuse patterns of light and so on. If the content is specified in this way, then clearly it will involve concepts that most perceivers either do not have or are never aware of employing. But if the basic perceptual content is of objects, then it could be specified in terms of spatial relations and size details. It would still be highly complex in the case of viewing a mountain, but there need be no reason to suppose that in the mountain case we are employing any more theoretical capacities than we do when we view the much simpler arrangements of everyday objects and situations. If this claim is true, then there is no need to suppose that scenario content is non-conceptual, for it may involve no more concepts than it is legitimate to suppose that ordinary perceivers already possess.

In employing objectifying concepts in perception, the links between perception, reasoning and action become

clearer. Our perceptions are beliefs about the appearance of external objects, so we may immediately relate them to our beliefs about those objects, and we can act on them when we observe that a certain condition is met: when the light turns green, and so on. When perception fails to locate any object, such as in after-images, we are aware of the indeterminacy of the colour field; we have no clear beliefs about **what** the apparent object is, and so when we consciously attempt to study it, it becomes no clearer or more determinate. I find sometimes that when I have an after-image, and I immediately attempt to discern its boundaries and colour, the effect is to actually make it **less** determinate - the apparent colour even oscillates between red, green or black. Presumably, my visual system, having no object to become clear about (other than a fading pattern of retinal excitation), is continuously altering and revising its judgements, and thus adjusting the content.

Of course sub-personal content - the intermediate and short-lived states that arise as the perceptual system responds to aspects of its stimuli - should be correctly called non-conceptual. But this is because it **is** sub-personal, below the level of consciousness. At the conscious level I am only aware of the belief states that functionally supervene on this sub-personal activity - or rather, I am aware of the complex informational states that contain a whole mass of different beliefs available for reflection.

Peacocke offers his scenario model both as an account

of this sub-personal content and as the phenomenologically basic component of perceptual consciousness (1992 p.134-5). I agree that at the sub-personal level there may be states and systems that represent a projection of sensory qualities in subject-centred spatial array; in my terms, some of these states will be qualitative and will possess qualia. But I reject the view that this is in any way the **basis** of perception. The existence of blindsight suggests that such a system, whilst a part of normal perceptual awareness, may be eliminated whilst still leaving the subject a (degraded) perceptual system. This indicates that the system does not occupy a causally vital link in the formation of perceptual beliefs, though it does also suggest that it plays a vital role in perceptual **awareness** - after all, blindsight patients believe themselves to be **guessing** facts about their surroundings.

In conclusion, there is no reason to suppose that perceptual contents are belief states with conceptual structure. Neither need we suppose that the subject requires special concepts in order for this to be so. The 'sensational' and 'scenario' models of non-conceptual content do not do the work that they were intended for, and we need not suppose that perceptual content is non-conceptual either because or in order that it play a foundational role in our awareness of the world. I will now argue that we still need to posit qualia to explain our use and understanding of sensory concepts, contrary to objections - some of which assume that qualitative content must be non-conceptual content.

2. The Nature of Sensory Concepts

So our perceptual awareness is a complex, conceptually structured belief state. Our consciousness of it presents a visual field of concrete, spatially ordered objects with secondary qualities. We are aware of these sensory qualities because our belief states contain sensory concepts, and the having of a perceptual belief state with a sensory content is necessarily an awareness of some external state instantiating a sensory property such as a colour.

What kinds of properties are colours? Once we reject sense-data and all indirect theories of perception, we have to take perception as some kind of **direct** relation to the perceived objects. These objects seem to have colour, so perhaps colours are objective properties of external objects.

But this will not do. Science shows us that colour experience is grounded in awareness of the reflectance characteristics of objects (I will discuss the exact relation shortly), but it also shows that this property of reflectance varies across a wider range than we are aware of: the visible spectrum is only a small portion of the whole electromagnetic spectrum. Similarly with hearing, we believe the basis of our aural experience is the existence of air oscillations. Yet normal hearers cannot detect any sound at frequencies above about 18 - 20 MHz. Intuitively, we should distinguish between our experience of these phenomena, in terms of sound and colour, and the physical events themselves: the sensory

quality is a property of our experiential state but does not correspond to an objective property of the physical event that causes it.

If we take this line we are distinguishing a category of properties as **subjective**, for they inhere in the subject's experience of the world but are not objective in the sense of being actual properties of the things and events that we experience as possessing them. These qualities are qualia. But do we need to posit them? Have we misconceived the nature of our experience?

Before reviewing the debate, I should make a note of the contemporary context. Many writers arguing in favour of qualia take them to pose a particular problem for functionalism (Shoemaker 1975, Block 1978) or even physicalism (Jackson 1977) as viable approaches in the philosophy of mind. The claim is that qualia are an unanalysable residue that defy functional characterisation. I argued in part I that we should reject sense-data theory (Jackson's solution) and that there are no particular difficulties in accommodating the existence of qualitative states in an 'homuncular' model of functionalism, contrary to Block and Shoemaker's worries. I nevertheless believe that these writers are correct in claiming that a complete theory of the mind needs to posit qualia. So I shall simply pass over arguments to the effect that subjective experience poses no problems for functionalism: I regard this side of the argument as settled.

