Essays on the Perception, Representation, and Categorisation of Colour

Will Davies
Balliol College, University of Oxford

THESIS SUBMITTED FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

Word Count (Excluding Bibliography): 73,470

Abstract

This thesis develops and explores a *constitutive approach* to colour vision, which serves as an alternative to the standard experiential view of colour vision operating in the philosophy of colour. The approach seeks to describe the *nature* or *essence* of colour vision *qua* psychological kind. I argue that it is constitutive of colour vision that an organism possesses the ability to achieve colour constancy. An important feature of my account is that colour constancy is characterised as the ability to discriminate differences in surface reflectance properties across changes in illumination conditions. This differs from the standard 'appearance invariance view', which characterises colour constancy by appealing to the phenomenology of apparent colour. I consider an important objection to the appearance invariance view posed by the argument from illumination, which might also seem to carry over to the reflectance discrimination view. The objection is based on the claim that in standard cases of colour constancy the phenomenology of apparent colour is partly illumination-dependent. I argue that the reflectance discrimination view is perfectly able to accommodate this point. As a case study in applying the constitutive approach to illuminate the distinctive nature of colour vision, I argue that a vivid feature of our ordinary experience of colour known as *categorical perception* should be dissociated from our colour vision abilities. Although colour ontology often is not at the forefront of discussion, these constitutive theses support the ontological view of colour known as *reflectance physicalism*. I critique the *argument from colour similarity*, which many take to pose the greatest threat to reflectance physicalism. The thrust of the argument is that colours phenomenally appear to stand in similarity relations that do not correlate with the similarities that are evident among reflectance properties. This argument lacks much force, however, as it fails to acknowledge the extreme context sensitivity of similarity and the presentation sensitivity of our knowledge of similarities.

Contents

Preface....5

0. Introduction....7

1. Colour Vision, Colour Constancy, and Discrimination....19

1.1. The Wavelength Discrimination View....22

1.1.1 Discrimination and Representation....26

1.1.2. The Centrality of Colour Constancy....40

1.2. The Reflectance Discrimination View of Colour Constancy....51

1.3. The Consequences for Colour Ontology....68

1.4. Objections and Replies....72

1.4.1. Non-Reflectance Colours....72

1.4.2. Dichromatic Colour Vision....77

1.4.3. 'Disregarding' the Illuminant....79

1.4.4. Access to Spatial Representations....82

1.4.5. Explaining Approximate Colour Constancy....83

1.4.6. Reflectance Properties *versus* Colour Categories....87

1.4.7. Vision for Action....91

1.5. Conclusions....92

2. Colour Constancy and Illumination....94

2.1. The Argument from Illumination....95

2.2. The Colour Appearance Claim....102

2.3. A Rearguard Action Against the Colour Appearance Claim....119

2.3.1. Noë and Cohen....119

2.3.2. The Reflectance Discrimination View....125

2.4. Conclusions....139

3. Colour Vision and Seeing Colours....141

3.1. Categorical Perception of Colour 143
3.2. Non-Categorical Perception of Colour 150
3.3. The Dissociation of Colour Vision from Categorical Perception of Colour 155
3.4. The View from Cognitive Neuropsychology 172
3.5. Conclusions 177
4. Colour Similarities at Different Levels 179
4.1. The Argument from Colour Similarity 180
4.2. Standard Reflectance Physicalist Responses 183
4.3. Categorical Perception and Similarity Effects 186
4.4. The Context Sensitivity of Similarity 188
4.5. Presentation Sensitivity and Knowledge of Colour Similarities 200
4.6. Conclusions 212
5. Concluding Remarks 214
Bibliography 219

Preface

Believe it or not, this thesis is really about vagueness. Or at least it was. My original thesis proposal was to write about the cognitive origins of vagueness in natural language. Early research into this topic led me to discover the vast body of empirical work on categorisation and similarity in cognitive psychology. Within this literature, certain cases relating to the categorical perception of colour seemed to be highly relevant to understanding the allure of sorites-type reasoning to otherwise semantically competent and rational individuals. As my research progressed, however, the connections with vagueness became more tangential and it seemed increasingly natural to situate my work within the philosophy of perception and philosophy of colour. This thesis marks the culmination of this somewhat unusual intellectual voyage. Perhaps one day I shall make good on that thesis proposal.

I have been extremely fortunate to have two outstanding and highly supportive supervisors during my time at Oxford. Tim Williamson guided this project from its first incarnation as an overambitious BPhil thesis on vagueness to the work that you see today. Tim epitomises the philosophical ideals of clarity and rigour to which all of us aspire. One simply could not hope for a better intellectual role model and mentor. My co-supervisor John Hawthorne proved to be an abundant source of extraordinarily imaginative questions, ideas, counterexamples, and 'toy land cases' that served to test my arguments and abilities to the limit. Above all, John always sought to make the very best of the work presented to him and I thank him sincerely for his patient and constructive spirit in guiding this project to completion.

In autumn 2008 I spent a term at Institut Jean Nicod in Paris under the auspices of Paul Egré's Cognitive Origins of Vagueness project. I found the multi-disciplinary

environment at Jean Nicod to be extremely stimulating and I am grateful to Paul in particular for his enthusiasm and encouragement. In autumn 2010 I spent a remarkably productive term at MIT under the supervision of Alex Byrne, who generously and patiently worked through every draft chapter that I sent his way. For financial support I am indebted to the Arts and Humanities Research Council, the Laces Trust, Balliol College, and the Philosophical Fellowship Fund. Since October 2011 I have been supported by a Junior Research Fellowship at Churchill College in the University of Cambridge. The input and advice of my new colleagues during the latter stages of the thesis has been most gratefully received. I have benefited in particular from conversations and exchanges with Tim Crane and Adrian Boutel in our research seminar on Consciousness, Colour, and Content in the spring of 2012.

Finally, I express my gratitude to Mum for her constant love and support, and Dad for his inspiration and guidance in matters both personal and philosophical. I reserve my deepest thanks for Gabriel, who has been an unwavering source of encouragement, driving me on when spirits and resolve were at their lowest ebb, and who will celebrate the completion of this thesis with me now as we look towards our future of married life together.

0. Introduction

This thesis discusses four interlocking issues: colour vision, categorical perception of colour, colour similarity, and colour ontology. How do these issues relate? And what is their philosophical importance? In order to tackle these questions, let us begin by discussing the *ontology of colour*, which is surely the central issue within the philosophy of colour. Historically the core dispute in this area has been between colour realists, who think that colours are genuine properties of material objects, and anti-realists, who deny this and hold that no ordinary worldly objects are coloured. This perennial disagreement is well summarised by Boghossian and Velleman (1991: 105),

> The dispute between realists about colour and anti-realists is actually a dispute about the nature of colour properties. The disputants do not disagree over what material objects are like. Rather, they disagree over whether any of the uncontroversial facts about material objects – their powers to cause visual experiences, their dispositions to reflect incident light, their atomic makeup, and so on – amount to their having colours. The disagreement is thus about which properties colours are and, in particular, whether colours are any of the properties in a particular set that is acknowledged on both sides to exhaust the properties of material objects.

In addition to this fundamental schism, there are of course numerous in house disputes between colour realists and anti-realists. Among the colour realists, for example, *dispositionalist* or *secondary quality theorists* such McGinn (1983) and McDowell (1985) claim that colours are dispositions of objects to cause certain types of visual experience, whereas *reflectance physicalists* such as Byrne and Hilbert (2003) claim that colours are the dispositions (or microphysical bases thereof) of objects to reflect incident light in certain ways. An important distinction here is between *subjective* realists such as dispositionalists,

who hold that colours are constituted in part by relations between objects and the visual experiential responses of subjects, and *objective* realists such as physicalists, who hold that colours are constituted by the microphysical properties that metaphysically determine how objects interact with incident light. Among the colour anti-realists, *sensational anti-realists* such as Boghossian and Velleman (1989) hold that colours are qualitative properties of the visual field such as sensations or sense data, whereas *eliminativists* such as Hardin (1983, 1988a) hold that colours are properties represented in visual experiences but which are nowhere instantiated. One thing that unites subjective realists and anti-realists, and which separates them from objective realists, is that they both take colour properties to be *in some sense ontologically dependent* on the visual responses or representational capacities of perceiving subjects. Objective realists such as reflectance physicalists thus stand alone in taking colour properties to be ontologically independent of subjects.

Colour vision plays a very special role, and occupies a foundational position, in philosophical debates about colour ontology. To be specific, it is more or less universally assumed that colour vision has *epistemic* and *explanatory priority* with respect to colour ontology. To say that *X* has *epistemic* priority with respect to *Y* is to say that our route to knowledge about *Y* goes via knowledge about *X*. The epistemic priority of colour vision over colour ontology thus implies that in order to acquire knowledge as to the reality or unreality of colours we must first seek to understand our ability visually to perceive colour properties. Somewhat relatedly, to say that *X* has *explanatory* priority with respect to *Y* is to say that our rationale or justification for believing *Y* will advert or appeal to features of *X*. The explanatory priority of colour vision over colour ontology thus implies that our reasons for believing one or another account of the ontology of colour will derive from features of colour vision.

These assumptions presumably have their roots as far back as Aristotle, who held that colour was the 'special object' of vision, which is to say that it is impossible to acquire knowledge of colour via any other sense modality. In contemporary philosophy, the priority of colour vision over colour ontology finds expression in the frequently made claim that in order to know what colour properties *are*, we must first understand how colours are *represented in visual perception*. Much of the contemporary debate over colour ontology has thus centred on the question as to which properties (and relations) are represented in colour vision and whether or not these properties (relations) correspond to genuine properties (relations) of the worldly objects that we perceive. This way of framing the debate is clearly outlined by Boghossian and Velleman (1991: 106),

> The role in which colours demand attention, of course, is their role as the properties attributed to objects by a particular aspect of visual experience. They are the properties that objects appear to have when they look coloured. What philosophers want to know is whether the properties that objects thus appear to have are among the ones that they are generally agreed to have in reality. [This] is partly a question about the content of visual appearances. When philosophers ask whether colours are real, they are asking whether any of the properties acknowledged to be real are the ones attributed to an object by the experience of its looking coloured; and so they are asking, in part, which properties are represented in experience – which is a question of its content.

This assumption about the nature of the debate seems to be held across all party lines. Although they are completely opposed to the sensational anti-realism of Boghossian and Velleman, for example, reflectance physicalists Byrne and Hilbert (2003: 5) nevertheless see the relationship between colour vision and colour ontology in exactly the same light,

> The problem of colour realism, then, concerns the representational content of colour experiences. Is this content – for example, that there is a red bulgy object on the table

> – sometimes true? And what is the property red that figures in the content of such experiences?

Related although perhaps less explicit sentiments are to be found in Strawson's (1979: 109) declaration that 'colours are visibilia or they are nothing', Jackson and Pargetter's (1987: 68) claim that 'our access to colour is through our visual experience of colour', and Hilbert's (1992: 359) assertion that 'colours that cannot be seen using colour vision are not plausibly colours at all'. It is common ground among all partisans, therefore, that the issue of colour ontology turns crucially on our view of colour vision and its representational contents.

The philosophically foundational status of colour vision with respect to colour ontology makes it absolutely essential that we have some agreed standard or conception as to what colour vision *is*, or what it *consists in*. This work accordingly is predominantly concerned with improving our understanding of colour vision and only secondarily with the matter of colour ontology. In philosophy an extremely pervasive thought is that we are all able to know what colour vision is, more or less, from our *first personal*, *conscious*, *visual experience* of colour. In experiencing colour, we are in a position to know that colour vision involves having experiences of *this* type, or so the thought goes. As the above quotes make abundantly clear, most philosophers working in this area accordingly hold the *experiential view of colour vision* on which the properties (relations) represented in colour vision are those which are 'manifest', 'apparent', or 'presented', in our ordinary visual experience of colour. Given the crucial role played by colour vision in debates about colour ontology, the experiential view entails that our route to knowledge and the explanatory basis for colour ontology will derive primarily from features of our ordinary conscious experience of colour. Such is the current and historical orthodoxy in the philosophy of colour.

A key aim of this thesis is to develop and explore an alternative to the standard experiential view of colour vision, which I call the *constitutive approach* to colour vision. As I see this project it is first and foremost neither an account of the ontology of colour, nor of our conscious visual experience of colour. Constitutive accounts in general seek to describe the *nature* or *essence* of a given entity or kind. In this case, the constitutive approach seeks an account of colour vision qua *psychological kind*, the kind of representational ability or capacity in common between all actual and possible colour perceivers. The account specifies conditions which are modally necessary and sufficient for the possession of colour vision, thus establishing both upper and lower bounds on this visual ability. That is to say, the constitutive account allows us to see which abilities an organism can, and which abilities it cannot, do without in order to class as possessing colour vision. On this view, although our conscious experience of colour continues to play a crucial 'reference fixing' role in helping us to identify colour vision as a distinctive psychological ability, the constitutive approach ultimately is not beholden to specific features of ordinary visual phenomenology. The belief is that reflection on the biological, evolutionary, representational, and computational function of perception in general, and colour perception in particular, provides a more promising alternative route to understanding the constitutive natures of these visual perceptual capacities. My assumption throughout is that although it is often difficult to interpret its philosophical relevance, empirical work in colour vision science should assume centre stage in the philosophy of colour in place of first-personal introspective or phenomenological evidence. I trust that the reader, even if not convinced about the centrality of colour vision science for philosophy, will share my conviction that in an area characterised by such perennial disagreement, more sources of evidence are at least better than one.

The specific view that I defend in chapter one is that it is constitutive of colour vision that an organism possesses the ability to achieve *colour constancy*. Discussions of colour

constancy are by no means new in the philosophy of colour. One particularly important and novel feature of my account, however, is that it employs a revisionary characterisation of colour constancy as the ability to *discriminate* differences in surface reflectance properties across changes in illumination conditions. This differs from the standard 'appearance invariance view' which characterises colour constancy as the phenomenon whereby objects *phenomenally appear* the same in respect of colour across changes in illumination. Ultimately I think that these foundational theses concerning the nature of colour constancy and colour vision stand alone as contributions to the broader philosophy of perception and therefore should be judged on their own merits. If correct, however, these views provide strong abductive support for an ontological view of colour known as *reflectance physicalism*, the objective realist view that colours are type identical with surface spectral reflectance properties. Rather than trudging over the debate between advocates of reflectance physicalism and its major competitors, this ontological view will remain in the background until chapter four, where I shall consider how the account stands up against its most troubling objection, the *argument from colour similarity*.

In the second chapter I address an important objection to the appearance invariance view, which might also seem to carry over to the reflectance discrimination view. The objection is founded on the somewhat contentious claim that in standard cases of colour constancy, there is an aspect of the phenomenology of apparent colour that is illumination-dependent and hence variant across changes in illumination. This clearly causes problems for the appearance invariance view, on which colour constancy is characterised specifically by the invariance of phenomenally apparent colour across changes in illumination. I shall argue that despite initial appearances, the reflectance discrimination view is in fact much better placed to accommodate and explain this putative fact about our visual phenomenology. The reflectance discrimination view of colour constancy thus presents a coherent, explanatory,

and naturalistically plausible view of colour vision, without deifying the phenomenology of apparent colour as the definitive source of knowledge as to the nature of distinctive visual ability.

Some people might continue to question the *motivation* for the constitutive approach to colour vision. Why seek an account of the modally necessary and sufficient features of colour vision? Why emphasise these features over the ordinary introspectible features of our conscious visual experience of colour? Why think that the constitutive account tells us *all* that we need to know, philosophically if not empirically, about colour vision? By way of response, the constitutive approach is motivated in part by dissatisfaction with the privileged evidential role accorded to the visual phenomenology by the experiential view of colour vision. In particular, I have come to believe that the experiential view fails to acknowledge the inherent and typically disguised *complexity* of our conscious visual experience. There is an increasing amount of evidence that despite presenting a unified front, as it were, the phenomenology of visual experience in fact reflects the upshot of a variety of perceptual, as well as recognitory, organisational, cognitive and conceptual, processes. On this point, Pylyshyn (1999: 362) argues,

> [T]he phenomenology of visual perception might suggest that the visual system provides us with a rich panorama of meaningful objects, along with many of their properties such as their colour, shape, relative locations, and perhaps even their "affordances"... Yet phenomenology turns out to be an egregiously unreliable witness in this case. Our subjective experience of the world fails to distinguish among the various sources of this experience, whether they arise from the visual system or from our beliefs.... Phenomenology is a good source of evidence about how vision works and we would not know how to begin the study of visual perception without it. But like many other sources of evidence it has to be treated as just that, another source of evidence, not as some direct or privileged access to the output of the visual system.

As Pylyshyn indicates, much is now known about the distinctive effects of belief and expectation on the visual appearance of objects and their perceptible features such as size and colour. Dissociations between object vision and object recognition in cases of *visual agnosia* also provide important evidence as to the differential effects on ordinary phenomenology of vision and visual recognition, as in the ability to see objects as belonging to certain ordinary types, such as *chairs*, *tables*, and *knives*. Although obviously unfortunate, dissociations of this sort enable us to differentiate abilities that are indistinguishable or conflated in our ordinary visual experience of the world, and to refine our related notions of perceptual function in ways that are simply not possible from naive reflection on phenomenology.

Within the philosophy of colour, there seems to be scant appreciation of the fact that our ordinary conscious experience of colour is likely to be similarly complex. Colour is uniformly considered to be a paradigmatic 'low-level', 'simple', or 'basic' perceptible feature. The constitutive approach to colour vision seeks to remedy this misconception and to downgrade the privileged status of colour phenomenology in assessing the contribution that colour vision makes to our overall cognitive economy. In particular, by focusing on actual and counterfactual dissociations between perceptual capacities normally conflated in the phenomenology of apparent colour, the approach seeks to refine our understanding of colour vision as a distinctive explanatory kind within visual psychology. In pursuit of this aim, in chapter three I deploy the foregoing constitutive account of colour vision in arguing that a vivid and pervasive feature of our ordinary visual experience of colour is not part of the constitutive nature of colour vision. The feature in question is *categorical perception of colour*, the distinctive banding, warping, or segmenting of the colour spectrum into categories such as *red*, *yellow*, *green*, and *blue*, which is familiar from our ordinary experience of the rainbow. The proposed dissociation of categorical perception of colour from colour vision proper warrants the seemingly counterintuitive claim that colour vision by its nature does not

confer the ability to *see colours*, in the natural categorically committed sense in which we say that we see the colours *red*, *yellow*, *green*, and so on.

A second and closely related motivation for the constitutive approach to colour vision is to provide alternative foundations for colour ontology. As noted above, contemporary accounts of the ontology of colour typically begin by asking which properties (and relations) are represented in colour vision. The mainstream philosophical approach to this question goes via reflection on our conscious visual experience of colour, which evidently presents or represents colours as shades of *red*, *orange*, *yellow*, *green*, and so on. Colour ontology thus turns partly on whether these ordinary colour categories and the relational structures that they imply correspond to anything physically respectable. These assumptions are firmly embedded within the philosophy of colour. Indeed, the very *idea* of founding colour ontology on anything other than our ordinary visual experience of colour seems the ultimate sacrilege. In my opinion, this alone provides sufficient reason to consider an alternative approach. If nothing else, then, the constitutive approach simply offers a new and interesting way to think about colour vision and its relationship to colour ontology. Quite aside from this point, however, the preceding points regarding the likely complexity of colour experience provide reasons for scepticism as to the suitability of visual phenomenology as a foundation for colour ontology. Let's suppose, as seems likely, that our ordinary visual experience of colour reflects the upshot of capacities or processes which are not strictly related to the constitutive function of colour vision. If this is indeed the case, it follows that the experiential approach to understanding colour vision runs the risk of introducing erroneous or unnecessary constraints on colour ontology. To put the point another way, in focusing on broader aspects of visual phenomenology that might or might not relate to the proprietary nature of colour vision, the experiential approach threatens to obscure or frustrate attempts to understand which properties colour vision *strictly speaking* enables us to perceive.

An analogy will help to illustrate this point. Suppose that we were fairly naive about *barometers* and completely ignorant about *barometer ontology*, that is to say, ignorant about which properties barometers represent. Let's call these properties, whatever they might be, *weathers*. We want to know whether or not weathers are physically respectable. Our default approach is first to examine some *ordinary barometer appearances* in order to establish constraints on this issue. The barometers that we examine are of a fairly unsophisticated household kind, possessing a circular dial split into regions corresponding to *fair*, *dry*, *changeable*, or *stormy* conditions. These regions present 'bands' or 'segments' of weather, with associated *categorical similarity effects* among weathers at their boundaries. That is to say, given three equally spaced points on the barometer's circumference, *a*, *b*, and *c*, where *a* and *b* fall within the 'fair' region and *c* falls within the 'dry' region, say, it would appear that the weather presented by *b* is more similar to that presented by *a* than that presented by *c*, despite the equal distance between them. (As will become clear in chapter three, this case is intended to mirror the similarity effects characteristic of categorical perception for colour.) After a little empirical investigation, we find some reliable correlations between the properties presented by ordinary barometer appearances and a range of physical properties, namely *atmospheric pressures*. We thus ponder the ontological hypothesis that weathers *are* atmospheric pressures. After some thought, however, we reach a negative verdict, for two reasons. First, despite the correlations between barometer appearances and atmospheric pressure, the properties presented by the former, such as *fairness*, *dryness*, and so on, simply do not feature in any meteorological descriptions of the latter. Second, and perhaps more worryingly, the banding or categorical similarity effects observed among weathers have no analogues among atmospheric pressures. In particular, the atmospheric pressures most closely correlated with points *a*, *b*, and *c*, on our household barometer simply do not stand in the similarity relation implied by the representational structure of barometer appearances. After

some considerable labour, we thus conclude that barometers do not represent physically respectable properties.

As will become clear, the above reasoning is intended to parallel that of the argument from colour similarity against reflectance physicalism, which many people take to pose the greatest threat to the view. This argument will be the topic of discussion in chapter four. The basic thrust of the argument is that colours phenomenally appear to stand in similarity relations that simply do not correlate with the similarities that are evident among reflectance properties. This seemingly undermines the view that colours are identical with these properties, and hence undermines the view that colour vision functions to represent objective, distal, features of the environment.

Returning now to the barometer example outlined above, I hope it seems obvious that in this case ordinary barometer appearances *mislead* us as to the representational objectivity of barometers. Given what we antecedently know about barometers, we can be absolutely confident that barometers *do* represent perfectly objective features of the environment, albeit in ways that introduce unnecessary, anthropocentric, structural and organisational features. The key thought here is that our investigation into barometer ontology would have been better served by a *constitutive account* of barometers; for what is *essential* to, or *constitutive* of, being a barometer is being a device or system that represents atmospheric pressure. That's what it takes to be a barometer; that's what barometers are *for*, that is their *nature*. This constitutive account says nothing about *how* atmospheric pressure is represented, or how that information is organised, structured, or presented, by any particular barometer. The account thus abstracts from contingent and anthropocentric representational and structural features of actual barometers, such as the segmentation of the barometer face into regions corresponding to *fair*, *changeable*, etc. weather, and the implied similarity relations between the properties represented by points in these regions. In abstracting from these features, the account

provides a more secure epistemic route to barometer ontology. We should note that the constitutive account of barometers does not imply that barometers *cannot* or *do not* represent or structure information about atmospheric pressure *via* categories such as *fair* and *changeable*; clearly many do. Nor does the account imply that barometers do not in some sense '*tell us*' something about such categories. The example is intended only to illustrate the point that such representational, structural, or organisational features are superficial, non-constitutive, features of barometers and that *because of this* they should not be taken as decisive constraints on barometer ontology. This is precisely the lesson that I want to carry over to the case of colour vision and colour ontology. The hope, then, is that a constitutive account of colour vision will enable us to avoid making similarly mistaken inferences about colour ontology due to inessential anthropocentric, representational, organisational, or structural features of ordinary colour experience. This central idea will be developed and clarified further as the thesis progresses. As far as our initial presentation and motivation of the constitutive approach to colour vision is concerned, however, we are now more than ready to proceed.

1. Colour Vision, Colour Constancy, and Discrimination

The aim of this chapter is to provide an account of the constitutive nature of colour vision, an ability with which all of us are familiar but whose nature remains unclear. As discussed in the introduction to this thesis, a constitutive account of colour vision seeks to specify the conditions that are modally either necessary or sufficient for any organism, possible or actual, to qualify as possessing colour vision. Such an account aims to distil the distinctive essence or nature of colour vision *qua* psychological kind. The aim is not to explain what is distinctive about visual perception as compared with other sense modalities such as audition and touch; I shall therefore assume some intuitive understanding of the taxonomic distinctions between the different senses. The aim is rather to characterise a particular *type* of visual perceptual ability, namely that of colour vision. Colour vision is a highly distinctive ability which is ordinarily easily separable from the ability visually to perceive other features such as size, shape, texture, movement, or distance. Setting aside strange cases such as synaesthesia, while one easily could imagine visually mistaking the shape of a circle for that of an ellipse or mistaking a yellowish hue for a greenish hue, it seems difficult to imagine visually mistaking the shape of a circle for its colour. Our question, then, is what minimal constitutive standards must an organism meet in order to possess this distinctive psychological ability? What capacities must an organism possess in order to be able visually to perceive colour, and which capacities could it do without?

The chapter is organised as follows. In order to orient the discussion and fix ideas, in the first section I shall evaluate the dominant and pervasive view that colour vision is the

ability to discriminate light wavelengths. Several commentators have objected that wavelength discrimination is insufficient for colour vision because an organism could discriminate light wavelengths while lacking any capacity for visual representation of colour. I shall argue that this objection is unsuccessful as it assumes an unnecessarily thin notion of visual discrimination. I introduce an alternative cognitive and representational notion of discrimination that serves as a potentially useful explanatory paradigm in understanding visual perceptual abilities. A second and ultimately more successful objection to the wavelength discrimination view is that states of colour perception do not seem to be well correlated with the wavelength composition of light reaching the eye. On the contrary, from casual introspection it is clear that objects tend visually to appear more or less the same in respect of colour across changes in the illumination conditions in the scene.

This last point highlights the importance and centrality of *colour constancy* to our colour vision abilities, which in turn provides *prima facie* motivation for the alternative view that colour constancy itself is constitutive of colour vision. As standardly characterised, colour constancy occurs when an object phenomenally appears the same in respect of colour despite changes in the illumination conditions in the scene. I shall refer to this as the *appearance invariance view* of colour constancy or *invariantism* for short. While the phenomenology of colour constancy is helpful in providing an initial fix on the distinctive nature of colour vision, the view that colour constancy is *constitutive* of colour vision is supported by a range of further arguments based on the broader representational and computational function of visual perception, evidence from comparative and evolutionary accounts of colour vision, and considerations of parity between colour vision and other major types of visual perception. The converging evidence is that colour constancy provides by far the most plausible and explanatory constitutive basis for colour vision.

In the second section I shall argue directly that the mainstream phenomenological, appearance invariance view of colour constancy is unsatisfactory. I shall introduce a novel alternative view of colour constancy as the ability to discriminate differences in surface reflectance properties across changes in illumination.[1] I refer to this as the *reflectance discrimination view* of colour constancy. This alternative characterisation of colour constancy makes use of the cognitive and representational notion of discrimination mentioned above, which turns out to have far more explanatory power than the notion of discrimination typically assumed in colour vision science and the philosophy of colour. The reflectance discrimination view shifts our focus away from colour phenomenology or conscious colour appearances and towards the core ability to acquire perceptual knowledge or to form reliable perceptual representations of objective differences. This shift in focus presents an apparent tension in my view, for as noted above the phenomenology of apparent colour provided our initial motivation for the view that colour constancy is constitutive of colour vision. Having appealed to phenomenology in motivating this view, it might seem illegitimate thereafter to redescribe colour constancy in non-phenomenological terms. In response I shall argue that while visual phenomenology serves as an aid in *identifying* or gaining an *initial fix* colour vision as a distinctive psychological kind, an account of the *nature* of this kind need not advert to these initial identifying features. This allows us to respect the clearly important role played by visual phenomenology in accounts of visual perception, while allowing that visual perceptual kinds need not be constitutively associated with specific features of visual phenomenology. I believe that the resulting view is both philosophically rewarding and more in keeping with the assumptions of mainstream colour vision science. After briefly considering the consequences of this view for colour ontology, I shall end the chapter by considering seven objections to my view and offering replies.

[1] As we shall see later, colour vision arguably involves the ability to discriminate differences in light transmission and emission properties as well as surface reflectance properties. A minor modification to the reflectance discrimination view suffices to accommodate this point.

1.1. The Wavelength Discrimination View

The aim of this chapter is to provide an account of the constitutive nature of colour vision, a highly distinctive type of visual perceptual ability. As an entry into this topic, I shall critique the prevalent view that colour vision is the ability to discriminate light wavelength properties. To those with some basic scientific knowledge, the view that colour vision requires the ability for wavelength discrimination often seems fairly intuitive. The folk know that the colour that an object appears to have depends, in some more or less complex way, on the light reaching the eye from the object. The more informed folk also know that light is constituted by photons at various different wavelengths. It therefore seems natural to suppose that light wavelengths are the crucial properties that organisms must be able to discriminate in order to class as possessing colour vision.[2] Our first task is to consider whether this seemingly attractive view could form the basis for a constitutive account of colour vision. Is the ability to discriminate between wavelengths of light *constitutively necessary or sufficient* for the possessing of colour vision? Is wavelength discrimination part of the *nature* or an *essential feature* of the ability to perceive colour? Discussing this view will give us an initial sense of the task at hand and will allow us to introduce some key desiderata and distinctions that will be important in what follows.

Given that we are interested in characterising the nature of colour vision *qua* psychological kind, a relevant place to start our discussion would be the descriptive practice of psychologists themselves. Among vision scientists, by far the most prevalent view on this matter is that possessing some degree of wavelength discrimination ability is the gold standard for colour vision. In an introductory article on colour vision, for instance, Gordon

[2] In colour ontology, a naturally associated objective realist view would identify colours with wavelength properties of light. Hilbert (1987: 7-8) calls this the 'wavelength conception of colour'. This view seems to have been held by Russell, Ayer, and C. I. Lewis; see Hilbert (1987: 48-49). Another advocate is Armstrong (1968); see Thompson (1995b: 111-115). Purely anecdotally, this also seems to be the most common view of colour ontology among the non-philosophers that I have spoken to about colour over the last few years.

and Abramov (2001: 93) claim that 'colour vision is the ability to discriminate among different wavelengths of light regardless of their relative intensities'. Exactly the same characterisation is offered by De Valois and De Valois (1975: 115) who state that 'if colour vision is present, then one should be able to discriminate between lights of, say 640 and 520 nm even if they were perfectly matched in luminance, or if they were mismatched in either direction'. Accordingly, scientific tests for colour vision typically measure an organism's ability to discriminate differences in wavelength, with degrees of colour vision roughly corresponding to the number of wavelength discriminations that a subject is able to make.[3] At one extreme, the inability to discriminate among *any* different wavelengths, independent of their intensity, implies total colour blindness. Monochromats fall into this category.[4] On the standard picture, dichromats, trichromats, tetrachromats, and so on, exhibit increasing degrees of colour vision in accordance with their increasing wavelength discrimination abilities.

The prevalence of the wavelength discrimination criterion among vision scientists, however, does not settle the current issue regarding the constitutive nature of colour vision. This is because in vision science the wavelength discrimination criterion merely provides an *operational standard* for colour vision. That is to say, the ability to discriminate between lights of different wavelengths is taken as an easily testable behavioural standard for the possession of colour vision.[5] Hilbert (1992: 353) thus remarks that wavelength discrimination

[3] Two qualifications: first, it is not simply the number of discriminations, but also the distribution of these discriminations across the visible spectrum that is taken as relevant in evaluating a subject's degree of colour vision. I doubt we would ordinarily consider someone able to make ten million ultra fine-grained discriminations within a 1nm range, and none elsewhere, to have better colour vision than someone able to make only one million discriminations, but spread across a 300nm range. To simplify, I assume that comparisons of wavelength discrimination abilities are always across the entire visible spectrum. Second, if we take the domain of wavelengths to be the real numbers in the range 400-700nm, then every (non colour blind) subject will be able to make uncountably many wavelengths discriminations. To simplify, I assume that wavelength discrimination abilities are measured on a discrete scale of wavelength values.

[4] Hsia and Graham (1965: 206) point out that dichromats therefore should be termed 'colour confused' rather than 'colour blind', given that they are able to make some wavelength discriminations.

[5] Hardin (1988a: 148). Thompson (1995b: 56) says that colour vision is 'behaviourally defined in vision science as the ability to discriminate wavelengths'.

provides the scientific 'grounds on which colour vision is to be attributed to an organism'. This point highlights an important difference between scientific and philosophical or constitutive accounts of colour vision. As we have already noted, a constitutive account of colour vision seeks to find modally necessary or sufficient conditions which must be satisfied in order for an organism to qualify as possessing colour vision. In contrast, and simplifying somewhat, the characterisations adopted by vision scientists are intended primarily as a convenient and measurable way to assess the type or degree of colour vision possessed by some actually encountered organism. By way of analogy, consider the distinction between scientific and constitutive accounts of *consciousness*. Medical scientists rely on a battery of easily practicable tests to determine whether or not a subject is conscious, assessing behaviours such as the subject's ability to respond to verbal command, visually track an object, and so on. Although these tests undoubtedly have been selected for their effectiveness in indicating the presence of consciousness, it would be extremely contentious to assume that any of these abilities was constitutive of being conscious. It is likewise far from straightforward to assume that the operational standard of wavelength discrimination could provide the basis for a constitutive account of colour vision.

The previous point notwithstanding, it would be equally premature to dismiss the operational standard of wavelength discrimination as irrelevant to the current project. As is commonplace in philosophy and indeed the foundations of science, by starting out with some widely accepted commonsense or empirical characterisation of a phenomenon or kind, we can attempt to extract a more general, modal, account of this phenomenon via the delightful rigours of counterfactual reasoning. In the present case, many people have been persuaded by the reasoning that, for any possible organism, if that organism displayed absolutely no discriminatory sensitivity to information about the wavelengths of light, then that organism would not possess any degree of colour vision. Vision scientists and philosophers alike

therefore accept that wavelength discrimination provides at least a *necessary* constitutive condition on the possession of colour vision,

> (NEC) Necessarily, for any organism *x*, if *x* possesses colour vision, then *x* is able to discriminate between lights of different wavelengths, regardless of their relative intensities.

This principle has been endorsed explicitly by Hilbert (1992: 358), Hardin (1992: 372-375), and Matthen (1999: 78), who says that 'the processing of wavelength-sensitive data is necessary for any kind of colour concept: that much is agreed on all sides'.[6] Given its relatively uncontroversial status I shall not discuss NEC further here.

It is a matter of significant debate, however, as to whether the wavelength discrimination criterion can serve as a *sufficient* constitutive condition on possessing colour vision:

> (SUF) Necessarily, for any organism *x*, if *x* is able to discriminate between lights of different wavelengths, regardless of their relative intensities, then *x* possesses colour vision.

I believe that SUF is false and that its failure points the way towards a more plausible account of the nature of colour vision. In the following subsections, I shall discuss two prominent arguments that have been presented against SUF. Although the first argument is unsuccessful, discussing it will allow me to introduce an important distinction between two notions of discrimination which will prove useful later on. The second argument against SUF is more successful and provides some initial motivation for an alternative constitutive account on

[6] Somewhat more cryptically, Hardin (1992: 380) claims that 'the hues that we human beings see express our system of coding wavelength information'. It is important to stress that Hardin is a colour eliminativist and holds that the hues that we visually perceive are nowhere instantiated. By 'coding wavelength information' Hardin therefore must mean something other than 'representing wavelength properties'.

which colour vision is necessarily equivalent to the ability to achieve *colour constancy*. This discussion will lead us into an extended argument in support of this alternative constitutive thesis.