This leaves the separate claim that we do not need the notion of qualia untouched. Unfortunately, I think much of the arguments levelled against qualia are in fact arguments against the notion of non-conceptual/non-representational perceptual states. For example, Tye (Tye 1992) argues correctly that cases such as those offered by Peacocke (in Peacocke 1983 ch.1) do not require any new categories of content or sensory states in order to be explained. But this is only an argument against qualia if we follow Peacocke and Crane in supposing that such states or content are needed to accommodate them. My argument in section 1 was that this is not the case, so such criticism misses the mark.

The substantive claim about qualia is that we need them to explain the facts of experiential similarity and dissimilarity, to state such obvious logical possibilities as that functionally similar creatures may experience qualitative differences, and simply to make it clear just what properties of things we are actually talking about when we mention their secondary properties. I shall now examine two lines of argument that do address these claims, and argue that they are unsatisfactory. Then I will review the facts about colour experience to argue positively for the existence of qualia.

If we are to avoid reference to qualia, we have to find some way of stating the facts of phenomenal difference and similarity without mentioning peculiarly subjective facts. That is, we have to be able to individuate experiential states not in terms of their

experiential properties, but of some objective feature. Neural facts will not do, because the question can be asked: 'Which neural states?' - we cannot reply 'Those that are identical with an experience that is...' on pain of circularity. Since our current position is that colours are objective properties of the objects that appear to possess them, rather than subjective properties of experiences, the obvious solution is to individuate in terms of presented objects.

The theory works like this. To have an experience of redness is to be in the same state one is in when presented with a red object. This is not circular, for on pain of begging the question we must assume that redness is an objective property. It may not be a scientific theoretical property, but then neither are sizes and shapes, which undoubtedly are objective properties. The mental state of experiencing-redness may be identified with the corresponding brain state. End of analysis.

Tye (Tye 1992 pgs.168-172) takes such a line. The objective world that we experience sets the facts of that experience. It is a mistake to pick out subjective properties as those in virtue of which experiences may be distinguished; we do not need to peel away the colour from coloured objects and treat it as a separate thing in itself.- we do not need the qualia this would give us.

However, things are not so straightforward. To individuate all possible experiential states, we require some objective state of affairs for each one, since we are claiming that to experience **that p** is just to be in

the state one would be in when confronted by the state of affairs **that p**. In such a way we accommodate hallucinations, for to have an hallucination that, e.g. a dagger is before one, is to be in the same neural state that one would be in if there were a dagger before one. But the actual world probably doesn't contain enough states of affairs to do all the individuating we need. Suppose it never occurs that a purple balloon drifts across a perfectly clear and blue summer sky. Nevertheless, I have an hallucination of such an event happening. What can we say about this experience? We must appeal to possible worlds. In this case, I am in the state that I would be in if, in some possible world, I saw a purple balloon drifting across a perfectly clear blue sky.

But we can have harder cases than this. Perhaps, under some extreme conditions, I have hallucinations of a world in which the laws of nature are different. It could happen, my imagination could be dominated by images from science fiction films. Now what individuates my experience is that it is the one I would have in the relevant world. Would I? Perhaps in that world human perception would function differently. Perhaps I would not be human, or at least my counterpart wouldn't (the interpretation of possible worlds semantics is not relevant to either side of the argument).

Turning away from the modal problems, we have trouble in interpreting some of the problem cases for which qualia have been proposed. There is the issue of the Inverted Spectrum. Let us suppose that someone demonstrates the

full range of human colour-discriminating behaviour. In spite of this, this individual's experience of colour is inverted with respect to ours: she perceives red objects as blue, and vice versa. This is perfectly coherent; after all, we already mark differences between colour perception. There are tests that can distinguish marginal variations from normal colour vision, cases which we would not class as colour-blindness in even a partial sense.

Tye (pgs.167-169) has to say that this is a case in which the subject's vision systematically misrepresents the true state of affairs. Although she is inclined to use the **words** red, blue, and so on in the same way as everyone else, in doing so she is mis-classifying her experience. Thus colour classifications are relative to populations: '[P]erhaps the most natural thing to say is that, relative to humans generally, [her] experience represents green but, relative to [her], it represents red.' (p.169).

Yet we can extend the example. Suppose spectrum inversion is the result of some disease, and it gradually spreads through the population until everyone has their spectrum inverted compared to their healthy selves. How do we express the change that has occurred? We must refer to our earlier, healthy state as definitive of the standard conditions for the correct use of colour terms. But what those standard conditions do is to define our **discriminative** capacities: which items we will class together as exemplifying the same colour. **Ex hypothesi,**

these response patterns have remained the same. The objects that we previously classed together as red (or blue) we now class together as blue (or red). We could distinguish the two populations due to some permanent neuronal change caused by the illness. Perhaps we should say that there has been an inversion in the functional architecture of the brain.