1.1.1 Discrimination and Representation

The first argument against SUF that I shall consider is due to Hilbert (1992) and has also been endorsed by Matthen (1999). In summary the argument is that SUF is false, for we can conceive of an organism that can discriminate differences in wavelength, but would not class as *detecting* or *representing* differences in colour and hence would not class as possessing colour vision.[7] In order to motivate this claim Hilbert (1992: 357) appeals to an actual empirical case which he takes to provide a 'variant' of the possibility just described,

> Some invertebrates display forms of behaviour that are wavelength-selective in the sense that the organism responds differently to stimuli based on their spectral power distribution and independently of the relative intensity of the stimuli. These organisms would seem to satisfy the [wavelength discrimination] criterion for the possession of colour vision. Their discriminatory behaviour may be the result, however, of their having different receptor types independently driving different response mechanisms. A receptor most sensitive to long wavelengths may be connected to a system that caused the organism to move towards the light source and a receptor most sensitive to short wavelengths may be connected to a motor system that causes the organism to move away from the light source. Over a wide range of relative intensities of a blue light and a red light the organism will always move away from the first and towards the second and consequently seem to be making a colour discrimination. Since the outputs from the two receptor types are never combined, however, there is no representation of colour or even of colour difference produced by the visual system and consequently little reason to credit these kinds of organisms with the possession of colour vision.

[7] A fundamental assumption here and throughout the thesis is that visual processes involve *mental representations* of objects and properties or property instances. As I understand them, (indicative) representational contents are characterised by their truth or veridicality conditions: a content specifies the world as being some way or other, and is true or veridical just in case the world is that way. I assume that visual processes sometimes *operate over* mental representations, in the sense that representational contents will be attributed to some states inputting to visual processes and at various intermediate stages or levels in the visual system, and that such processes will *output* mental representations, in the sense that the end product of a given visual process will be a representation of some object or feature. As far as possible, I want to remain neutral on the operative theory of representational content. I therefore shall have nothing to say on whether perceptual contents are propositions, proto-propositions, scenarios, or something altogether different.

Hilbert first claims that these creatures satisfy the wavelength discrimination criterion in virtue of their wavelength-selective or wavelength-sensitive behaviour. Hilbert then argues that it is implausible that these creatures possess colour vision given that they are unable to combine their photoreceptor responses and hence form unitary perceptual representations of colour. Matthen (1999: 52) endorses Hilbert's argument and reiterates the view that wavelength discrimination is insufficient for colour vision because 'it requires very little by way of an internal representation of colour'.[8]

In order to evaluate this argument we need to make explicit the key assumptions that lead to its conclusion.

1. Wavelength discrimination is a behavioural notion.

In the quoted passage, Hilbert refers to the 'discriminatory behaviour' exhibited by the invertebrate organism when it 'responds differentially to stimuli based on their spectral power distribution'.[9] Hilbert evidently takes this behaviour to warrant the attribution of wavelength discrimination abilities to the organism. Hilbert (1992: 357) later confirms this reading, glossing his argument as establishing 'the inadequacy of any attempt to provide a criterion for the presence of colour vision that relies only on discriminatory behaviour'. The following assumption also plays a key role in the argument:

2. Colour vision requires the ability to represent colour information, which in turn requires some means of combining or comparing the output signals of photoreceptors.

[8] Matthen (1999: 53) later on endorses the related but weaker proposal that colour vision 'is the ability to extract from light visual representations of qualities associated with its wavelength composition'.
[9] Hilbert is not alone in assuming a behavioural notion of discrimination. Hardin (1988a: 148) also assumes that discrimination is a 'behavioural' criterion. See also Noë (2004: 138). Hilbert never fully clarifies the operative notion of 'behaviour', although it is clear that he means to include neuronal responses to different wavelengths under the heading of 'differential responses' (1992: 355).

It is well known from vision science that all known biological photoreceptors respond to some degree to light at any wavelength in the visible spectrum.[10] A corollary of this is that a single photoreceptor will respond in exactly the same way to lights of different wavelengths when these lights are suitably adjusted for intensity, a fact known as the 'principle of univariance'. This is why monochromats, subjects possessing only one type of photoreceptor, are said to be completely colour blind.[11] By extension, it follows that even organisms possessing more than one type of photoreceptor – that is, photoreceptors with differing spectral sensitivities – will lack colour vision if the outputs of these different photoreceptors cannot be combined or compared in some way. This general sort of processing of receptor outputs is known as 'opponent processing'.[12] The invertebrates discussed above lack the ability for this sort of processing because the outputs of their two photoreceptor types drive separate motor responses and thus are never combined. As Hilbert (1992: 357) remarks,

> In order to possess colour vision it is necessary to be able to perceptually acquire information about what colour an object is and not just to be able to tell that two objects that differ in colour are different. The information acquired through the different photo-receptor mechanisms must interact somehow.[13]

[10] The 'visible spectrum' of light will vary across different species due to differences in the sensitivity of their photoreceptors.

[11] Hardin (1992: 372) notes that 'a few photons at a wavelength to which the cell is maximally sensitive will elicit the same response as a large number of photons at a wavelength for which it is less sensitive'. People sometimes gloss this point by saying that the output of a single photoreceptor itself is therefore 'colour blind', but for obvious reasons it is extremely misleading to attribute sightedness or blindness to photoreceptors or their signal outputs.

[12] Useful overviews are provided by Hardin (1988a: Chapter 1) and Gordon & Abramov (2001). As Hardin (1992: 374-375) notes, comparative processes of this sort 'lose information about absolute signalling rates of receptors. Instead, they preserve differences in the signalling rates of the receptors of different types that subserve the same visual area (the information required for wavelength discrimination), as well as differences between the signalling rates of cells subserving adjoining areas (the information required for spatial discrimination).'

[13] Hilbert's use of the verb 'tell' is perhaps unhelpful as it might suggest that the organism can *see* that the two objects differ in colour. Given the context the verb needs to be read in a behavioural sense.

Due to this lack of interaction between photoreceptor responses, Hilbert thinks that these invertebrates are unable to 'perceptually acquire information' about colour, which is to say that they are unable to form any representation of colour. He infers that there is scant reason to attribute colour vision to these creatures. Given that these invertebrates meet the assumed behavioural standard for wavelength discrimination, the principle SUF fails.

How could the defender of SUF respond to this argument? One might question the second assumption, the claim that opponent processing of photoreceptor signals is a necessary condition on the ability to represent colour information and hence on colour vision.[14] As Matthen (1999: 69) rightly points out, although all known biological colour vision systems do seem to exhibit opponent processing of some form, this does not establish that opponent processing is a *constitutive* or *modally* necessary condition on colour vision. Matthen (1999: 55-57) describes a fairly complex hypothetical case of a primate that lacks opponent processing but seems to possess colour vision abilities that are comparable to our own in at least some respects. Despite its obvious interest, however, it is beyond the scope of this thesis to discuss the details of this case here.

The most obvious weak spot in Hilbert's argument rather seems to be the assumption of a behavioural notion of wavelength discrimination. To discriminate is undoubtedly to do something – the verb is active – but is it to behave in some way? As with any term, people can of course define the term 'discrimination' to mean whatever they want. Williamson (1990) has argued persuasively, however, that our *ordinary* notion of discrimination has a central *cognitive* component.[15] To fail to discriminate *a* from *b* is to fail to find something out

[14] Compare Thompson (1995a: 9) and Hardin (1988a: 114), as discussed by Matthen (1999: 54).

[15] There is a pejorative sense of 'discriminate' on which some behavioural conception arguably is primary. In the pejorative sense, to discriminate against a person on the basis of their sex or race, for instance, arguably is to treat or behave towards that person differently, or otherwise act in a way that is different – and typically worse in some respect – than the way that one treats or behaves towards people of another sex or race. This is not the sense of 'discriminate' that I have in mind here. We should also distinguish the logical notion of discrimination operating in discussions of the identity of indiscernibles, as discussed by Quine (1976) on which 'two objects

about *a* and *b*, namely that they are *distinct*. One likewise cannot discriminate *a* from *b* if *a* is in fact identical to *b*. If I judge that two wines are of different types but in fact they are both Shiraz, then I have not discriminated these wines but have rather mistakenly judged the two to be distinct. As Williamson (1990: 6) notes, 'in general, a good way of refuting someone's claim to be able to discriminate between this and that is to show that this is that.' Discrimination in the ordinary sense therefore seems to be a *success* notion in that discriminating *a* from *b* entails that *a* is not identical to *b*. Discrimination is also an *intensional* or *presentation sensitive* notion. I may be able to discriminate two varietals when they are presented visually, (the wines differ slightly in colour), for example, but not when presented orally, (my taste buds lack the requisite refinement). For these reasons Williamson argues that discrimination is most plausibly equated with a type of *knowledge*; more precisely, a type of knowledge *activation* or *acquisition*.[16] Specifically, he assumes that *S* discriminates *a* from *b* just in case *S* acquires the knowledge that *a* and *b* are distinct.[17] This accounts for specific *acts* or *events* of discrimination, but a subject also can stand in the dispositional state of being *able* to discriminate *a* from *b* even if the subject has never encountered *a* and *b* and thus lacks the relevant knowledge, Williamson (1990: 6) claims, 'simply because my present cognitive capacities would permit me to discriminate if I did encounter them'. Let's call this the *knowledge account of discrimination*.

---

are discriminable if there is an open sentence, in one free variable, that is satisfied by one of the objects and not the other.'

[16] See also Martin (2004: 50) and Hellie (2005: 483). Siegel (2004: 111) offers a similar argument in favour of the knowledge account: 'suppose X and Y were discriminable by S at t just in case S was disposed to judge X distinct from Y. Then X and Y could count as discriminable for S at t, even in the case that X = Y. And this seems to stretch the notion of discriminability unacceptably far beyond the ordinary notion. If this infelicity were fixed by adding that the S at t be disposed to make a correct judgment, then it starts to look like it involves an appeal to knowledge (or something very much like it, given the link between knowledge and reliable judgment) after all.' Hardin (1988b) adopts a distinct statistical notion of discrimination that is discussed by Williamson (1990: 12-14).

[17] Although I am using objectual variables in characterising discrimination here, I assume that subjects can discriminate the same objects in different respects, and discriminate between events, properties, property instances, and perhaps more besides. This will be clarified further later.

The knowledge account of discrimination differs significantly from the operationalised behavioural notion assumed in colour vision science and adopted by Hilbert in his argument against SUF. One pertinent difference is that on the knowledge account discrimination entails the capacity for successful or veridical representation of differences in environmental features. This follows from the assumption, which I take for granted, that knowledge is a representational state. The account thus implies that organisms, plants, or objects incapable of genuine representation do not discriminate. This might seem to conflict with some ordinary applications of the verb, such as when we say that plants 'discriminate' between night and day. The knowledge account explains away such uses as analogical extensions of the central cognitive notion of discrimination. As Williamson (1990: 8) says, 'the use of the word [in such cases] is not wholly inept, for such conceptualised sensitivity to difference bears some similarity to the cognitive paradigm. Knowledge is missing, but information may be received'.[18]

Another pertinent feature of the knowledge account of discrimination is that discriminating subjects must possess some sort of first personal or whole organism level access to, or monitoring or registration of the relevant representations of environmental differences, such as would allow the acquisition of the relevant knowledge of distinctness. This follows from the fact that the knowledge account of discrimination is a *personal* or *whole organism* rather than *sub*personal account: in the ordinary sense of 'discriminate', it is subjects or persons that discriminate, and not the subsystems, organs, or any other putatively information-sensitive parts of that person. Thus while an organ such as a subject's skin might register differences in the intensity of light (think *sunburn*), the absence of the capacity for knowledge acquisition entails that her skin would not thereby *discriminate* these differences

[18] Somewhat relatedly Burge (2009, 2010) distinguishes between *perceptual* systems properly so-called, which achieve representational constancies, and merely *sensory* systems, which are systems that function solely to encode or register information received in the proximal stimulus. For more on the sensory registration/perception distinction and for several illustrative examples, see Burge (2010: 421-430).

in the paradigmatic sense at issue. It seems plausible that in paradigm cases, the relevant type of access or registration is often subserved by the organism being *phenomenally conscious* of the relevant environmental features. This feature is not mandatory, however, and I allow that discriminations can be subserved via systems of monitoring or registration that are non-conscious, even in organisms with the capacity for phenomenal consciousness.

Does the knowledge account of discrimination provide a suitable basis for theorising about *visual* discrimination? In certain respects it does, but in others it does not. On the one hand, as just noted, the requirement that discriminating subjects possess some type of first personal or whole organism access to, or monitoring or registration of the relevant representations of difference means that acts of discrimination are attributable to the *subject*, *individual*, or *whole animal*, rather than to its subsystems. This feature of discrimination is highly conducive to explaining visual perception as it allows us to respect the fact that visual perceptual states also are paradigmatically attributed to subjects or whole animals, rather than their visual subsystems.[19]

On the other hand, if visually discriminating *a* from *b* requires the capacity for knowledge that *a* and *b* are distinct, it follows that creatures lacking epistemic or rational capacities would be incapable of visually discriminating between any objects, properties, or events whatsoever. This is a somewhat unwelcome conclusion, for many people – myself included – would deny that visual perception involves conceptual or rational capacities and hence would not want to assume that visual discrimination is epistemic or cognitive in nature.[20] All the same, the foregoing discussion indicates that our ordinary notion of discrimination requires more than merely differential behavioural responses. What is needed,

[19] See Burge (2010: 369-376).
[20] This point is noted by Siegel (2004: 97-98). Martin (2004: 76) similarly notes that there must be some sense in which non-cognitive animals' experiences are discriminable from one another, but that this cannot be a cognitive sense. See Dretske (2004) for related discussion of different notions of perceptual discrimination.

then, is a notion of discrimination intermediate between the fully fledged epistemic notion and the deflationary, broadly behavioural, conception operating in vision science. I think that this project is of considerable importance in the philosophy of perception and warrants extended discussion. This will be a topic for future research. For present purposes, however, when care is needed I shall assume that *S* visually discriminates *a* from *b* just in case *S* acquires or forms, via some sort of access, monitoring, or registration system, a representation with the veridicality conditions that *a* and *b* are distinct, and this representation is veridical. A further necessary condition on any specific act of visual discrimination is that it be a manifestation of an ability that is reliable within some suitably limited range of actual and counterfactual cases.[21] This account preserves many of the central features – representation, first-personal access or registration, reliability – of the knowledge account, without the unwanted conceptual or rational baggage. It is inevitable, however, that attempts to recover the structure of the knowledge account of discrimination in non-epistemic terms will ultimately fall short in some way or other.[22] Some explanatory notions are simply so rich that it can be worthwhile continuing to use them even in cases in which the fit is less than perfect. As such, when less care is needed I shall employ the knowledge account of discrimination on the strict understanding that the view provides an explanatorily useful if ultimately inadequate basis for theorising about visual perceptual discrimination.

[21] The relevant standards of reliability will depend on context. In exceptional cases, it is plausible that an organism may class as a discriminator of some feature, even if it is *highly* unreliable in representing differences in that feature. For relevant discussion, see Burge (2010: 299-303). On this point, Burge (2010: 407) notes that 'individuals' discriminatory abilities operate in a *restricted* context of environmental alternatives. To represent specific kinds, an individual need not be able to represent principles, laws, or criteria that segment off the kinds from all other kinds consistent with the individual's discriminatory abilities. It is enough that the individual have perceptual capacities that discriminate environmental attributes within ranges that have figured causally in the formation of the states and that are relevant to biological needs and activities...'

[22] For example, the foregoing representational characterisation will be subject to Gettier-style cases. Suppose that John tokens a representation that *a* differs from *b* in respect of shape via a normally reliable visual process, and that this representation is veridical. As it happens, John is in an unusual environment, (unusual, that is, with respect to the environment in which his visual abilities evolved and normally operate), in which objects are constantly shifting in shape, but only in ways that are undetectable via normal human visual capacities. In this case, it seems intuitive to think that John fails visually to discriminate *a* from *b* in respect of shape despite the fact that he tokens a veridical representation of their distinctness via a normally reliable visual process.

Let's return to our main thread concerning Hilbert's argument against SUF. Adopting the alternative representational notion of discrimination outlined above, Hilbert's argument that wavelength discrimination is insufficient for colour vision now fails to go through. Returning to the invertebrate example discussed above, suppose that the organism tokens a state *r1* in response to the stimulation of its long wavelength receptor by light *a*, and a state *r2* in response to the stimulation of its short wavelength receptor by light *b*. The structure of the case is that the two states *r1* and *r2* drive two different motor response mechanisms *m1* and *m2*, but are never integrated or combined. Hilbert would argue via principle 2 that these organisms thus lack colour vision. Now, for these organisms to provide a counterexample to SUF, they must satisfy the antecedent wavelength discrimination condition. While these organisms satisfy a behavioural conception of wavelength discrimination given their different motor responses to *a* and *b*, however, they fail to satisfy the alternative representational notion of discrimination. On this alternative view, discriminating *a* and *b* in respect of wavelength requires tokening a representation with the veridicality conditions that *a* and *b* differ in respect of wavelength, and moreover that this representation is veridical. This condition in turn requires A) that the organism can represent the wavelength properties of *a* and *b* and B) that the organism possess a monitoring or registration system capable of combining or integrating these representations in a complex representation that the wavelength of *a* differs from the wavelength of *b*. The organisms described by Hilbert fail on both counts. Regarding point A), as Hilbert himself argues, there is no reason to attribute a capacity to represent wavelength information to these organisms. Regarding point B), even if we were to grant that states *r1* and *r2* were representational states, *ex hypothesi* the organism lacks any means of combining or integrating the information in these states and would therefore be unable to form a state with the veridicality conditions that *a* and *b* differ in

respect of wavelength. These invertebrates thus lack the ability for wavelength discrimination on the revised representational view and thus fail to provide a counterexample to SUF.[23]

Hilbert might respond by pointing out that the assumption of a behavioural notion of discrimination was motivated by the fact that advocates of the wavelength discrimination criterion for colour vision *themselves* employ such behavioural or operational standards for discrimination. Hilbert might reasonably insist that we therefore should assess the wavelength discrimination view on these terms, which would rule out the appeal to the representational notion of discrimination outlined above.[24] My view, however, is that this risks confusing the scientific and the philosophical projects of characterising colour vision. While operational or behavioural criteria for discrimination undoubtedly play a prevalent role in scientific attributions of colour vision, this is not a good reason to adopt such a behavioural notion in a philosophical or constitutive account of colour vision. In contrast, having endorsed Williamson's (1990) argument that our ordinary notion of discrimination has a central cognitive and hence representational component, it seems eminently plausible that this should be the default notion for any philosophical account of our visual abilities. I think that Hilbert should be happy, on reflection, to concede this point. Hilbert's argument is intended specifically to establish that wavelength discrimination – a particular type of discrimination – is insufficient for colour vision – a particular type of vision. If we assume that discrimination is a behavioural notion, however, then we surely would have a much simpler argument for Hilbert's conclusion: namely that *no* type of discrimination is adequate to characterise the nature of *any* visual ability. This simple argument follows from the general principle, which I assume to be well established, that there are no adequate behavioural reductions of *any*

[23] An important assumption here is that the *organism's behaviour* does not count as integrating or combining the outputs of the two distinct receptoral mechanisms in the relevant sense. The precise notion of informational integration or combination at issue here would benefit from further clarification. I thank David Hilbert for noting this point.