If Tye follows the story in this direction, then I think he should acknowledge subjective qualia. For the story shows that the connection between qualitative content and discriminative response is contingent: we can suppose that we might class the same objects as resembling each other, yet do so in response to a different common property. So we may isolate the aspect with respect to which we discriminate the different classes - the particular colours they possess. We have to make this distinction in order to state the significance of the neural change. Otherwise, with respect to what do we regard the two populations as differing? Let us suppose that the later population, having realised the great change that has befallen them, change to using colour words to describe the classes which they now perceive as so-coloured. That is, they now use 'blue' instead of 'red', and so on. Then when they discover the underlying neuronal difference, they should say that the inversion was with respect to **qualia**, not discrimination or language use. It would be the better explanation.

There is one final objection I will make to Tye's account, not connected to colour experience. Even if we assume no difficulties in appealing to other possible worlds, we cannot correctly individuate certain experiences. The experiential state caused by a Muller-Lyer figure, to a normal percipient under normal conditions, is identical to that caused by a similar diagram in which the two lines actually are different lengths. But the theory depends on external states of affairs to individuate experiences, and here we have two different states of affairs that are experientially indistinguishable. This introduces a general problem for specifying 'normal conditions': we would want to say that they are those under which the experience matches the presented scene. Muller-Lyer diagrams look quite different from the way they actually are, under normal viewing conditions.

The second line of argument against qualia that I shall consider denies that we need to attribute special qualia-identifying skills in understanding our grasp of colour concepts, and thus that there is any role for qualia in our sensory concepts. Levin (1986 pgs.480-486) offers an argument of this kind. But I should first make some distinctions. Levin is responding to claims by Jackson and Nagel (1979) to the effect that physicalism cannot account for the subjective qualities of experience, in the sense that a physicalist account of some form of mind would not give one the knowledge of **how** its subjective states presented themselves in its consciousness. This is a different claim from the one

that I am advancing in this essay. I think there is no problem with placing qualia in a physicalist theory of mind, I also think we can be justified in making some inferences about the qualia experienced by other species reasonably similar to ourselves. Nagel's concern is that awareness of qualia is subjective, and so a fully objective description of an individual mind, since it would not be necessarily stated in a language tied to describing the subject's qualitative states, would not be able to acknowledge the existence of these facts. I think Nagel's position is bound up with his commitment to metaphysical realism, and I will discuss it in more detail in part III. What interests me is that Levin's response denies that experience of qualia plays a constitutive role in the possession of sensory concepts. It is this claim that I shall now contest.

Levin initially distinguishes between 'having a concept and having the wherewithal to apply it' (p.480) Being able to apply it is not required for full knowledge of the facts that the concept relates to. 'Knowledge' in this sense is objective knowledge. Its alleged failure to encompass subjective facts is Nagel's problem, but it is not **my** problem (see part III). What interests me is the subjective case. Suppose someone devoid of colour experience, who nevertheless knew about the full range of normal human colour discriminations (which objects are classed together, what the relations of similarity are between the different colours, and so on), were suddenly made able to have colour experience,

Would we say that an awareness of qualia now gives her full possession of colour concepts, or that she now gains the wherewithal to apply concepts she previously had? If the latter, we should wonder if qualia play a substantial role in our use of sensory concepts - and if they **don't**, then we have less reason to accept them at all.

The relevant part of Levin's argument is:

[T]he failure to immediately identify red and green upon first being shown any colors at all may be taken to show a deficiency in [the] ability to apply color concepts to ...experience, and not a deficiency in those concepts themselves. As in the case of...Molyneux's man born blind, [she] will have the relevant color concepts as long as she has sufficient information about the structure of the perceptual field, the similarities and differences among the experiences in it... (p.485)

The reference to 'Molyneux's man' is to a classic empiricist thought experiment. Molyneux wondered: if a man born blind, who experienced the world through touch, were suddenly made to see, would he be able to distinguish a sphere and a cube immediately on seeing them? Locke thought not (see Evans 1985). **The issue is whether** we should suppose him able to generalise his concepts of shape to his new sensory modality. In the empiricist model, we suppose that sensation and thus knowledge of external objects comes through the varieties of sensory impression one receives. We would then think that tactual sensations were structured by their own category of concepts; thus the blind man would have to **acquire** the new visual concepts needed to discriminate shapes. We should reject such a model. The correct approach to sensory awareness is to

assume a unified spatial awareness, with a single set of spatial concepts employed in mental representations derived from the workings of all the senses - such is the moral of modern cognitive research. Levin argues analogously that we can suppose colour concepts to be in nascent form in the absence of experience. With sufficient advance knowledge, one can be able to orient oneself appropriately with respect to the new stream of experiential input, so a capacity to identify qualia correctly is not an essential part of the possession of colour concepts. It rather belongs to 'having the wherewithal to apply them'.