[24] In personal communication David Hilbert informed me that he intended in his paper (1992) to argue that the wavelength criterion was insufficient for colour vision *by the vision scientists' own lights*. He therefore took the scientists' behavioural notion of wavelength discrimination at face value.

psychological abilities; ergo there are no adequate reductions of visual abilities to discriminatory behaviour. If the wavelength discrimination view were to fall on this more general anti-behaviourist principle, Hilbert's detailed consideration of invertebrate visual systems would be dialectically redundant. It is therefore fairer not only to the wavelength discrimination view but also to Hilbert to adopt a non-behavioural notion of discrimination.[25]

Consider the following further rejoinder to the proposal that we adopt the knowledge account of discrimination in order to salvage SUF. Suppose we grant that discrimination involves knowledge of distinctness, or the reliable formation of mental representations with the veridicality conditions noted above. One might still argue that wavelength discrimination is insufficient for colour vision, because the relevant wavelength representations could be tokened in some *non-visual* system within the organism. Suppose that I lack any visual capacities whatsoever, but acquire the ability for the reliable representation of wavelength properties via some *auditory* mechanism that analyses and recodes light information into a series of sound waves. Suppose in addition that I possess the relevant abilities to combine, compare, or integrate these auditory representations into complex representations with the veridicality conditions that *a* and *b* differ in respect of wavelength. Presumably this would suffice for wavelength discrimination on the revised representational account, but it seems obvious that in this case I would lack colour vision. Similar examples could be constructed involving worlds where wavelengths and smells or tastes, for example, are perfectly correlated so that we might be able to discriminate between light wavelengths via olfaction or gustation alone, which again would be insufficient for colour vision.

In response, this rejoinder does not create further problems for our proposal, as it simply highlights an omission in our working characterisation of discrimination. Discrimination was introduced above as a three-place relation between a subject and two

[25] I thank Alex Byrne for stressing this point.

objects. This characterisation omits three additional parameters that are implicit in our ordinary attributions of discriminatory abilities. Our ordinary notion of discrimination is in fact (at least) a *six*-place relation between an organism, a time, two objects, a *respect* in which the two objects are discriminated, and a *basis* for the discrimination, such as vision, audition, olfaction, testimony, etc. The claim that *S* discriminates wine *a* from wine *b*, for example, is thus elliptical for the claim that *S* discriminates wine *a* from wine *b* at *t* in respect of acidity and texture, say, on the basis of taste. Characterised more precisely, our attributions of wavelength discrimination therefore should take the following form: *S* discriminates *a* from *b* at *t* in respect of wavelength on the basis of vision. The above rejoinder fails because it does not provide an example of an ability to discriminate light wavelengths *on the basis of vision*, but rather *on the basis of audition*, *olfaction*, or *gustation*. Wavelength discrimination in the sense at issue thus constitutively involves the ability to token representations of wavelength in a system which can antecedently be described as a visual system. Although this characterisation of wavelength discrimination appeals to a *visual basis* for discrimination, this does not render the wavelength discrimination account of colour vision circular or vacuous. As noted in the introduction to this chapter, my present aim is to provide an account of a particular and distinctive *type* of visual ability, not an account of the distinctive nature of visual perception as compared with other sense modalities. In this endeavour we therefore must be entitled to assume an antecedent notion of a visual modality. Our investigation into the nature of colour vision otherwise would regress into a discussion of the individuation of the sense modalities – an important question, no doubt, but not one that we should be expected to answer here.[26]

A closely related objection to the present proposal is that an organism might qualify as discriminating *a* from *b* in respect of wavelength on the basis of vision and yet *still* fail to

[26] On the individuation of the senses, see MacPherson (2010).

possess colour vision, because the organism discriminates differences in wavelength *via* her discrimination of some other visually perceptible feature that is completely unrelated to colour. For example, consider a world in which whenever an organism looks at two objects emitting light of different wavelengths, say, the spatial region between these objects visually appears to *wobble* at a particular frequency. In this world, organisms completely lacking the ability for colour vision nevertheless could discriminate two lights in respect of wavelength on the basis of vision *via* their ability visually to discriminate differences in the frequency of these attendant wobbles.

My response to this objection is that on the account of discrimination under discussion, discriminating *a* from *b* in respect *F* (shape, wavelength) on the basis of vision requires the organism to possess the capacity to represent determinates of the determinable feature *F* (particular shapes, wavelengths) within a system that we could antecedently describe as a visual system. It is implausible, however, to suppose that the organisms described in the above example possess the ability visually to represent wavelength properties. Their ability to discriminate wobbles seems only to require that they be visually sensitive to the frequency of the supposed spatial vibrations between objects that emit different wavelengths of light. Presumably this visual sensitivity could be achieved without any receptors or processes capable of decomposing the wavelength composition of light reaching the eye, for example, which is surely necessary for forming representations of wavelength information.[27] As such the example fails to establish that an organism could qualify as discriminating *a* from *b* in respect of wavelength on the basis of vision and yet still fail to possess colour vision.

[27] It would also be insufficient for discriminating *a* from *b* in respect of wavelength that one could tell that *a* had some wavelength but *b* did not. In this case, one would be discriminating the presence or absence of an instance of the determinable *wavelength*, whereas the notion in question requires one to discriminate between determinates of this determinable.

Another possible rejoinder to the present proposal is that the knowledge account of discrimination cannot be correct, for I can possess a discriminatory capacity even though I believe that I don't. For example, I once believed that I had an unsophisticated palette for wine and that I was completely incapable of discriminating one type of wine from another. Having recently completed a blind tasting of different varietals, however, I was pleasantly surprised to learn that I could in fact discriminate between almost all common types of wine. One might argue, however, that my belief regarding my lack of discriminatory ability would have defeated any claims to *know* that these wines were of different types, even though it seems clear enough that I successfully and reliably discriminated between them.[28]

This rejoinder fails, however, for it rests on the mistaken assumption that knowing that *p* entails that I *believe* that I know that *p*, or that being able to know that *p* entails that I *believe* that I am able to know that *p*. There are many things that I know, or that I am able to know, that I believe I do not know, or that I am incapable of knowing. I go into a maths exam firmly believing that I have forgotten and hence ceased to know the quadratic formula. To my immense surprise and pleasure, when prompted I recall the formula perfectly: I knew it all along. I go into a chemistry lesson firmly believing that I am completely incapable of acquiring scientific knowledge. To my intense surprise and satisfaction, I turn out to be the star pupil: I was able to acquire such knowledge all along. By the same token, the possibility of erroneous beliefs regarding one's own discriminations or ability to discriminate poses no barrier to the knowledge account of discrimination.

In this section I have discussed a prominent objection to SUF, the thesis that wavelength discrimination is constitutively sufficient for the possession of colour vision. The objection is that the wavelength discrimination condition could be satisfied by organisms that are incapable of representing colour information and hence unable visually to perceive

[28] This objection was put to me by David Hilbert.

colour. I have argued that this objection assumes an unduly thin behavioural notion of discrimination. Once this notion is rejected and replaced with a representational notion of discrimination, the objection is blunted. During this discussion I made some important distinctions between different notions of discrimination and clarified what I take to be the most plausible cognitive and representational account of discrimination. This account of discrimination will play an important role later in the chapter. My immediate concern, however, is to consider a second and seemingly more successful objection to SUF. This objection paves the way towards an alternative constitutive account of colour vision as the ability to achieve *colour constancy*.

1.1.2. The Centrality of Colour Constancy

A more promising objection to SUF is that the ability to discriminate between different light wavelengths could not be a sufficient constitutive condition for colour vision, because the colours that objects phenomenally appear to have typically do not correlate well with light wavelength properties. This objection requires us to introduce some more technical vocabulary. We can characterise the light reaching the eye or *proximal stimulus* via its wavelength composition or *spectral power distribution* (SPD).[29] The SPD of the proximal stimulus is a joint product of the ambient illumination, that is to say, the SPD of the primary source(s) of illumination in the scene, and the *surface spectral reflectance* (SSR) of objects in the environment, that is, the dispositions of objects to reflect incident light at certain wavelengths.[30] Due to the effect of illumination on the SPD of the proximal stimulus, the wavelength composition of light reaching the eye can vary significantly with changes in

[29] The *spectral power distribution* (SPD) of a light specifies its intensity at each wavelength in the visible spectrum, (around 400 to 700 nm), where intensity is measured by the amount of energy (in electron volts) of photons of that wavelength emitted per second.

[30] More formally, Byrne & Hilbert (1997b: xiii) note that 'the *reflectance* of an object (or surface) at a given wavelength is the ratio of the light (number of photons per unit area per unit time) it reflects at that wavelength to the incident light at that wavelength. The *surface spectral reflectance* (SSR) of an object is the reflectance of the object at each wavelength (in practice, narrow bands of wavelengths) in the visible spectrum'.

illumination across different times of day, changes in weather conditions, and so on. From introspection on visual experience, however, it seems clear that objects phenomenally appear relatively invariant in respect of colour across such changes in illumination, even when these changes are significant enough to produce discriminable differences in the wavelength composition of the proximal stimulus. This remarkable feature of colour vision is known as *approximate colour constancy*. Figure one provides an illustration of this phenomenon. In these images, a house is pictured in two different illumination conditions: in normal daylight and at dusk. While there is clearly some sense in which the two images differ in appearance, I hope you share the view that the colours of the surfaces in these images appear to remain relatively constant across the change in illumination.[31]



Figure 1: Colour Constancy in a Natural Scene (Brainard 2008)

Hilbert (1992) has argued that the approximate colour constancy of human colour vision provides strong reason to reject SUF. The argument can be spelt out as follows. If

[31] Hilbert (1992: 364-365), Thompson (1995b: 111-115). It is important to stress that colour constancy is by no means perfect. As Hilbert (2005: 143) notes, 'perfect colour constancy is not possible. There is no method that will work across all scenes and all illuminants'. We have only degrees of colour constancy. As these images make plain, objects also do not appear in *exactly* the same way under changes in illumination. This point forms the basis for an important objection to the appearance invariance view of colour constancy which I shall discuss in detail in chapter two.

wavelength discrimination were sufficient for the possession of colour vision, then discriminating a difference in wavelength typically should be sufficient for a difference in apparent colour. That is to say, discriminable changes in light wavelength typically should result in a discriminable difference in the apparent colour of an object, and consistency in light wavelength typically should result in the apparent colour of an object remaining the same. Approximate colour constancy indicates that this is not the case, at least in humans with normal colour vision abilities. In testing for colour constancy, subjects are typically presented with a uniformly painted surface that is successively, (or simultaneously, in different regions), illuminated by lights with differing wavelength compositions.[32] Things are set up so that the SPD of the light reaching the eye in each illumination condition is discriminably different. That is to say, it is guaranteed that this difference in wavelength composition will result in different outputs of the retinal photoreceptors.[33] Nonetheless, in such cases subjects typically report that the colour appearance of the surface remains the same across the change in illumination. In this case the differences in light wavelength presumably are discriminable due to the different outputs of the subject's photoreceptors, but subjects report no difference in the appearance of the colour of the surface. This indicates a systematic dissociation between a subject's discrimination of differences in light wavelength and her visual discrimination of a difference in colour. This is not at all what we should expect if wavelength discrimination were a sufficient constitutive basis for the possession of colour vision.

[32] These qualifications highlight the difference between *successive* colour constancy, which occurs when the changes in illumination are diachronic, and *simultaneous* colour constancy, which occurs when changes in illumination are synchronic. The experimental design in testing for each type of colour constancy will be subtly different.

[33] As the foregoing discussion makes clear, this is not to say that differences in photoreceptor outputs suffice for wavelength discrimination. We should rather say that the initial registration of wavelength information by these photoreceptors makes wavelength discrimination possible in a subject capable of forming visual representations of wavelength properties, and possessing the relevant first-personal access to these representations.

We could put this argument against SUF another way. It is well established that colour vision exhibits approximate colour constancy. Moreover, approximate colour constancy seems to be an absolutely central feature of our ordinary ability to perceive colour: an object's colour typically does not phenomenally appear to vary significantly under changes in illumination. Any putatively sufficient constitutive condition on the possession of colour vision therefore should be sufficient for the possession of some degree of colour constancy. However, an organism could possess even quite sophisticated wavelength discrimination abilities without being able to achieve colour constancy. Consider again the case described in the previous paragraph, in which a uniformly painted surface is presented under different illuminants, such that the wavelength composition of the proximal stimulus is discriminably different in each case. In this example, the ability for wavelength discrimination alone will not be sufficient for colour constancy, for the outputs of a wavelength discrimination system would differ significantly across the change in illumination, whereas colour constancy requires a system in which the outputs are at least approximately the same across the change in illumination. A pure wavelength discrimination system therefore would lack the capacity to produce the stable colour appearances encountered by normal subjects across changes in illumination. The ability to discriminate differences in wavelength is therefore insufficient for the possession of colour constancy abilities, and hence, by our supposition that colour constancy is central to our ordinary notion of seeing colours, for the possession of colour vision.

The above argument contends that the approximate colour constancy possessed by subjects with normal colour vision abilities provides strong reason to reject SUF. It seems extremely plausible that any putative sufficient constitutive condition on colour vision should also be a sufficient constitutive condition on colour constancy. The wavelength discrimination condition fails this test, because a subject could possess the ability to

discriminate wavelengths while lacking colour constancy. Out of this negative conclusion, however, we can draw a more positive proposal. In bringing out the centrality of colour constancy to colour vision, the argument provides strong *prima facie* reason to consider whether colour constancy *itself* could provide a more promising basis for our constitutive account of colour vision. In the remainder of this chapter, I shall make good on this *prima facie* consideration by arguing for the following principle,

(CC) Necessarily, for any organism *x*, *x* possess colour vision iff *x* possesses the ability to achieve colour constancy.

While the ordinary phenomenology of colour constancy is helpful in providing an initial fix on the distinctive nature of colour vision *qua* psychological kind, the view that colour constancy is *constitutive* of colour vision will now be supported by a range of further arguments based on the broader representational and computational function of visual perception, evidence from comparative and evolutionary accounts of colour vision, and considerations of parity between colour vision and other major types of visual perception. These arguments further demonstrate the inadequacy of the wavelength discrimination view of colour vision. As one might expect given the nature of the subject matter, my reasoning in support of the colour constancy view of colour vision will be more abductive than deductive. It therefore will be important to evaluate the overall success of the view in light of the responses to objections that I shall consider in the final section of this chapter.

A deep rationale for the colour constancy view of colour vision, as directly contrasted with the wavelength discrimination view, is that the ability for wavelength discrimination signifies sensitivity only to features of the *proximal* stimulus, namely the SPD of light reaching the eye. It is overwhelmingly plausible that visual perception functions to provide sensitivity to *distal* stimuli, however, for it is distal stimuli that are relevant to an organism's

biological ends, such as foraging, initiating or avoiding predation, seeking out a mate, locating shelter, and so on. This view of the significance of colour as a distal feature of the environment in vision science can be traced back to Helmholtz (1924: 286-287):

> Colours have their greatest significance for us in so far as they are properties of bodies and can be used as marks of identification of bodies. Hence in our observations with the sense of vision we always start out by forming a judgement about the colours of bodies, eliminating the differences of illumination by which a body is revealed to us.[34]

More recently Palmer and colleagues (2003: 314) have expressed a similar view,

> In our opinion, constancy provides the single most crucial landmark in visual processing. It is the set of visual operations whose presumed job it is to convert visual representations that encode image-based (retina-based) features into ones that encode environment-based (object-based) features... [M]any of the most crucial inferences the visual system must make are concerned with the logical leap from 2-D representations to some more ecologically useful representations that contain explicit information about properties of external, environmental objects.

In a recent wide ranging work Burge (2010) has made significant further contributions to our understanding of the significance of perceptual constancies and distal representation in psychological explanation, and outlines numerous lines of support for the claim that constancies are constitutive of our visual perceptual abilities. As Burge (2010:413) himself concludes, 'perceptual constancies are paradigmatic marks of objectification. I think that their presence in a sensory system is necessary and sufficient for the system's being a perceptual system.' Although it is beyond the scope of this thesis to review Burge's arguments, his work presents an overwhelmingly strong case for the view that perceptual constancy is *the* key

[34] As quoted by Hilbert (2005: 157).

explanatory notion in understanding the distinctive nature our visual perceptual capacities. As no doubt will be clear, the present thesis is indebted in many respects to Burge's work.