The analogy is spurious. From the vantage point of science we can recognise that Molyneux's blind man already possessed spatial concepts that he could generalise to visual experience if he acquired it, but nothing follows from this about the structure of the corresponding colour case. We credit the blind man with spatial concepts because both the visual and tactual cases relate spatial information, and they place instances of solidity - primary qualities - into a single unified space. However in the colour case, we are given a new range of experience about objects, and with it a new **category** of property to be aware of. This category was entirely absent from the earlier colourless experience, unlike the primary qualities present in both tactual and visual experience.

Even if we grant knowledge of 'constitutive' truths such as 'Nothing can look red and green all over at the same time (Levin p.485), this gives no purchase on the question of **which** colours are red and green (unless one

is just told which objects possess which colours. But to do that is to abandon the exercise rather than settling it).

It is possible that one could infer which colours are which from the relative similarities, though I am doubtful of this. Colours differ from each other in three dimensions: hue, brightness, and saturation. An account of similarity relations would mention all three aspects. But why should we suppose that someone without colour experience could apply **those** concepts without prior experience? At best, I think the debate about the 'Knowledge Argument', which has grown up since Nagel and Jackson's attacks on physicalism, is inconclusive.

I suspect that when people argue against qualia in this way, the thought is that colours can be distinguished by certain assymetries in our experience of them: our greater capacity to discern shades in particular parts of the visible spectrum. Thence comes the idea that knowledge of similarity relations is sufficient for possessing colour concepts (Dennett 1991 ch.12 seems to be influenced by this thought). If this is so, it is a purely contingent fact, and it can hardly constitute a conceptual truth.

I now want to turn to some empirical facts about our colour experience, and argue that we **need** qualia, philosophically, to understand what is going on.

Contrary to a long-established tradition in science and philosophy, colours do not correspond to the reflecting of particular wavelengths of light, or to the existence of common textural properties in similarly coloured objects

(Westphal 1991 ch.6 disposes of a range of proposed 'physicalist reductions' of colour). The property that similarly-coloured objects do have in common is their reflectance profile - the proportion of incident light reflected at each particular wavelength (Campbell 1982). But this is a relationship between the nature of the incident light as well as the reflected light. Also, it is not clear why colour experience should also extend to light **emitters** as well. Perhaps our visual system simply treats them as reflectors and responds accordingly. After all, for almost all of the period in which humans evolved, only natural light sources existed.

As reflectance is a relation between two kinds of light, our awareness of it must be some kind of judgement about the overall contrast effect. I don't mean that the visual system actually 'decides' which colour is present, but the relevant representational state must be determined by some sort of comparative process. We can study colour perception experimentally by presenting subjects with varying mixtures of the three primary colours (red, blue, yellow). By varying the flux intensity of these three, we can produce the full spectrum of visible colours in the observer.

To study individual colour response in more detail, scientists have recently turned to examining the response systems from the retina inwards. The human retina contains three different kinds of pigmented cells, which have their own distinct wavelength absorption behaviour - they consistently absorb different intensities of light at particular wavelengths. Once we know how each system

happens to respond across the range of visible wavelengths, we can construct a **wavelength mixture space** (Clark 1993 ch.2). We have three axes representing the intensity of light (in numbers of quanta) absorbed by each system; we then place a point for each incident wavelength at the position defined by the absorbed intensities in the three systems. Now it turns out that we have a model of the subject's colour experience from which we can predict her response to any combination of wavelengths at any intensity. For it turns out that we can represent particular wavelengths by vectors from the origin to the corresponding point in wavelength space, and this point defines a unique colour response in the subject. Thus we can represent any mixture of wavelengths by the relevant combination of vectors (the length of the vector represents its intensity) and the point that the combination arrives at causes the identical colour response that would be caused by a single wavelength of the appropriate intensity (Clark pgs.39-44).

Such is the current scientific position, and it refutes any naive views we might have about qualia being beyond investigation. But such research doesn't give us the whole story about colour perception, for a simple reason that Clark himself acknowledges: the entire approach depends on the subject contributing her own judgements of qualitative similarity.

This is most apparent when we generalise across different perceivers. What remains constant is the structure of the individual's colour 'quality space'.

We can calibrate an individual's particular colour experience against incident wavelengths by asking her to identify the spectral combinations that give rise to the experience of unique hues - e.g. green that is not at all yellowish or bluish (Clark p.165). Given such a position, we can superimpose a wavelength structure on to the general colour space structure, and be able to predict which shades the subject will identify at certain spectral combinations. But for this to work it is essential that we have the subjective judgement about the nature of the induced experience:

A given 'place' within the psychological colour solid may be occupied by different classes of stimuli in different people - even among those who are neither colour-anomalous nor deficient in their discriminative capacities...[D]ifferent individuals will consistently arrive at different [wavelength calibrations], and the distribution even in the 'normal' population is wider than one might expect. (Clark p.165)

Such variation demonstrates that colours cannot be objective properties correlated with certain wavelengths. Colour experience is rooted in the physical nature of the observer. Nevertheless we can make sense of it having a common structure in most normal subjects, and there being properties that most of us are capable of identifying and making comparative judgements about. I say that these properties are qualia, and that they are physically realised in qualitative brain states in the perceptual system.

There are three final points I want to make with regard to Clark's synthesis of modern research into sensory property perception. Firstly, the model applied

to colour perception can be used for other sensory qualities as well. We may draw up taste quality spaces, auditory quality spaces, and so on. In each case, we find a generalisable structure of qualitative similarities. The number of dimensions in which this space is articulated can be related to the physical structure of the sensory system. Just as the three kinds of retinal receptor in vision yield a three-dimensional quality space, so the four kinds of taste receptor yield a four-dimensional taste space. (This is a great simplification of taste science, though the underlying principle is correct. I cannot pursue the subject here. See Clark pgs.140-144).

Secondly, the need for subjective judgement in calibrating individual quality spaces amounts to an empirical refutation of the Levin view that experiential content is not an essential aspect of sensory concepts. One could argue that the qualitative nature of a sensory concept could be identified from a knowledge of the general human colour space, and a little exposure to colour experience. Clark himself argues that colour space contains certain 'landmark' shades that can be located in pure relational terms. Given knowledge of 'the locus through colour space along which humans can make the greatest number of discriminations' (p.200), one could come to relate which of one's experiences were of the colours one had been told about.

But it is a purely contingent matter that human quality space has these features. It could have been that

it was 'relationally symmetric', without any of these qualitative 'landmarks', as Clark concedes. If this were the case then acquiring colour concepts would have to be a process of learning them through the direct teaching of others, with the use of example-objects to indicate which kinds of experience instantiate which properties. In such a case it would also be impossible ever to verify whether or not a subject has an inverted spectrum, short of inspecting their neural properties. But I do not think that such a case can be taken as a **reductio** of the notion of qualitative states. They are still needed to fill a theoretical role in our model of perceptual experience.

Thirdly, the fact that colour similarities have a finer structure than we normally remark on - they vary with respect to hue, saturation, and brightness - indicates that we need to regard sensory content as conceptual. We possess the concepts because we can observe the relevant qualitative orderings that they give rise to, and their presence in perceptual content ensures its determinacy in these respects.

Clark takes his psychophysical account of sensory qualities as settling the scientific need to explain the phenomenal character of experience. I agree that his model shows how we can place qualitative states into a general physicalist theory of the mind. Yet there still remain conceptual issues about the place of qualia in human experience that have to be settled. I now turn to the question of how we are able to **share** our awareness and understanding of them.

3. How we understand qualia

Given that we have qualitative states, we need to know how it is that we are able to describe them with a common vocabulary. The sense-data theory made the whole of the mind's perceptual relations with the world into transactions involving subjective states. It then foundered on the difficulties in explaining how it is that we can share a language that describes such states, and how we can consistently apply standards of correctness in making judgements about them. Both these problems stem from the **privacy** of sense-data. We might respond to this failure by rejecting the notion of there being **any** intrinsically subjective element in perception, but this would be overplaying the argument. For us to be able to agree in our use of colour words, it cannot be that we all experience qualitatively-identical, objective colour properties, because, as a matter of fact, we don't. Intersubjective differences in colour vision can be empirically demonstrated, as Clark's research shows. We need a model of how qualitative states relate to our descriptive language that can deal with this fact.

Qualitative states are subjective, but the non-qualitative aspect of perceptual content isn't. I suppose a sufferer from astigmatism would have her spatial perception systematically distorted, so we might say that it also had subjective idiosyncrasies; but such cases are rare so I shall just gloss over them. What counts in all cases is that we agree in our objective judgements about the existence of a world of objects.

If someone disagrees substantially about its constitution, we can soon identify her as suffering from hallucinations or delusions. But at least for most of the time, we are in agreement about what objects the world contains, and this gives us something to orient our experiences with respect to.

Having particular subjective experiences is also part of our normal condition. There is, as Clark observed, wide variation between the actual states that specific lighting conditions give rise to throughout the population. But broadly speaking, we have similar colour experiences. Person X may experience wavelength W as pure red whilst person Y has a sensation of orange-tinged red. Nevertheless, they both agree that the sensation is definitely red. I think it is just because most of us agree about the experience at this level of generality that colour language can take hold.

This remark needs expanding. When I say that we agree about the **experience**, obviously I don't mean that we are able to compare the actual experiences. What we can compare though are our discrimination reactions. For language to take hold, I mean that the fact of our mostly sharing such reactions creates the possibility of our constructing linguistic conventions to describe them. If we didn't have similar reactions, we could not devise such conventions, for there would be no salient aspects of our behaviour with respect to colour properties about which we could converge in establishing a linguistic norm.

It is sufficient for such linguistic practice to develop that **most** of us agree about colour, and we **broadly** agree about it. I suppose this is why our colour language has the degree of precision it does. 'Red' and 'blue' and the others pick out wide sections of our quality spaces. It turns out that our personal judgements of 'unique hue' vary widely over these regions (see Clark, quoted earlier). This is understandable, and it fits with my linguistic conjecture. We have a basic colour vocabulary at just the right level of precision for us to generally agree in its application. We might experience a certain red as a variety of shades, but we all agree that it is more similar to other red things than to any blue or green objects. We differ in our experience because our experience is caused by the observed wavelength pattern, and there is enough variation in human brain structure to cause a corresponding variation in qualitative response. If it were the case that we had a much finer agreement about our colour experiences, we might have developed a colour vocabulary that distinguished different shades and elements of the quality space with finer detail than our actual one does.