Further support for the constitutive thesis CC derives from the fact that perceptual constancies occupy centre stage in *computational theories of vision*, which have the strongest claim to the title of Kuhnian 'normal science' of all competing approaches to the scientific study of visual perception.[35] While the notion of perceptual constancy is of course not proprietary to computational theories of vision, the approach is particularly well-placed to aid understanding of constancy-related visual abilities. This is partly due to the fact that the foundational assumption of computational theories is that the most general function of visual systems is to solve the 'inverse problem' of visual optics.[36] As I assume is well known, the inverse problem is the computational or information processing problem of recovering a detailed representation of the three-dimensional layout of the objects and properties in the subject's environment from the relatively informationally impoverished two-dimensional proximal stimulation of the retinal transducers.[37] This poses a classic 'ill posed' or 'under constrained' problem for the visual system. Visual systems are thus modelled as operating under a variety of *constraints* that serve to reduce the degrees of freedom involved in this problem, which Thompson (1995b: 179) explains as follows,

> Such constraints in the model are taken to correspond to "assumptions" or "expectations" that the visual system has about the environment, the assumption being

[35] The modern computational approach began with Marr (1982). See also Rock (1985) and more recently Palmer (1999a). Early applications to the study of colour vision included Land (1986) and Maloney (1984); see the references in Hurlbert (1998). There are subtle conceptual differences between Marr's computational theory and Rock's Helmholtzian inferential theory; see Churchland (1982: 225-226) and Epstein (1989). Contrasting programs of visual research include Gibson's (1986) theory of ecological optics and enactive theories of perception as in Noë (2004).

[36] See Palmer (1999a: 23, 133), and Burge (2009).

[37] Some ecological or enactive theorists such as Gibson (1986) would deny that the proximal stimulus is informationally impoverished. This debate needn't worry us: even on the view that there is no loss of information between distal cause and proximal stimulation, the inverse problem of *extracting* information pertaining to particular aspects of the environment remains.

> that, because of evolution by natural selection, expectations about certain pervasive regularities in the physical world have been internalised in a representational form...[38]

The key assumption of such approaches is that the visual system functions to generate representations of objective features of the world given only the patterns of light impinging on the retina, subject to certain operative constraints or assumptions as to regularities in the environment. The computational theory of vision holds that the proper functioning of a visual system is marked by the achievement of perceptual constancies, which are understood as the veridical or successful representation of constant environmental features across changes in viewing conditions or perspectives, such as distance from the stimulus, viewing angle, and illumination. As Thompson (1995b: 43) notes, the achievement of colour constancy is thus 'the central information-processing problem' studied by computational theories of colour vision. The fact that mainstream computational colour vision science is centrally preoccupied with understanding the nature of colour constancy lends substantial further support to the thesis that colour constancy is part of the essence or core nature of this distinctive visual ability.

Additional support for the constitutive thesis CC derives from the fact that it ensures *parity* for colour vision with respect to other main forms of visual perception. Although it is typically not made explicit, in the philosophy of colour one often gets the sense that colour vision is seen as *exceptional* or *anomalous* in comparison with other types of visual perception such as shape or motion vision. Certainly many people would claim that colour vision differs from shape vision, for instance, in respect of the purported objectivity of its

[38] The debate is ongoing as to exactly which constraints are applied by the colour vision system in the process of achieving colour constancy. For a useful summary see the table in Hurlbert (1998: 295). An exemplar of the general constraints-based approach to modelling psychological abilities is Chomsky's (1980, 1986) work on language acquisition and syntax. The approach has been extended to features of semantics, or concept/category acquisition. For the 'constraints/principles' approach to concept acquisition, see Golinkoff, Mervis, & Hirsh-Pasek (1994), Mervis & Bertrand (1994). For a summary of competing approaches, see Hollich et al. (2000: 1-16) and Hirsh-Pasek et al. (2004).

associated representations. This sort of exceptionalism is entirely at odds with the emerging view of colour vision provided by vision science. It has been known for some time that the mammalian visual system comprises several relatively autonomous, anatomically and functionally segregated, physiological *pathways* or *channels*.[39] Each visual pathway is dedicated to processing information about a specific visual feature: form (i.e. shape and size), movement, depth, texture, and colour. The structural, organisational, functional, and physiological properties of these relatively distinct feature-processing pathways form the fundamental objects of study for cognitive, computational, and neurophysiological vision science. From this scientific perspective, the colour vision pathway or system is thus utterly unexceptional in the sense that it is simply one of several, more or less parallel, visual information processing systems dedicated to extracting information about a specific perceptible feature in the environment. It is highly notable, then, that perceptual constancies are seen as the defining feature of all other dedicated feature-processing pathways found in the visual system. Perceptual constancies are the central objects of study in computational theories of size, shape, depth, orientation, position, and motion perception.[40] It seems overwhelmingly plausible that the achievement of perceptual constancies with respect to size, shape, depth, orientation, position, and motion, is constitutive of the ability visually to perceive size, shape, depth, orientation, position, and motion, respectively. In all cases, the achievement of perceptual constancies is less than perfect. Nonetheless, perceptual constancies are taken as *the* paradigmatic mark of an objectifying visual system – a system which, while not always successful, purports to represent genuine objective features of the

[39] According to Palmer (1999a: 43) the view originates in Zeki (1978) and further physiological details were provided by Livingstone & Hubel (1987, 1988) and DeYoe & Van Essen (1988). I stress that these pathways are only *relatively* autonomous. While the various visual pathways operate more or less in parallel, there are significant interconnections between their associated, proprietary, processes. See Palmer (1999a: 40-43).

[40] For a summary of these perceptual constancies, see Palmer (1999a: 311-342; 487-504). Although 'motion constancy' might seem like a misnomer, the notion is understood as per other constancies as the ability accurately to represent the motion of objects in the environment despite variations in the movement of one's head, eyes, or body.

environment. There simply seems to be no good reason to think that colour vision is exceptional in this regard.[41] By far the most plausible view, then, is that colour vision – like all other basic visual capacities – is characterised by the ability to achieve perceptual constancy with respect to a particular, proprietary, perceptible feature.

Important further evidence for the constitutive thesis CC derives from comparative studies of colour vision. We have only started to understand the nature and role of animal colour vision in the last thirty to forty years. Our inventory of colour vision abilities instantiated across the animal kingdom is far from complete. It is nevertheless significant to note that that all animals currently known to possess colour vision seem to demonstrate some degree of colour constancy, even those species that have evolved along quite separate paths. Comparative studies have established colour constancy in the honeybee (Neumeyer, 1981; Werner et al., 1988), goldfish (Dörr and Neumeyer, 2000; Neumeyer et al., 2002), butterfly (Kinoshita and Arikawa, 2000), and diurnal and nocturnal hawkmoth (Balkenius and Kelber 2004), to name but a few.[42] Colour constancy in human and other primate subjects has been well established. Of course the merely actual commonality of colour constancy does not in itself warrant the *modal* claim that colour constancy is constitutively both necessary and sufficient for colour vision. As in all empirical investigations into the nature of kinds, however, elucidating the core features in common between all encountered instances of a kind provides one of our best guides as to its ultimate nature. On this point Neumeyer (1998:

[41] Exceptionalism with regards to colour vision presumably derives in part from the lack of clarity regarding the ontology of colour. Whereas it is relatively uncontroversial that shape and motion vision function to generate stable representations of objective features of the environment, many people would dispute that colour vision shares a similar representational function. This view strikes me as prejudicial, narrow, and as I have made clear, inconsistent with the view emerging from vision science.

[42] Neumeyer (1998: 326-29) provides descriptions of tests for colour constancy in animals.

347) concurs that the wide ranging comparative evidence cited above 'indicates that colour constancy... is indeed essential for colour vision'.[43]

A related source of evidence for the constitutive thesis CC derives from hypotheses about the evolutionary significance of colour vision. This remains an area of considerable debate and disagreement. A reasonable consensus is emerging, however, that the most important stage in the evolution of colour vision was the acquisition of abilities visually to filter or discount changes in the ambient illumination, in order to derive a more stable representation of features of objects in the environment. Maximov (2000), for example, argues that the very earliest colour vision systems were likely to be adaptations to shallow water environments in which surface waves and ripples cause constant, rapid, changes or 'flickers' in underwater illumination. Acquiring the ability to discount or filter such changes would have provided a significant advantage for predator detection, as the sudden changes in the proximal stimulus resulting from an approaching predator could be more accurately distinguished from the changes resulting merely from flickers in the illumination conditions. Discussing the evolution of later trichromatic forms of colour vision, Vorobyev (2004: 232) suggests that 'forest dwelling animals, such as many primates, probably use [i.e. evolved] colour vision to detect, segregate and discriminate unevenly illuminated objects.'[44] Colour constancy *just is* the ability to filter or discount changes in illumination, or to extract information about the uniform surface properties of unevenly illuminated objects. The view

[43] Although they agree on the centrality of colour constancy to colour vision, Thompson et al. (1998: 380) claim that colour vision also 'involves' additive colour mixture and colour contrast. They do not explain why additive colour mixture should be necessary for colour vision. As to whether colour contrast is partly constitutive of colour vision, it is interesting to note that many researchers now believe that colour constancy and colour contrast phenomena reflect a single underlying perceptual ability. The seminal article on this issue is Lotto and Purves (2000). If correct, this would eliminate some of the tension between my view of colour vision and that of Thompson et al. (1998).

[44] Also see Kelber & Osorio (2010: 1620).

expressed by these authors therefore is that the ability to achieve colour constancy is perhaps the single most significant adaptation provided by colour vision.[45]

Let's take stock. I have reviewed two arguments originating with Hilbert (1992) for the conclusion that, while wavelength discrimination provides an adequate necessary constitutive condition on the possession of colour vision, it cannot serve as a sufficient constitutive condition. The first argument was unsuccessful because it rested on an unnecessarily anaemic behavioural notion of wavelength discrimination. The second argument seems to have been more successful. The argument is that wavelength discrimination is insufficient for colour vision because it is insufficient for *colour constancy*, a highly significant feature of normal colour vision. This argument provided *prima facie* motivation for an alternative constitutive thesis on which colour vision is equivalent to the ability to achieve colour constancy. Going well beyond this initial phenomenological motivation, however, I have outlined several independent sources of support for this constitutive thesis, drawing on evidence from computational, comparative, and evolutionary accounts of colour vision.

1.2. The Reflectance Discrimination View of Colour Constancy

In the previous section colour constancy was introduced as the visual phenomenon whereby objects appear approximately the same in respect of colour across changes in illumination. Cohen (2008) labels this the 'appearance invariance view' of colour constancy or 'invariantism' for short. As Cohen (2008: 64) observes, 'invariantism has become the de facto standard understanding of colour constancy in both philosophical and scientific work on colour'. The operative notion of 'appearance' here is *phenomenal*: invariantism characterises colour constancy via an aspect of how things phenomenally appear to subjects in their

[45] Neumeyer (1998: 323) also highlights the suggestion by Campenhausen (1986) that colour vision evolved to compensate for the spectral changes of natural daylight.

conscious visual experience of colour.[46] In this section I shall argue that this characterisation is inadequate and that colour constancy is best conceived as a *discriminatory* capacity. I shall argue for the following constitutive thesis,

> (DISC) Necessarily, for any organism *x*, *x* possesses colour constancy if and only if *x* has the ability to discriminate visually between the surface spectral reflectances of objects across changes in ambient illumination.[47]

The operative notion of visual discrimination here will be the broadly cognitive, representational, notion outlined above. As we have seen, this representational notion has much more explanatory potential than the standard behavioural notion operating in vision science and much of the philosophy of colour. DISC can be split into the following two theses:

> (NEC*) Necessarily, for any organism *x*, if *x* possesses colour constancy, then *x* has the ability to discriminate visually between the surface spectral reflectances of objects across changes in ambient illumination.

> (SUF*) Necessarily, for any organism *x*, if *x* has the ability to discriminate visually between the surface spectral reflectances of objects across changes in ambient illumination, then *x* possesses colour constancy.

[46] We might thus say that the invariantist view of colour constancy is *psychophysical* rather than purely phenomenal, in that it states a relationship between an aspect of the environment – changes in illumination conditions – and an aspect of our conscious perceptual experience – the stability of phenomenally apparent colour. As should be clear, however, the putative stability of visual phenomenology forms the core part of the characterisation.

[47] It is not only *simultaneous* or *synchronic* discrimination that is at issue. A subject may discriminate *a* and *b* on *diachronic* or *successive* presentations, given the additional function of memory. The principle DISC thus raises important questions about the relationship between memory and perceptual constancy, and thus with visual perception. Walsh (1999: 13594) makes the intriguing suggestion that 'it may not make sense to think of constancy as being separate from memory: the question "is that the same colour?" need only be answered if one is matching a colour with a representation of an already seen colour.' I disagree with Walsh on this point, but further discussion must await another occasion.

As I make clear below, these theses imply that there are degrees of colour vision ability corresponding to degrees of reflectance discrimination ability. At the limit, I assume in principle that an organism capable of even the most minimal reflectance discrimination would class as possessing some degree of colour vision. In practice, however, I assume that different contexts of evaluation and attribution will set the minimal level for the possession of colour vision at some more sophisticated level of reflectance discrimination. In what follows I shall present some general arguments in support of the discriminatory conception of colour constancy, before presenting some arguments more directly in support of NEC* and SUF*. First, however, I must offer some clarificatory remarks.

A notable feature of the thesis DISC is that it makes no explicit appeal to colour appearances or our conscious visual experience of colour. The operative notion of discrimination moreover is not constitutively associated with phenomenality or consciousness, but rather with the ability to acquire knowledge, or to form contextually and counterfactually stable representations, of differences in features of the environment. DISC thus marks a significant departure from the appearance invariance view of colour constancy, which is explicitly phenomenological in character. This shift somewhat mirrors the empirical distinction between phenomenological and computational models of colour constancy. According to Hurlbert (1998: 284), whereas phenomenological models 'explicitly attempt to reproduce human colour matches under varying conditions', computational models 'attempt to transform receptor signals into surface descriptors that represent estimates of spectral reflectance and so remain constant under changing illuminant. They are designed to achieve a presumed goal of the human visual system, rather than to explicitly predict human performance on colour-matching experiments'. In a similar way, the reflectance discrimination account deemphasises the phenomenology of human colour constancy and

shifts our attention towards a representational capacity that is shared, I shall argue, to some degree by all organisms capable of colour constancy.

It is important to note, however, that the foregoing point does not imply that states or events of visual discrimination, and hence colour constancy, are necessarily non-conscious or subpersonal. On the contrary, I allow that the relevant states can be, indeed often are, states of phenomenal consciousness. As remarked above, the notion of discrimination that I employ necessarily involves some sort of first-personal or whole organism access to, monitoring or registration of, representations of the worldly objects or features in question. I allow that in paradigm cases of discrimination this first-personal access can be mediated by our phenomenally conscious awareness of the relevant features. I also allow the possibility, however, that visual discriminatory abilities can be subserved via systems of monitoring or registration that are non-conscious. The type of access, monitoring, or registration system involved in visual discrimination may be more basic than phenomenal consciousness, even in organisms capable of phenomenal consciousness.[48] The view on offer, then, is that while phenomenality or consciousness is a *highly regular* and perhaps even *paradigmatic* feature of visual discrimination in normal human subjects, it is nevertheless a *non-constitutive* feature.

When coupled with principle CC, DISC likewise entails that *colour vision* is not constitutively associated with features of visual phenomenology. In the introduction to the thesis this was billed as an attractive feature of my account of colour vision, due to the complexity of conscious visual experience and my related desire to seek alternative epistemic foundations for colour ontology. I suspect, however, that many people will object that by downplaying the role of phenomenology, we threaten to rob our prospective account of colour vision of its distinctive philosophical subject matter. As noted earlier, the basis for our

[48] Also see Burge (2005: 19). In discussing this point, Siegel (2004: 97) notes that 'it is a difficult question how to understand the nature of such access... Perhaps there is a more primitive form of introspective access of some sort'.

intimate familiarity with colour vision *qua* distinctive psychological kind is undoubtedly via our ordinary conscious experience of colour. It therefore might seem *inconceivable* that we could adequately characterise colour vision in a way that makes no explicit appeal to features of our visual phenomenology. Even more worryingly, in the previous section it seems that *my very own* proposed account of colour vision as constitutively involving colour constancy made just such an appeal. The reader will recall that in arguing against the wavelength discrimination view I claimed that it was evident from *introspection on visual experience* that colour vision centrally involves colour constancy, which was characterised as the phenomenon whereby objects phenomenally appear to have approximately the same colour across changes in illumination. This phenomenological observation was presented as a *prima facie* motivation for the constitutive thesis CC. Having adopted this phenomenological characterisation of colour constancy in *motivating* my view, however, it is surely deeply inconsistent or unstable now to claim that colour constancy is not constitutively associated with phenomenality or consciousness.

In response to this objection, I want to explain how the constitutive approach to colour vision consistently can maintain that colour vision does not constitutively involve any particular phenomenological features, while continuing to respect the inalienable role of phenomenology in helping to characterise our visual perceptual abilities. As I made clear in endorsing the quote from Pylyshyn (1999: 362) in the introduction to this thesis, I completely accept that 'phenomenology is a good source of evidence about how vision works and we would not know how to begin the study of visual perception without it'.[49] I also do not dispute that ordinary visual experience provides us with our primary initial *identifying characteristics* of colour vision as a distinctive psychological ability. As I see it, we are all

[49] In experimental contexts, conscious experiences are typically seen as providing 'feedback' on perceptual processes, (Boynton 1996: 1610), and conscious sensations serve as 'the metre', or measure, for underlying perceptual mechanisms, (Hurvich 1982: 69). Gordon & Abramov (2001: 108) likewise 'use colour sensations to constrain analyses of neuronal responses and guide creation of models'.

familiar or acquainted with colour vision *qua* psychological kind primarily via our conscious experience of colour. I argued above that the visual phenomenology of ordinary colour experience provides *prima facie* support for the view that the key psychological ability with which we are thereby acquainted is that of colour constancy. It is therefore entirely natural that our first working characterisation of colour constancy should appeal to features of our visual phenomenology.