Of course I cannot **prove** the general similarity of our qualitative experience, but I can infer to it as the best explanation of our having a shared colour language. The explanation runs through the sort of story about the evolution of linguistic convention I have just outlined. The existence and basis of these conventions also explain how we are able to learn and

transmit our colour concepts as and to children. Because they have similar dispositions as ourselves, they are receptive to the use of colour exemplars - particular coloured objects - in teaching the language. Agreement does not need to be universal for the conventions to come into existence: we can still identify varying degrees of colour-blindness, through a subject's persistent failure to respond appropriately.

This explains how we can have a common language to describe subjective states, in virtue of widespread inter-subjective similarity manifested in discriminative responses. A further issue is how qualitative states relate to the subject's own consciousness.

This is not the same issue as the structure of perception, which I will deal with in the next section. We need to get clear about a perceiver's relation to her subjective states. Sense-data theory treated this awareness as incorrigible or ineffable, usually in pursuit of some wider epistemological goal. Taking such a position leads to difficulty however. If it is the case that we are never **wrong** about the immediate content of our awareness, then we cannot justifiably rank our experiences according to reliability: certainty is a property that attaches to all of them. Of course, in sense-data theory certainty only attaches to the sensory field contents themselves. The thought is that we may be wrong about what is going on, but we cannot be wrong about how it **appears to be**.

This approach is misconceived. I treat my normal

perceptual experiences as certain and reliable because, under normal conditions they represent the truth. But this does not exclude the possibility that under some extreme conditions (there is a bang, I turn round, everything flashes by suddenly in a blur...) my resultant perceptual content is confused, or fails to resolve any clear objects. Then I would want to report: 'I don't know **what** I saw', and I would be glad to defer to someone else's opinion as to what happened to me, and thus what I saw. So we have no need to suppose that qualitative contents have any of the properties that caused sense-data to be problematic.

Accepting such a possibility of error does not impugn my account of perceptual content as conceptual. The brain is a physical system, and complex physical processes such as perception need time to occur properly (Dennett 1991 ch.6). If a chaotic event occurs quickly, the system may not respond adequately, leaving me in a confused, complex belief state (see section 1). These beliefs may still be conceptually structured, but the overall informational state is so disordered I can make nothing of them.

4. The projective theory of perception

Finally I come to consider how to fit all the elements together into a full theory of perception. An adequate theory must explain how qualitative content enters into consciousness, how it achieves its spatial articulation, and account for the effect that neurological disorders have on the phenomenology of perception.

Such disorders include not just blindsight cases, but also the parallel condition of 'space-blindness' (Baldwin 1992 pgs.191-5). Here, the subject has awareness of coloured objects before her, but has no clear awareness of the spatial relations between them. This is difficult to visualise, but by their own testimony sufferers are able to see the actual **objects** before them, whilst having no inkling about their positions.

Clearly this condition indicates, like blindsight, that a particular function of vision is localised in a separate system, so that certain patterns of cerebral damage interrupt the working of the overall perceptual system differentially. It seems natural to suppose that there are different functional elements articulating sensory qualities, and the spatial organisation of the visual field, which are disrupted in blindsight and space-blindness respectively.

Another aspect of the visual field that we must take notice of is the phenomenon of 'filling-in', emphasised by Dennett (1991 ch.11) in his critique of sense-data. The retinas both contain blind spot regions, due to the openings of the optic nerves at the rear of the eyes. These regions should take up a 6° degree wide area of the visual field, yet we never notice them. Why? Because the visual system compensates by 'filling-in' the empty region with an appropriate texture. But this process is not a clumsy business of patching the image up with an area of colour. If the blind spot interrupts a patterned area of space, then the pattern is continued across it.

Clearly the visual field that we are aware of is the product of intelligent and detailed interpretive processes. We need room for such capacities in our model.

The projective theory is so-called because it posits a **projective** relation between the perceiver and external objects; the objects are directly involved in the perceptual state, yet the state also involves the mind projecting its own qualities onto the object. This is not to say that qualia are actually squirted into the world in perception. It means that in the act of perception, the object is presented as possessing subjective qualities deriving from the nature of the subject's experience of it.

To make the notion of projection clearer, consider its role in anti-realist theories of probability and moral properties. Such theories hold that chance and goodness certainly seem to be part of our experience, but they should not be given the realist's literal interpretation. They are qualities we have projected into our experience of the world; they are not objective features of it. The only objective facts about chance are those pertaining to our subjective propensities to anticipate certain events. Equally, on a projective theory, the facts about moral natures are just facts about our dispositions to discriminate conduct and events in certain ways. 'Only' should not be taken here as implying some repudiation of the concepts at stake. Rather, the anti-realist intends to offer the correct analysis of our use of such concepts; whether we decide to do without

them is a separate question that requires a separate argument. The projective argument simply establishes that the realist's ontological commitments are not required to make sense of these concepts.

All of which chimes well with the position I shall offer in part III. For the moment, we need to understand how projection relates to perception, and also whether it can be fitted into the naturalistic framework of physicalism.

On the first point, I suggest that projection is a high level functional property of the system. It supervenes on the functioning of the sub-personal, computational systems that yield various intermediate states representing aspects of the external array of objects. As these states are at the sub-personal level, we may suppose that they have non-conceptual content. They may even have content similar to Peacocke's 'scenario content' (see section 1).

The neurological disorders suggest that in the absence of qualitative awareness, some beliefs about external objects may still arise. These beliefs are not conscious, however. What must happen is that the existence of sub-personal content, though not drawn into the formation of a conscious perceptual state, nevertheless forms a sub-conscious informational state, from which the blindsight patient derives her 'guesses' on questioning.

In space-blindness, a conscious state is formed, and a degraded act of projection occurs. Qualitative states are formed by the sub-personal systems that respond to reflectance levels; the particular qualia

that are present are determined by the subject's neural peculiarities. At some level, the articulation of the colour quality space with respect to wavelength mixture space will be determined, presumably due to the fine structure of the neural system type-identical with it. But for the space-blind subject, the full spatial organisation of the visual field is either not determined, or causally cut off from the projective system. So the resulting perceptual state has no spatial relations.

If there are no neurological disorders, then the projective state does have the appropriate spatial properties.

The projective process can be regarded physicalistically. Projection is at the top of the hierarchy of functional properties; it is the process that brings the organism into consciousness of its surroundings. Thus we have a complete physicalist theory of perception, including the position of subjective properties within it.

III. The Metaphysics of Qualia

The admission of subjective facts into our ontology does not conflict with the methodological goals of physicalism, or cause us to abandon a basically functionalist approach to its realisation. The effect is rather to cause a revision of our **metaphysics**. In this chapter, I will argue that accepting subjective facts requires the abandonment of metaphysical realism as our overall interpretation of our knowledge claims. This does not require a revisionary attitude to our current scientific knowledge. However, we should recognise limitations in our capacity to generalise from our physicalist theory of our own mind and experience to other possible forms of mind.

I take the arguments of part II as showing the indispensibility of the notion of qualitative states in accounting for the subjective properties of experience. Furthermore, the arguments of part I showed that such states can be taken to be physical, and that the properties that individuate the states with respect to each other - the qualia types - should also be taken as physical. The subjective nature of these properties is given by their entering into conscious events in a suitably structured functional system. What individuates these properties are the phenomenal characters they present to the subject of the experience; an experience with a different character would be one involving a different quale.

The claim of metaphysical realism is that the world is determinate in its properties irrespective of our capacities to know the nature of those properties. More exactly, realism is defined by accepting both of two theses about content (Luntley 1988, Introduction):

- (1) The characterisation of content requires the subject's possession of a conception of a world beyond that which is experienced.
- (2) The contents we grasp are contents that have a determinate truth value independent of our knowledge of that value.

Thesis (1) expresses the notion of contents representing an objective world beyond the subject possessing them. It amounts to a basic idea of objective truth. Thesis (2) expresses the truly contentious idea at stake with realism: that the world may be **unverifiably** different from our putative knowledge of it.

I contend that accepting qualia as physical properties - much as we accept electromagnetism and gravity, to fulfil a theoretical need - brings us into conflict with the realist thesis (2). Since we cannot avoid appeal to qualia, on pain of beginning to theorise about the mind all over again from the beginning, we should drop (2).

The problem does not arise because we have incorrigible access to our qualitative states, or anything like it. It is simply that realism claims, for all aspects of the world, that there may be a gulf between their reality and the appearance they present to us. But the whole reason for characterising qualia as subjective is to

express the fact that they are nothing but appearances, and that such a gulf does not exist even in principle.

The contradictory implications are:

(3) There are facts relating to our experience, whose nature is constituted by their contribution to our experience.

(4) For all properties, their real nature may be beyond our capacity to experience it.

We can understand (4) in terms of objective scientific properties comparatively easily: it expresses the thought that properties like gravity and mass may not be as we believe them to be, and the truth may in principle forever elude us (we suppose). In terms of the subjective properties of qualia, it cuts directly against (3), which just is the claim that qualia exist.

What would we lose if we gave up (2) and thus realism? Not much. Anti-realism can still acknowledge (1) and thus the notion of objective fact. What it does not accept is that facts are determinate beyond our capacity to detect them; that for some content whose truth value we may never be able to decide, there exists a truth value.