These points are entirely compatible, however, with the view that colour constancy, and by association colour vision, are not constitutively associated with features apparent in our ordinary conscious experience of colour. In order to see this point, it is important to bear in mind the distinction between features via which we *identify* or gain an *initial fix* on a kind, and features which are *constitutive* or part of the *nature* of that kind. It is well known that explanatory kinds are often if not typically initially identified via features that ultimately turn out to be merely superficial features of these kinds. Once an initial identification has been made, an account of the *nature* of a particular kind may allow instances to be identified or hypothesised in circumstances in which the initial means of identification are absent.[50] To pick a hackneyed example, although water was initially identified via features such as being the potable clear liquid in our rivers and lakes, and continues to be identified or recognised as such in most ordinary cases, an account of the chemical nature of water enables instances of water to be identified or hypothesised in situations in which these features are absent, as in the case of non-liquid water. This is the model by which I propose that we think about the relationship between our ordinary experience of colour, on the one hand, and colour constancy and colour vision, on the other. While the phenomenology of colour constancy arguably is the main identifying characteristic by which we get an initial fix on colour vision *qua* distinctive psychological ability, our subsequent theorising about the *nature* of this kind

[50] For an extremely interesting recent application of this thought to the study of phenomenal consciousness itself, see Shea (2012).

may appeal to quite different considerations or explanatory factors. This in turn might allow instances of this visual ability to be identified or hypothesised in cases where many of the initial identifying features are absent. To return directly to the objection raised above, then, I have indeed argued that the phenomenology of colour vision provides *prima facie* motivation for my key constitutive thesis CC, but I have been careful to stress that the main support for this view derives from considerations that are independent of phenomenology. Having appealed to phenomenology in order to fix our attention on the phenomenon of colour vision, there is no inconsistency or instability in subsequently developing this constitutive thesis in a way that does not appeal explicitly to these features of the phenomenology of apparent colour. This point is absolutely central to this thesis and should be kept in mind as we progress.[51]

With these important clarificatory issues out of the way, I shall now highlight some general motivations for the discriminatory conception of colour constancy. The discriminatory conception is implicit in philosophical work on the nature of colour constancy, and perceptual constancies more generally. Consider the following quote from Burge (2010: 274),

> Perceptual constancies are capacities systematically to represent a given particular or attribute as the same despite significant variations in proximal stimulation – despite a wide variety of perspectives on the particular or attribute. Such constancies are explanatorily associated with systematic filtering mechanisms that yield sensitivity to a single environmental particular or attribute.

[51] Although it is beyond the scope of this thesis to discuss this further, the proposed characterisation of colour constancy also allows us to make sense of data from cognitive neuropsychology regarding apparent dissociations between colour constancy and visual conscious awareness of colour. Walsh (1999) discusses cognitive neuropsychological evidence that conscious awareness of colour and colour constancy are doubly dissociable. For a contrasting interpretation of this data, see Stoerig & Cowey (1997).

In this passage, Burge associates perceptual constancies with capacities to represent a given attribute, and mechanisms which 'filter' proximal information in order to 'yield sensitivity' to that attribute. As noted above, the ability to discriminate some attribute or feature is here taken to entail the ability to *represent* that feature. The ability to discriminate also clearly requires selective and reliable 'sensitivity' to changes in a given attribute. In particular, discriminatory sensitivity requires one to 'filter' the proximal stimulus so as not to mistake changes in the proximal stimulus for changes in the distal attribute. Given these points, the mechanisms for constancy to which Burge alludes should be considered discriminatory mechanisms.

Another motivation for the discriminatory view is that the ability to maintain stable colour appearances across changes in illumination could be achieved simply by assigning the very same phenomenal attribute to every object, regardless of illumination. This seems to trivialise the appearance invariance view of colour constancy. The objection is well stated by Abrams and colleagues (2007: 2611),

> [C]onstancy is useful only if colour vision also supports some amount of discrimination performance; in the absence of any requirement for discrimination, constancy can be achieved trivially by a visual system that assigns the same colour descriptor to every object in the scene.[52]

I think that this poses a clear counterexample to the appearance invariance view. At the limit, the model fails to accommodate the thought highlighted by Burge that perceptual constancies signify sensitivity to attributes of objects in the environment. It is important to note that the discriminatory conception of colour constancy faces no such problems of triviality, because

[52] See also the papers by Robilotto & Zaidi (2004) and Hillis & Brainard (2005).

discrimination is a *success* notion.[53] A visual system which assigned a different reflectance value to *every* pair of visual inputs, for example, could not sustain a capacity for discriminating reflectance properties because the system would lack the requisite counterfactual reliability in detecting which surfaces differ in reflectance. Even if the world were such that every pair of visual inputs did indeed correspond to different reflectance properties, thereby rendering the visual system's representations of differences in reflectance always veridical, the system would be liable to global error in worlds in which surfaces typically possess relatively uniform reflectance properties, as in the actual world. By analogy, I would not class as a discriminator of Shiraz and Cabernet Sauvignon by virtue of responding 'different' to every two samples of wine with which I am presented. Even if the situation was engineered so that I was always correct in these pronouncements, this would be entirely a matter of luck rather than my being reliably sensitive to the relevant differences between the two types of wine.

Another motivation for the reflectance discrimination view of colour constancy is that it performs much better than the appearance invariance view in handling the problems posed by the *argument from illumination*. I intend to discuss this argument in detail in chapter two and hence my presentation here will be brief. The invariantist characterises colour constancy as the phenomenon whereby objects phenomenally appear the same in respect of colour across changes in illumination. As will no doubt be clear, however, there is clearly *some* sense in which objects viewed across changes in illumination do appear different. Many people interpret this difference in appearance as a change in the object's *apparent colour*. If this interpretation is sound, it follows that the invariantist view is at best incomplete and at worst straight out false as a description of the phenomenology of apparent colour across changes in illumination. This objection also seems to cause problems for the reflectance

[53] In personal communication, David Hilbert objected to the reflectance discrimination view that it would face similar problems of triviality.

discrimination view. In focusing primarily on our ability to discriminate *reflectances*, the view seemingly assigns no significant role to information about the illumination conditions in understanding colour constancy. As such the reflectance discrimination view might struggle to accommodate the putatively illumination-dependent aspect of apparent colour observed in normal cases of colour constancy. In the next chapter I shall argue that the reflectance discrimination view is in fact well placed to handle this difficulty, which counts as another point in its favour.

A final motivation for the discriminatory conception of colour constancy is that when coupled with the thesis CC it provides a view of colour vision that departs only minimally from that commonly adopted in vision science. As we saw above, in empirical contexts colour vision is typically characterised via an operationalised notion of wavelength discrimination. The wavelength discrimination criterion ultimately failed as a constitutive standard for colour vision due to the fact that wavelength discrimination is insufficient for the possession of colour constancy, a particularly important and central feature of colour vision. In considering the first objection to the wavelength discrimination view above, however, we saw that once it is given a suitably robust representational characterisation, the notion of discrimination itself is perfectly well suited to the task of explicating the nature our visual perceptual abilities. The constitutive thesis DISC makes use of this revised notion of discrimination in characterising colour constancy as the ability to discriminate surface reflectance properties. In appealing to discrimination in this way, the thesis enables a philosophical characterisation of colour vision that is only a short distance away from the dominant standard operating in colour vision science. We could easily imagine this philosophical conception of colour vision being converted and adapted for empirical purposes. The resulting empirical, operationalised, standard for reflectance discrimination could serve as a behavioural proxy for the full-blown philosophical notion. This would

greatly facilitate cross-talk between philosophical and scientific accounts of colour vision and would allow philosophers to make better sense of empirical attributions of colour vision in cases in which discriminatory abilities are evident, but the character or even the presence of phenomenal consciousness of colour is unknown.

I shall now present a direct argument by contraposition for NEC*. The principle states that, necessarily, for any organism *x*, if *x* possesses colour constancy, then *x* has the ability to discriminate visually between the reflectances of objects across changes in ambient illumination. Now consider an arbitrary organism which is extremely poor at discriminating surface reflectances across changes in ambient illumination. Evaluated across a suitably limited range of differing illumination conditions, let's say, the organism is unable reliably to token or form representations with the veridicality conditions that surfaces *a* and *b* have distinct reflectances, via some sort of access, monitoring, or registration system. Of course, an organism might lack this discriminatory ability for all sorts of reasons: it might be incapable of representation, or it might lack the relevant introspective or monitoring system; indeed, it might lack any means of photoreception whatsoever. These are not relevantly challenging cases. Consider instead an organism such that whenever it is presented with the same reflectance type under different illuminants, it mistakes this for a difference in reflectance. Such an organism clearly would not be reliable in representing differences in surface reflectance properties across changes in illumination, as it would consistently yield false positives when presented with the same reflectance type. It is also clear, however, that this organism would lack any significant degree of colour constancy as assessed by *asymmetric matching tasks*, which are by far the most common method of testing colour constancy performance. In such tasks, the subject is presented with two fields with the same surface reflectance properties under different illuminants and instructed to adjust the chromaticity values of the 'test field' until they achieve a visual match with the 'standard

field'. The extent or degree of the subject's colour constancy is inferred from the amount of adjustment needed in order to match the two fields. Roughly speaking, smaller adjustments signify a high degree of colour constancy, whereas larger adjustments signify a low degree of colour constancy.[54] Normal human subjects perform relatively well on this task, although by no means perfectly: *some* adjustments are needed in almost all cases in order to achieve a match. Humans are thus good but imperfect reflectance discriminators: many changes in illumination are mistaken for changes in reflectance, but given some adjustment such mistakes can often be corrected. In contrast, our hypothetical organism would *never* be able to achieve a perceptual match, no matter how much she adjusted the test field. *Ex hypothesi*, the organism will always mistake a difference in illumination between the two fields for a difference in reflectance, and the only way to eliminate these mistakes would be to present the fields under equal illuminants. The organism will thus possess no significant degree of colour constancy by the lights of the asymmetric matching task.[55] Generalisation on this case suffices to establish the principle NEC*.[56]

[54] For a useful summary of methods for testing colour constancy, see Brainard et al (2003). I shall discuss the details of asymmetric matching experiments, and the key notion of 'chromaticity', in the next chapter. As we shall see, there are numerous difficulties in interpreting the results of these tasks in relation to philosophical accounts of colour constancy. For the time being I simply take these results at face value.

[55] It is crucial to note that assessing an organism's discrimination abilities therefore is not simply a matter of assessing which differences a subject can detect. Discrimination requires *reliability* which in turns requires the *minimisation of false positives*. This is important given that in the asymmetric matching test for colour constancy, the surface reflectance properties in both fields are identical and hence cannot be discriminated. The task nevertheless is indicative of an organism's discriminatory abilities as it demonstrates how prone the organism is to signal false positives.

[56] Although the point is often overlooked, asymmetric matching tasks assess only one aspect of colour constancy: specifically, they test constancy under conditions in which the illumination conditions change while everything else in the scene, including surface reflectance features, is held constant. Maloney (1984) calls this more limited ability *scene colour constancy*, while Gilchrist (2006) calls it *illumination-independent constancy*. They distinguish this ability from a more general notion of colour constancy in which both the illumination and surface conditions in the scene may change. As Brainard & Maloney (2011: 4) note, 'because of the conceptual prominence of illuminant changes in the formulation of the constancy problem, the bulk of research on human colour constancy has focused on illumination-independent colour constancy.' I believe that the narrow focus on illumination-independent constancy, and the associated asymmetric matching tasks, partly explains the popularity of the appearance invariance view of colour constancy, which appeals only to the stability of apparent colour across changing illumination conditions and says nothing about the ability of an organism to detect changes in reflectance.

I now present an argument by contraposition for SUF*, which is the more contentious of the two constitutive principles. I want to establish that, for some arbitrary organism lacking colour constancy, that organism would also lack the ability to discriminate visually between the reflectance properties of surfaces across changes in ambient illumination. Let's suppose that lacking colour constancy means lacking the perceptual ability reliably to filter information in the proximal stimulus to form distinct representations of the reflectance properties of objects that are relatively independent of the illumination conditions at the relevant point in the scene. (Again, there are numerous other ways to lack colour constancy: not possessing a visual system would be one such way, being dead another. These are not relevantly challenging cases.) Now consider the basis that this organism would have for discriminating reflectance properties. Its only basis for such discriminations would be information conveying the *combined* effect of surface reflectances and illumination in the scene. Given this basis, the organism would have no reliable way of determining which changes in the proximal stimulus were due to changes in reflectance, and which were due to changes in illumination. The organism therefore would be liable to mistake any difference in illumination for a difference in surface reflectance. The organism therefore would be unable reliably, or often veridically, to token or form representations with the veridicality conditions that surfaces *a* and *b* have distinct reflectances, via some sort of access, monitoring, or registration system. Generalisation on this case suffices to establish SUF*.

Consider the following objection to this argument. If the illumination conditions in an organism's environment were always constant, then the only changes detected in the proximal stimulus would be due to changes in surface reflectance properties. Under these conditions changes in the proximal stimulus would be perfectly correlated with changes in the reflectance properties in the environment. Under these circumstances an organism might lack colour constancy, (that is to say, lack the ability reliably to separate out information in

the proximal stimulus into distinct representations of the reflectance properties of objects), but still be able to detect differences in the reflectances of objects via its ability to detect differences in the wavelength composition of the proximal stimulus. That is to say, the organism could be reliably sensitive to differences in surface reflectances without being able to recover distinct representations of these properties that are relatively independent of the illumination conditions in the scene. Surely this case establishes that an organism could be able to discriminate surface reflectances and yet still lack colour constancy, which is enough to undermine the thesis SUF*.

In response, I deny that the organism described above would possess the ability to discriminate surface reflectances in the sense at issue here. Let's specify distinct surface reflectances by *R1*, *R2*, *R3*... etc., and the correlating proximal stimuli by *PS1*, *PS2*, *PS3*, ... etc. The objection is that an organism could discriminate *R1* from *R2* by virtue of being able to discriminate *PS1* from *PS2*, because *PS1* and *PS2* are perfectly correlated in that organism's environment with *R1* and *R2*. In general, however, claims of this sort are not true. In general no organism qualifies as being able to discriminate *F*s from *G*s by virtue of being able to discriminate *F**s from *G**s, where the *F*s and *F**s, and *G*s and *G**s are perfectly but only contingently correlated. For example, one arguably would not acquire the ability to discriminate renates from non-renates by learning how to discriminate cordates from non-cordates. Given that renates and cordates are actually co-extensive, one would certainly be able to sort actual animals into extensionally correct groupings of renates and non-renates on the basis of discriminating the cordates from the non-cordates. If one counterfactually were presented with two cordates of which only *one* was a renate, one would mistakenly judge that both were renates, thus failing to discriminate the renate from the non-renate. A similar point applies in the case of surface reflectances and proximal stimuli. If the hypothetical organism described above was placed in an environment in which illumination was *not* constant and

differences in the proximal stimulus therefore often did not correlate with differences in surface reflectances, then this organism would be unable reliably to detect differences in surface reflectance. This failure is crucial in the present case, for in order to qualify as possessing colour constancy an organism's reflectance discrimination abilities must be reasonably reliable *across changes in ambient illumination*. This organism's abilities are not reliable in this way and hence it does not pose a counterexample to SUF*.

Another objection to SUF* is found in Hilbert (1992). Hilbert shares the view that colour constancy is equivalent to colour vision, but would dispute that colour constancy is a discriminatory capacity on the grounds that discrimination is insufficient to establish degrees of *similarity* between reflectances (1992: 364),

> In a nutshell, the function of human colour vision is the extraction of information about the reflecting properties of surfaces from the visual stimulus. In addition, this information includes at least a partial specification of the variation of these reflecting properties with wavelength. Included in this functional characterization is the idea… that it is not enough to just get information about sameness or difference of reflecting properties. The information must at least support classification in terms of degree of similarity in a dimension independent of brightness.

By analogy, it might seem intuitive that the mere ability to represent that two lines differ in length would be insufficient for the perception of length, as if we were given lines of 2cm, 3cm, and 4cm, we would be able to detect that there are three different lengths present, but we would be completely unable to determine visually, for instance, whether the 2cm line is more similar in length to the 3cm or the 4cm line.

I offer two points in response. First, as a purely formal matter, it might be possible to derive a similarity ordering among reflectances given only the resources of a discrimination relation among them. Goodman's (1977) discussion of the 'problem of order' is highly

relevant here.[57] Goodman's problem is that of constructing a similarity ordering on a set of phenomenal qualities given only relations of sameness ('matching') and difference ('non-matching') between these qualities. In brief summary, Goodman shows how it is possible to define formal relations of comparative similarity given these scant resources. The technical details of this proposal are beyond the scope of this thesis. We might insist, however, that there is no problem *in principle* with constructing a similarity ordering among reflectances given only relations of (in)discriminability among them. It is a moot point whether the resulting similarity ordering would reflect our intuitive sense of visual similarity among these features. As we shall see in chapter three, I believe that many phenomenally apparent similarities among the colours are not linked to our fundamental, constitutive, discriminatory colour vision abilities. Nevertheless, if *some* similarity ordering among the reflectances is all that is required in order to qualify as possessing colour vision, then Goodman's formal apparatus will provide the resources to respond to Hilbert's objection.

A second response is that discrimination as understood here is more demanding than simply 'getting information about the difference of reflecting properties'. On my view discrimination is not just a matter of difference detection, as Hilbert seems to assume.[58] Crucially I have argued that visually discriminating surface reflectances entails the ability to *visually represent* these properties. In response to Hilbert's objection, I would be happy to grant that it is a necessary condition on visually representing some feature *F* (shape, reflectance), and hence on discriminating among *Fs* (shapes, reflectances), that one is able to represent *at least some* similarities among *Fs*. To put the point another way, it presents no problem for my view to concede that the visual representation of *F* entails the existence of

[57] Goodman (1977), especially chapters IX and X. Goodman's discussion centres on the individuation of phenomenal qualities, which should be clearly distinguished from the present issue.
[58] Once again Hilbert (1992: 357) clearly has in mind a deflationary behavioural notion of discrimination.

some or other abstract *quality space* in which our representations of *F* can be modelled.[59] What I would deny, and what Hilbert's objection does not establish, is that there is any *particular* or *privileged* quality space in which our visual representations of *F* necessarily must be represented. In particular nothing that Hilbert says implies that our visual representation of *colour* must preserve the distinctive differential similarity effects associated with our ordinary visual perception of the colour categories *red*, *orange*, *yellow*, and so forth.[60] This point will become relevant in chapter three, where I shall argue that the similarity effects characteristic of categorical perception of colour are not associated with our constitutive colour vision abilities.

In concluding this section, henceforth I shall assume that the possession of colour constancy, the ability to discriminate between the surface reflectance properties of objects across changes in ambient illumination, and the possession of colour vision, are all pairwise necessarily materially equivalent. As noted above, an attendant advantage of this view is that the degree of colour vision possessed by an organism can be characterised via the degree of reflectance discrimination ability that the organism possesses. We can capture this thought in the following principle:

> (DEG) Necessarily, for any organisms *x*, *y*, *x* and *y* are alike in respect of their colour vision abilities iff *x* and *y* are alike in respect of the visual discriminations they are able to make between surface reflectance properties across changes in illumination.

This principle will play a central role in chapter three.

[59] Note that this is a *metarepresentational* condition on visual representation: the quality space is a topological representation of the structure implied by our system of visual representation, and is not itself a medium of visual representation.
[60] For related comments, see Matthen (1999: 65, 69, 78).

1.3. The Consequences for Colour Ontology

As noted in the introduction to this chapter, the view developed here has clear consequences for the ontology of colour. The ontological view most naturally associated with this view of colour vision is *reflectance physicalism*, which can be loosely characterised by the three following features.

1. Colours are *surface spectral reflectance properties*.

As we have seen, the surface spectral reflectance of an object is its disposition to reflect incident light at varying wavelengths. Reflectances feature prominently in optics, the physics of light and its interaction with matter, hence the label reflectance *physicalism*.[61]

2. Colours are not in any sense ontologically dependent on the visual responses or representational capacities of perceiving subjects.

As noted in the introduction to this thesis, reflectance physicalists stand completely apart from both subjective realists and anti-realists about colour in holding that colour properties are entirely ontologically independent of subjects. For example, reflectance physicalists would reject the mainstream *dispositionalist* claim that necessarily, for all *x*, *x* is red iff *x* is disposed to produce a certain reddish kind of visual experience in normal perceivers in normal contexts.[62] The reflectance physicalist takes colours to be subject-independent properties, akin to shape, size, mass, and other so-called *primary qualities*.

[61] Hilbert (1987), Byrne & Hilbert (2003). For detailed discussion of taxonomic issues in philosophy of colour, see Cohen (2009: Ch. 1). The label 'physicalism' is perhaps unfortunate, as it might suggest that rejecting the view entails rejecting physicalism *simpliciter*. I assume, however, that all other candidate theories of colour are compatible with metaphysical physicalism. For example, even on the anti-realist subjectivist view that colours are properties of neural events, physicalism about the mental would suffice for a form of physicalism about colour.

[62] Byrne & Hilbert (1997a: xx-xxii) provide a brief overview of dispositionalist theories of colour. On the contrast between classic dispositionalist and *functionalist* theories of colour, see Cohen (2009: Ch. 8.1). *Realiser* forms of colour functionalism use subject-involving descriptions or functional roles to fix the reference of colour terms on their physical realisers, with which the colours are strictly identical (Jackson & Pargetter 1987; Jackson 1998; McLaughlin 2000). If these realisers turn out to be surface reflectance properties, and if the

3. Colours are *type identical with* physical properties.

Reflectance physicalism rules out weaker relations of metaphysical determination, supervenience, or token identity between the colours and the physical.[63] The general claim is that every determinate of the determinable *colour* is type identical with some determinate of the determinable *surface reflectance*. Not all reflectance physicalists will assume the converse claim, however, that every determinate of the determinable *surface reflectance* is type identical with some determinate of the determinable *colour*: some reflectance physicalists allow for surface reflectance properties that are not colours.[64] By analogy, type identity theories of mind claim that every determinate of the determinable *mental state* is type identical with some determinate of the determinable *brain state*, but can allow that not every determinate of the determinable *brain state* is type identical with some determinate of the determinable *mental state*.

It has been noted in various places in the literature that the phenomenon of colour constancy provides one of the main motivations for reflectance physicalism. As Cohen (2008: 61) points out,

> All the vocal proponents of [reflectance physicalism]... have taken colour constancy as a central pillar of support, and even those who reject the view typically allow that colour constancy is something they need to discount or play down in order to meet their opposition...

---

relevant identities are *type* rather than *token* identities, then realiser functionalism will end up looking very much like the reflectance physicalist's type identity theory; see Cohen (2009: 187-188). One might also compare Braddon-Mitchell & Jackson (1996) on realiser functionalism as a route to the identity theory of mind. *Role* functionalism takes colours to be identical with subject-involving functional roles (Cohen 2009: Ch. 7). Given the central role played by subject responses in individuating colour properties, role functionalism is not in my view a form of reflectance physicalism. See Cohen (2003: §2) for contrasting discussion.

[63] Token identities are compatible with forms of colour functionalism. Colour *primitivism* is a view that takes colours to be merely supervenient on physical properties. See Campbell (1993).

[64] This depends partly on whether one thinks there are colours that no human could perceive. See Byrne & Hilbert (2003: 20, fn. 28).

To date all arguments for reflectance physicalism of this form have assumed the appearance invariance view of colour constancy, on which objects appear approximately the same in respect of colour across changes in illumination. This view indicates that apparent colour is correlated with some illumination-independent property. The argument from colour constancy to reflectance physicalism then proceeds via two inferences to the best explanation. First, it seems reasonable to infer that if the phenomenological quality of apparent colour correlates with illumination-independent properties, then *colours* themselves are a type of illumination-independent property. Second, if colour properties are illumination-independent properties of some sort, then the most natural candidates are surface spectral reflectances. Tye (2000: 147-148) thus argues that 'the fact that objects appear to retain the same colour through a wide variety of changes in illumination conditions (though certainly not all) strongly suggests that colours are illumination-independent properties of those objects'. In a similar vein, Hilbert (1987: 65) argues that 'the existence of colour constancy should lead us to suspect the existence of some illumination independent property of objects that is correlated with colour... The obvious candidate for such a property is... surface spectral reflectance'.[65] Byrne and Hilbert (2003: 9) likewise argue that in assessing candidate ontologies for colour,

> We can narrow the field further by noting that the colour vision of human beings and many other organisms exhibits approximate colour constancy... for instance, tomatoes do not seem to change colour when they are taken from a sunny vegetable patch into a kitchen illuminated with incandescent light. Assuming that our perceptions of colour are often veridical, we therefore need a physical property of objects that is largely illumination-independent—a physical property that an object can retain through changes in illumination... The property that initially suggests itself is surface spectral reflectance: the proportion of incident light the object is disposed to reflect at each wavelength in the visible spectrum.

[65] This presentation of the argument and the related quotations are borrowed from Cohen (2008: 76-77).

All of these authors make the same inference from the relative illumination-independence of apparent colour to the probably illumination-independent of colour, and from this to the probable identity of colours with surface reflectance properties.

My primary intention in this chapter has been to develop an account of the nature of colour vision and the issue of colour ontology therefore has been relegated to the background. Having placed colour constancy centre stage in my account, however, it is predictable that my sympathies should lie with reflectance physicalism, which I take to be the best candidate view of the ontology of colour. Unlike other reflectance physicalists, however, I reject the appearance invariance view of colour constancy and hence the route from my account of colour vision to reflectance physicalism must take a subtly different form to the standard argument outlined above. On my view, colour constancy and hence colour vision constitutively involve the ability to discriminate, and thus reliably form representations as of, the reflectance properties of objects in the environment.[66] It is natural to infer that if colour vision constitutively involves the ability to discriminate some particular range of properties, then these very properties should be identified with the colours. In this regard, my account of the nature of colour vision provides an even more direct argument for reflectance physicalism than the standard route *via* the appearance invariance view of colour constancy. As such the fates of the two views are tied even more closely together. To some people this might seem to be an unattractive feature of the view. I do not consider this close association to be a disadvantage, however, any more than I would consider it a disadvantage of our view of visual perceptual shape constancy that it should be inherently tied to a background theory of physicalism about shape properties. In my view the very notion of perceptual constancy presupposes a background of objective distal properties that visual systems function to recover. Without this presupposition, it would be difficult if not impossible to understand the

[66] As we shall see shortly, this view must be modified slightly in order to accommodate the colours of light transmitting and emitting bodies.

biological and evolutionary significance of perceptual constancy in enabling organisms to adapt, respond, and interact with objects in the environment in the service of their basic biological needs.[67] Standard phenomenological accounts of perceptual constancy do not fully embrace or respect this fact. In contrast my account of colour constancy accommodates this fact via the proposed constitutive link with *visual discrimination*, which I have understood to be a *success* notion. Visual discrimination thus entails knowledge, or at least reliably veridical representation, of differences in the features of objects in the environment. From this standpoint the adoption of a physicalist account of colour is entirely natural although admittedly not mandatory. Many others undoubtedly will espouse different theoretical priorities and adopt different emphases. As I have made clear, however, in this thesis I do not intend to enter into repetitious discussion of competing, alternative, views of colour ontology. My strategy rather will be to take this putative motivation for reflectance physicalism at face value and then in chapter four to consider how well the view stands up against a particularly troubling objection that is posed by the *argument from colour similarity*. Ultimately I shall argue that this objection is not nearly as troubling as is ordinarily assumed. This greatly increases the plausibility of reflectance physicalism and by association the account of colour constancy and colour vision that I have defended in this chapter. I shall now move on to consider objections to the constancy based constitutive account of colour vision defended above.

1.4. Objections and Replies

1.4.1. Non-Reflectance Colours

Objection: On the proposed account, colour vision constitutively involves the ability reliably to form representations of differences in the surface spectral reflectances of objects. The

[67] For extended discussion, see Burge (2010).

closely associated ontological thesis, *reflectance physicalism*, implies that the surfaces of opaque light-reflecting objects are the only things which truly have colours and that their spectral reflectances are the only properties relevant to determining what colours they have. This view is at odds with the sizable range of materials, gases, and fluids which we perceive as having colours and the existence of numerous different causal bases of our perception of colour besides the reflection of light by surfaces. Examples include crystals such as ruby and emerald, where perceived colour is determined largely by the selective transmission of light; radiation effects, such as flame incandescence, where perceived colour is determined by the emission of light, as opposed to reflection or transmission; scattering effects, such as the perceived blueness of the sky, which results from the diffraction behaviour of molecules and dust particles in the atmosphere; and optical effects such as rainbows, which result from the dispersive refraction of light by raindrops.[68] As Thompson (1995b: 104) complains, the account therefore 'play[s] favourites among the different ways that colours can be encountered', privileging the reflection of light by opaque surfaces over the many other bases for our perception of colour.[69]

Reply: One possible response is that the foregoing examples of non-surface, non-reflectance based colours all involve *misperceptions* or *visual illusions* of colour. The visually apparent colours of transparent objects, gases, fluids, or optical phenomena, are explained away as cases in which our visual systems 'misinterpret' the nature of the distal stimuli in the scene, mistakenly attributing reflectance properties either to surfaces that do not possess these properties or perhaps even when no surfaces are present. Matthen (1988: 25), for example, adopts this line in arguing that we misperceive the sky as blue because the sky is

[68] Nassau (1980).
[69] For similar points see Broackes (1992: 217) and Hardin (1988a: 2-7).

not a surface and therefore does not have whatever surface spectral reflectance properties our perceptual states attribute to it.[70] Let's call this the *chauvinistic* response.

The chauvinistic response preserves the straightforward reflectance discrimination of colour constancy and the associated physicalist ontology, but at the cost of reducing their plausibility in accounting for the nature of colour vision and colour ontology. Throughout the chapter so far I have sought to promote the virtues of the constancy view of colour vision by appealing to comparative, biological, and evolutionary evidence. I think that all parties to this debate would agree that surface reflectance properties play a crucial role in many biologically significant activities of organisms and therefore would agree that surface reflectance properties *at a minimum* figure among the causally relevant bases of our visual perception of colour. It remains an open empirical question, however, as to whether colour vision evolved to enable the detection and representation of a much *wider* range of properties, including the selective transmission of light by fluids and gases, the emission or radiation of light, and the scattering of light by atmospheric particles. Rather than give hostage to fortune over the outcome of this ongoing empirical debate, it might be wise to consider a more inclusive approach.[71]

Fortunately I believe that a fairly straightforward modification to the reflectance discrimination view of colour constancy and reflectance physicalism will suffice to accommodate this point. To be specific, the proposal is to extend the standard characterisation of surface reflectance properties to include other ways that surfaces, liquids, or gases can be disposed to interact with incident light. Byrne and Hilbert (2003: 11-12) outline the approach as follows,

[70] Matthen's response in discussed in Thompson (1995b: 181).
[71] As standardly conceived, computational theories of colour constancy do not extend to cover the perception of the colours of crystals, metals, and other such materials. See Hurlbert (1998: 291). This seems to be a feature of simplification rather than a theoretical commitment to reflectance focused views of colour constancy.

> Earlier, we gave a standard definition of reflectance: the proportion of incident light the object is disposed to reflect at each wavelength in the visible spectrum. However, we could just as well have characterised reflectance slightly differently, in terms of the light that would leave the object, rather than the light that the object would reflect. [Let's say] that the productance of a surface is its disposition to produce (i.e., reflect or emit or transmit) a specific proportion of incident light. For opaque non-luminous surfaces this will be equivalent to the original definition of reflectance in terms of reflected light. For surfaces that emit or transmit light, however, the productance and the reflectance will sharply diverge.

In summary, the *productance* of an object is a dispositional property that determines the amount of light leaving that object at each wavelength in the visible spectrum, as a function of the reflection, transmission, and emission of light incident on the surface of the object. Like surface spectral reflectance properties, productances will be supervenient on the microphysical properties of a surface, liquid, or gas, that determine how it will interact with light at varying wavelengths. The productance of an object is also illumination-independent in the sense that the microphysical properties that determine how much incident light an object will reflect, transmit, or emit remain constant across changes in ambient illumination conditions. The proposal, then, is that the official view of colour constancy could be modified so as to involve the discrimination of differences in productance across changes in illumination. The resultant ontological thesis would be *productance physicalism*, on which colour properties are type-identified with productances. This relatively straightforward modification allows us to accommodate the view that objects such as crystals, transparent or translucent materials and fluids which act as filters, and light emitting phenomena such as fluorescents, are all coloured.[72]

[72] For objections to the productance proposal, see Jakab & McLaughlin (2003: 34-35), Mausfeld & Niederée (2003: 41), Decock & van Brakel (2003: 27), Dedrick (2003: 28), and Kuehni (2003: 36). Byrne & Hilbert (2003: 53-54) offer replies. This proposal doesn't fully avoid the charge of chauvinism because productances do not extend to phenomena dependent on the scattering, diffraction, or refraction of light, such as the blue sky or

There remain significant details of the productance proposal that have yet to be worked out. One particular concern is that it is impossible to predict whether biological visual systems are able to recover representations of productances from the information contained in the proximal stimulus, or whether the relevant inverse problem is simply insoluble for such subjects.[73] Whatever the outcome of research into this issue, we can be confident that the resulting computational account will differ markedly from the standard reflectance based models of colour constancy. To pick one example, all but one of the standard computational models described in Hurlbert (1998) assume that colour vision systems operate under the constraint that there is only a single source of illumination in the scene. As Byrne and Hilbert (2003: 12) point out, however, in order to determine the contribution of *light transmission* properties the visual system must assume that the object is illuminated from the front *and* back relative to the viewer. If the object were illuminated only from the front, for example, then the light reaching the eye would depend solely on reflectance and not transmittance and if the object were illuminated only from the back, the converse would be true. The assumed presence of two distinct sources of illumination in the scene significantly increases the difficulty of the computational problem which the visual system must solve in order to generate a representation of productance properties.[74] The productance proposal therefore presents greater theoretical challenges to colour vision science than reflectance based accounts of colour constancy. These issues will form the basis for further research. Throughout the remainder of this thesis, however, I shall assume that some such extension of the reflectance discrimination account and reflectance physicalism will be possible.

---

rainbows. Byrne & Hilbert (2003: 20, fn. 24) concede that such phenomena will 'often produce colour illusions', that is, mistaken or non-veridical attributions of productance properties to objects or media which lack such properties. The view undoubtedly constitutes a step in the right direction, however, and accommodates the vast majority of apparently coloured objects.

[73] Byrne and Hilbert (2003: 54) thus 'wouldn't presume to say whether productance, as defined by us, will be useful to vision scientists'.

[74] One strategy might be to retain the assumption of a single light source, adding the further assumption that illumination is constant across all points in the scene, (Hurlbert 1998: 290, 304).

1.4.2. Dichromatic Colour Vision

Objection: Several authors (Maloney and Wandell 1986; Shepard 1987, 1990) have argued that colour constancy requires *trichromacy*, the possession of at least three different types of retinal photoreceptor. If it is correct that colour constancy is essential for the possession of colour vision, it follows that dichromats do not possess colour vision by virtue of their inability to achieve colour constancy.[75] This consequence is extremely implausible, as we know from both behavioural studies and qualitative, phenomenological, accounts of dichromacy in humans that such subjects retain significant colour vision abilities.[76] Comparative behavioural and neurophysiological studies also indicate that colour vision is present in dichromatic animals such as squirrels, rabbits, tree shrews, possibly cats and dogs, and some monkeys, (Thompson 1995b: 145). The view that colour constancy is constitutive of colour vision thus makes radically false predictions as to the distribution of colour vision among human subjects and the animal kingdom at large.