Anti-realism does not however challenge the realist interpretation of scientific theory. The ontological claims of theories may be taken seriously, and the methodological principle of inference to the best explanation employed. The only restriction on the realist model is that putative truth claims should be claims to

some kind of **experientiable** truth. But 'experientiability' is open to a liberal interpretation. The states of affairs that are experientiable are:

(a) Any state of affairs perceptible with human sensory organs;

and:

(b) Any state of affairs whose constituents or their properties may enter into lawlike causal relations with a state of affairs that is experientiable in the sense of (a).

(a) covers everyday life, (b) permits oscilloscope readings and the results obtained on the screen of an electron microscope.

It could now be objected that the qualitative states of non-humans are not open to our experience, and so their existence must be rejected by these criteria as unknowable. Not so. In so far as non-humans obey psychophysical laws, their experiential states will be causally linked to their overt behaviour. Admittedly we cannot experience the particular qualia, but we can at least make sense of there being some true qualitative attributions applying to them, although we do not know their exact detail.

One limitation that is placed on our understanding of other minds is that they must be reasonably similar to ourselves, for us to make reasonable mental state attributions of them, and consider them possible possessors of experiential states. The anti-realist world-view is ultimately a human-centred one. It is in relation to our own experience or capacity for experience that we

should judge possible knowledge claims, and this includes claims to know of other minds.

I think that a further consequence of this is that Nagel's (1986 p.18) project for understanding human mentality in terms of a wider phenomenon of consciousness is misconceived. Nagel adopts a realist position, and then wishes to have a general, unified theory of the mental in all its instances, rather as physicists may wish for a unified theory of forces. 'We must think of mind as a phenomenon to which the human case is not necessarily central'.

But I think it is central to our notion of mind, and it sets the limits to what we might consider to be mind as well, namely that it must in some way resemble ourselves, and manifest its mindedness to us.

Consider this fantasy. We don't attribute mental states to individual neurons, yet millions of them causally interacting constitute out mental states. Millions of us causally interact everyday, in fast-changing structures of cities, nations, and so on. Could it be that **these** structures instantiate mental states, enjoy qualia, have propositional attitudes?

I don't think that this suggestion makes sense, I don't understand what it would actually mean to accord mentality to such entities. I think it would be as warranted as according it to inanimate objects, a bizarre and almost meaningless extension of our ordinary concepts into extraordinary circumstances. Or rather, the suggestion is perfectly meaningful, but I see no reason to suppose

that it has a truth value. Such mental states could never manifest themselves in our lives - or perhaps if they did (or do), we have no hope of ever knowing of this.

Nagel's realist metaphysics leave such options as live ones. I think this is another reason in favour of anti-realism.

BIBLIOGRAPHY

- Baldwin, T. (1992) 'The Projective Theory of Sensory Content', in The Contents of Experience, ed. T. Crane
- Block, N. (1978) 'Troubles with Functionalism' reprinted in Lycan 1990
(1983) 'Mental Pictures and Cognitive Science' reprinted in Lycan 1990
- Broackes, J. (1993) 'The Autonomy of Colour', in Reduction, Explanation and Realism ed. Charles & Lennon
- Campbell, K. (1982) 'The Implications of Land's Theory of Color Vision', reprinted in Lycan 1990
- Churchland, P.S (1986) Neurophilosophy
- Clark, A. (1993) Sensory Qualities
- Crane, T. (1992) 'The Non-Conceptual Content of Experience' in The Contents of Experience, ed. T. Crane
- Davidson, D. (1970) 'Mental Events', reprinted in Essays on Actions and Events
- Dennett, D. (1978) 'Beyond Belief', reprinted in Brainstorms
(1991) Consciousness Explained
- Dummett, M (1978) Truth and Other Enigmas
- Evans, G. (1985) 'Molyneux's Question' printed in the Collected Papers
- Fodor, J. (1975) The Language Of Thought
(1981) Representations
- Hill, C. (1991) Sensations
- Horgan, T. & Woodward, J. (1985) 'Folk Psychology is here to stay', reprinted in Lycan 1990
- Jackson, F. (1977) Perception
- Luntley, M. (1988) Language, Logic and Experience
- Lycan, W. (1988) Consciousness
(1990) Mind and Cognition (ed.)
- MacIntyre, A. (1981) After Virtue
- Mellor, D.H. (1991) Matters of Metaphysics
- Nagel, T. (1979) 'What is it like to be a bat?' reprinted in Mortal Questions
(1986) The View From Nowhere

- Peacocke, C. (1983) Sense and Content
(1992) 'Scenarios, concepts and Perception'
in The Contents Of Experience, ed. Crane
- Pears, D. (1987/8) The False Prison (2 volumes)
- Shoemaker, S. (1975) 'Functionalism and Qualia',
reprinted in Identity, Cause and Mind
- Stich, S. (1978) 'Autonomous Psychology and the Belief-
Desire thesis', reprinted in Lycan 1990
(1983) From Folk Psychology to Cognitive Science
- Tye, M. (1992) 'Visual Qualia and Visual Content'
in The Contents of Experience ed. Crane
- Westphal, J. (1991) Colour: A Philosophical Introduction
- Wittgenstein, L. (1953) Philosophical Investigations