Reply: I agree that it would be both implausible and contrary to current scientific opinion to hold that dichromatic humans, squirrels, rabbits, and so on, lack colour vision. The key issue is whether these organisms possess colour constancy. Fortunately for my view and contrary to the above assertion, there is strong evidence to suggest that dichromats achieve a degree of colour constancy not significantly different to that achieved by trichromats. Rüttiger and colleagues (2001) compared the colour constancy performance of subjects with red-green colour blindness, including the two forms of dichromacy protanopia and deuteranopia, to that of normal trichromatic subjects. They found that although the dichromatic subjects did perform worse than trichromats on colour constancy tasks, the difference was not nearly as large as we might expect given the reduced spectral sensitivity

[75] This conclusion is tentatively accepted by Hilbert (1992: 365), and discussed in Thompson (1995b: 194).
[76] The question of which colours, if any, are seen by colour blind subjects is a delicate one. For discussion, see Byrne & Hilbert (2010) and Broackes (2010).

implied by their reduced number of photoreceptor types.[77] These results undermine the sweeping claim that at least three photoreceptor types are necessary for colour constancy and relieve the pressure on the proposed association between colour constancy and colour vision. While brute deficits in retinal physiology clearly undermine colour constancy performance to some degree, there are numerous other post-receptoral mechanisms and computational strategies which contribute to the achievement of colour constancy and which plausibly would be retained by dichromatic organisms.[78] This is entirely consistent with the present constitutive thesis, which allows that colour constancy *qua* reflectance discrimination, and hence colour vision, comes in degrees.

By way of analogy, the relationship between dichromatic and trichromatic colour constancy abilities might be similar to the relationship between monocular and binocular organisms in respect of depth constancy. Like colour constancy, the achievement of depth constancy in biological systems is multifaceted and makes use of a wide range of mechanisms and strategies to form representations of the relational distance of a point or object relative to the organism. One important mechanism is the analysis of binocular disparity, the relative displacement of the images projected from a distal object onto the left and right retinas, which provides important information as to how far away the object is from the subject. We can therefore conjecture that monocular subjects would achieve a degree of depth constancy inferior to that of binocular subjects due to the absence of this source of depth information. Nobody would claim, however, that monocular subjects completely lack depth constancy by virtue of lacking just one of the many mechanisms implicated in this visual perceptual ability. Other cues not related to stereoscopic information such as motion

[77] See also Foster & Linnell (1995), Foster et al. (2003), and Amano et al. (2003). Using different techniques, Morland et al. (1997) showed systemic deficits in the colour constancy performance of protans and deutans. The task sensitivity of these results suggests that the situation, unsurprisingly, is more complicated than we have made out.

[78] See Smithson (2005) for a helpful review of retinal and post-retinal processes involved in colour constancy.

parallax, convergence of parallel lines, texture gradients, and shading information, would remain in place to contribute to the achievement of depth constancy by the monocular subject.[79] Presumably something similar holds in the case of colour constancy in dichromatic subjects. Although lacking an obvious and important source of information as to differences in the reflectance properties in the scene, dichromats retain a battery of other strategies that seem to allow them to achieve a significant degree of colour constancy.

1.4.3. 'Disregarding' the Illuminant

Objection: On the standard computational view of colour constancy, colour vision systems 'discard' or 'disregard' information about the illumination conditions contained in the proximal stimulus in the process of extracting a representation of the stable surface reflectance properties in the scene. As Hilbert (2005: 157) notes, on such views 'the illuminant is merely a means of revealing stable surface properties and not a property of interest in its own right'. Thompson (1995b: 102-103) objects that this underplays the significance of illumination conditions to organisms in their natural environments,

> Many computational approaches seem to rest on the assumption that colour vision is concerned primarily with the reflecting properties of surfaces. As a result, estimates of the illuminant are often treated largely as something to be "discounted" in the task of recovering reflectance. This "surface-fixated computational viewpoint"... neglects the point that natural colour vision might be concerned with illumination conditions in their own right, for these provide indications about weather conditions, time of day, and so forth...[80]

By association the reflectance discrimination view of colour constancy likewise fails to acknowledge the biological significance of information about illumination conditions. The

[79] For further details, see Palmer (1999a: 200-253).
[80] The objection seems to be adapted from Jameson & Hurvich (1989: 178).

view therefore forms an inadequate basis for theorising about colour vision *qua* biologically relevant kind.

Reply: This objection conflates two distinct claims. The first claim is that standard reflectance based models of colour constancy imply that information about the illumination conditions is *completely discarded by the visual system at large* and as such plays no role whatsoever in explaining *any* of our visual capacities. The second claim is that standard reflectance based models of colour constancy imply that information about the illumination conditions is discarded by the *colour vision system* and as such plays no role in explaining our ability to perceive *colour*. At the very worst, the frequently made claim that colour vision systems 'discard' or 'disregard' information about the illumination conditions commits us to the second claim. The objection seems to assume, however, that reflectance based accounts of colour constancy are committed to the stronger first claim. This is clearly mistaken. Even if the *colour vision system* functions to 'discard' information about the illuminant in forming representations of reflectance, this clearly leaves room for the view that information about the illumination conditions plays a significant role in the visual system *at large*. It seems clear, for example, that some aspects of the illumination conditions are represented within the *luminance* vision system, a distinct visual pathway which functions to extract properties of the intensity of light.[81] The luminance vision system is implicated in the perception of shading and shadows cast by objects, which in turn contributes to our perception of object form, depth and orientation, (Palmer 1999a: 184-185, 243-246). A crucial task for the visual

[81] The anatomical and functional separation of colour and luminance information seems to occur very early in the visual system, with colour being processed largely in the pathway leading from parvocellular layers of the lateral geniculate nucleus (LGN), whereas luminance is processed in the pathway leading from magnocellular layers of the LGN. See Gordon & Abramov (2001: 105-106) and Hubel & Livingstone (1987). Van Essen & DeYoe (1995) stress the difficulties in making a clean functional separation of the visual pathways. We can of course expect there to be significant anatomical overlap between the colour and luminance systems. As throughout the whole visual system, we can also expect significant sharing of information between distinct, parallel, feature-based visual processes. The ultimate computational story may well involve complex interactions between algorithms for the extraction of colour (reflectance/productance) information and luminance information.

system is to distinguish between changes in luminance or light intensity resulting from changes in the reflectance properties of a surface, (so called *reflectance edges*), and changes corresponding to different amounts of light falling on a surface of homogeneous reflectance, (*illumination edges*), such as are found at shadow edges. In completing this task, it is eminently plausible that the visual system employs not only representations of the probable reflectance conditions across a surface but also representations of the likely illumination conditions in the scene. Palmer (1999a: 132), for example, claims that the visual system employs 'illumination maps' for such computations, which I take to be representations of the point-by-point light intensity across the visual scene. There is simply no reason to suggest that information about illumination is 'discarded' by the visual system *per se*. The claims made by advocates of reflectance based accounts of colour constancy about 'discarding' illumination information need to be viewed in this broader context and accordingly taken with a large pinch of salt.

As noted above, at the very worst the claim that colour vision systems 'discard' or 'disregard' information about the illumination conditions commits us to the view that this information plays no substantial role in the functioning of the *colour vision system* and as such no role in explaining our ability to perceive *colour*. In the next chapter I shall consider whether even this weaker consequence is acceptable. Several authors have objected to this view that colour appearances and hence our visual representations of colour are partly illumination-dependent. If sound this objection provides *prima facie* reason to think that information about the illuminant plays a significant role within the colour vision system itself, which puts pressure on all reflectance based accounts of colour constancy, including the reflectance discrimination account.

1.4.4. Access to Spatial Representations

Objection: The reflectance discrimination view holds that colour constancy constitutively involves the ability to represent the reflectance properties of the surfaces of objects in the environment. This view presupposes that the processes involved in achieving colour constancy have access to representations of the spatial layout of the scene, in particular representations that segment the scene into regions corresponding to distinct surfaces and objects. The problem with this view is that colour information itself is known to contribute to processes of visual perceptual discrimination and segmentation of objects in the scene. As Thompson (1995b: 92) notes, 'to assume that the scene has been segmented prior to the deployment of the algorithms for colour constancy... misses in part the role that colour vision plays in object discrimination and identification'. Thompson (1995b: 95) thus thinks that 'the processes involved in colour constancy and in spatial segmentation are likely to be intertwined, not just biologically, but also computationally'.[82] This undermines the reflectance discrimination view of colour constancy, which simply takes spatial representations as given.

Reply: This objection concerns the purported insulation of colour constancy processes from those related to object discrimination and spatial segmentation. The objection is that the reflectance discrimination account of colour constancy takes an overly simplistic view of the processes involved in generating representations of surface reflectance, disregarding the probable 'sharing' of information between the colour vision system and object vision. In response, I would argue that some degree of theoretical simplification or idealisation of this sort is both useful and necessary in advancing our understanding of the distinctive nature of our various visual perceptual abilities. It is entirely normal that the achievement of a particular type of perceptual constancy should be treated theoretically as an

[82] See also Thompson (1995b: 198).

independent and 'informationally encapsulated' visual process.[83] In reality, however, we should of course expect significant anatomical, functional, and computational overlap and exchange between the visual pathways responsible for the achievement of perceptual constancies for the various visual perceptual features. In the service of generating a rapid and unified perceptual representation of the objects and features in the scene, the visual system runs numerous feature-based processes in parallel. Information from a given processing stream can influence and interact with numerous other visual processes in a lateral and subsequently top-down fashion. Our ultimate picture of the visual system no doubt will have to account for and explain these complex interdependencies. There also seems little doubt that accounts of visual phenomenology will reflect the upshot of these varied interactions between visual processes. At early stages of understanding, however, taking such a holistic approach would result in our theory of colour vision becoming a theory of *everything*. By focusing on the constitutive nature of colour vision *qua* psychological kind, the aim is rather to acquire an understanding of the distinctive contribution that colour vision makes to our perception of the world.[84]

1.4.5. Explaining Approximate Colour Constancy

Objection: It is well known that the degree of colour constancy achieved by biological organisms is never perfect. All naturally encountered examples of colour vision involve the ability to achieve merely *approximate* colour constancy. On the reflectance discrimination view, the approximate nature of colour constancy is interpreted as a *flaw* or as *sub-optimal performance* relative to the constitutive ideal of perfect reflectance discrimination. Indeed we may presume that this constitutive ideal could never be instantiated by any biological

[83] See Fodor (1983) for discussion of informational encapsulation and its relationship to psychological modularity.

[84] Discussing computational approaches to colour constancy, Wandell (1989: 163) relatedly notes that 'rather than address the biological system immediately, with all of its complexity and competing demands, we first try to understand how a simple, ideal system might perform'.

organism. We might wonder, however, why we should view all biological colour vision systems as sub-optimal relative to this ideal constitutive standard. Wouldn't it be more plausible rather to view approximate colour constancy as optimal *relative to the numerous and varied tasks that colour vision evolved to perform?* Thompson and colleagues (1998: 366-367) explains this thought as follows,

> If we wish to design a visual system that exhibits complete constancy, and the system exhibits only approximate constancy, then we are justified in saying that the system does not perform optimally, that it fails to achieve the task for which it was designed. But because natural colour vision presumably resulted from evolutionary tinkering involving "trade-offs" rather than the optimal design, why should the approximate constancy of natural colour vision be explained by appealing to such a strong, engineering notion of optimality?

Rather than viewing approximate constancy from the 'top-down' as a flawed or sub-optimal realisation of a constitutive ideal, then, we should view this phenomenon from the 'bottom-up' as the result of numerous evolutionary trade-offs between perceptual tasks and the competing demands placed on an organism's physiology. This point significantly undermines the highly idealised constitutive approach to understanding colour constancy and hence colour vision.

Reply: The objection presents a false dichotomy between 'top-down' idealising or constitutive accounts and 'bottom-up' biological or evolutionary accounts of colour constancy. I see these two approaches as entirely complementary and believe that they both play an essential role in explaining the approximate nature of colour constancy in biological visual systems. As should be clear, in biology as elsewhere in science it is entirely normal to explain the actual abilities or capacities of organisms as sub-optimal relative to some idealised, typically evolutionary, function. This approach is entirely compatible, however,

with appealing to trade-offs between competing tasks or factors in the organism's actual biology or physiology in explaining why it performs this function sub-optimally. Often the explanation will be that the organism performs sub-optimally relative to some ideal standard because it employs elements of its physiology to perform functions for which these elements are sub-optimally designed or structured. The relevant physiological subsystem might also be subject to evolutionary trade-offs between processing efficiency and energy efficiency, for example, or between the competing demands of various distinct information processing tasks that the subsystem performs, with the result that no particular task is performed optimally.[85] All organisms simply make use of the raw biological materials available to them in order to adapt and survive and it is predictable that these biological materials should sometimes have features that frustrate the functions for which they are employed.

To take an illustrative example from the field of vision research, consider the arrangement of cells in the mammalian retina. It is presumably not contentious that the primary adaptive function of the retina is to transduce light transmitted to the eye from the scene around the organism. This is a top-down, engineering-type, idealised constitutive conception of retinal function. It is both modally necessary and sufficient for some mass of tissue to *be* a retina that the tissue transduces light in this way.[86] The mammalian retina is suboptimal in fulfilling this function. A partial explanation of this fact is that the cells of the retina have an inverted arrangement, whereby the light transducing photoreceptors are embedded deeper in the surface of the retina than the ganglion nerve cells which carry information to the brain. Light must therefore travel through the nerve cells prior to reaching the photoreceptors. One significant consequence of this arrangement is the blind spot, the point in the retina in which all retinal nerve cells meet and pass along the optical nerve.

[85] The brain itself is a voracious consumer of energy, requiring more than any other organ and accounting for approximately twenty percent of the human body's total consumption. Accordingly the development of larger, more energy intensive, brains comes at a *significant* cost to the organism.
[86] Artificial retinas developed outside of the organism would, in my view, therefore still class as retinas.

Objects presented in an organism's blind spot will not be detected. Some other species do not display this inverted retinal structure and in this respect their retinal physiology is better suited to the function of transducing light.[87] In the mammalian case, however, the putatively sub-optimal retinal design can be explained by appealing to a trade-off between information processing efficiency and the energy needs of the retinal photoreceptors. The process of generating an electrical response to an absorbed light photon requires a *huge* amount of energy and indeed the energy required by photoreceptive cells is among the largest of any known tissue. The seemingly inverted arrangement of retinal cells therefore could be due to the significant need to keep photoreceptors 'bathed' or well supplied with arterial blood.[88] In this case, then, it is entirely consistent to hold that the performance of the retina is sub-optimal with respect to some idealised evolutionary, constitutive, function, while simultaneously explaining this imperfect performance by appealing to trade-offs between competing needs and demands on the organism's physiology.

The situation in the case of colour constancy is entirely analogous. On my view colour constancy performance in biological organisms is deemed sub-optimal relative to a top-down, idealised, constitutive standard. Part of the explanation as to why organisms perform sub-optimally in this respect will proceed in a bottom-up fashion by appealing to the competing demands placed on an organism's visual subsystems by the various information processing tasks that they have evolved to perform. A well known example is that humans possess only three cone types, which results in a limited capacity to sample or encode the information received in the proximal stimulus. Given that objects with distinct surface reflectances can reflect light which stimulates the three cone types to exactly the same degree, the input into the vision system will often equivocate on features of the

[87] This seems to be the case in Cephalopods. For discussion, see Novella (2008: 493-494).
[88] See Denton (1999).

environment.[89] As such, our discrimination of surface reflectance properties and hence colour constancy performance will be less than perfect. This performance can be partly explained by a trade-off between the demands of chromatic informational sensitivity and visual *spatial* acuity. Roughly speaking, the more distinct types of chromatic photoreceptor there are present in the retina, the lower its spatial resolution. Similar forms of explanation apply to every type of visual perception. There is simply no incompatibility, therefore, between the top-down understanding of approximate colour constancy as sub-optimal performance relative to a constitutive ideal, and the bottom-up explanation of this performance via trade-offs between various competing demands on an organism's actual physiology.[90]

1.4.6. Reflectance Properties *versus* Colour Categories

Objection: The reflectance discrimination view of colour constancy holds that surface reflectance properties, or perhaps productances, are the only distal properties implicated in understanding the nature of colour vision. Although surface reflectance properties or productances are clearly causally implicated in colour vision processes, however, these properties are not relevant to understanding the colour-directed *behaviour* of organisms in their natural environment or ecological niche. The properties that are relevant to understanding behaviour are the colour *categories* such as *red*, *yellow*, *blue*, and *green*. Colour categories are the most perceptually salient features of our visual perception of colour. Colour categories are also the most relevant types implicated in object detection and identification, and in guiding behaviour such as seeking out food and conspecifics, and social interactions such as mating. In order to explain the role played by colour constancy in guiding

[89] This is known as the issue of *metamerism*.

[90] The importance of interplay between the computational, algorithmic, and implementation levels of explanation has been highlighted since the earliest days of computational theory. For example, see Marr (1982: 330): 'I have laid particular stress on the level of computational theory, not because I regard it as inherently more important than the other two levels – the real power of the approach lies in the integration of all three levels of attack – but because it is a level of explanation that has not previously been recognised and acted upon.'

biologically relevant behaviour, colour constancy therefore must be understood as generating stable, illumination-independent, representations of colour *categories* rather than surface reflectance properties.

Thompson (1995b: 196) presents this argument in some detail,

> The key to colour constancy is to be found in this joining of object identification and categorical colour perception. Rather than providing constant perceptual indicators of surface reflectance, the primary role of colour vision is probably to generate a relatively stable set of perceptual categories that can facilitate object identification and then guide behaviour accordingly. The detection of surface reflectance is certainly useful for object identification, but, as Jameson and Hurvich note, "surface colour recognition... is adequately accomplished by category matching and does not require precise matching-to-sample by the three colour variables of hue, brightness, and saturation"... In general, categorical perception involves a many-to-one mapping from the physical to perceptual dimensions... in the particular case of colour vision, this means that many physically distinct surface reflectances will be lumped together into one colour category. It is the approximate invariance of the categorical colour with variations in the illumination that is crucial for object identification and subsequent behaviour, not constant hue-saturation-lightness correlates of surface reflectance.[91]

In a very similar vein, Jameson and Hurvich (1989: 194) claim that 'it seems predictable that computational approaches to the old issues of colour constancy will not for long continue to seek direct and precise perceptual correlates of constant surface reflections, but will increasingly embody the more realistic approach of object identification through approximate invariance of colour category'.[92] In concluding this argument Thompson (1995b: 197) remarks,

[91] The quote from Jameson & Hurvich can be found at (1989: 195).

[92] Also see Troost and De Weert (1991: 596): 'If colour constancy would be described from a more natural perspective, it can be argued that the purpose of colour constancy is to prevent erroneous object-colour perceptions due to variations in illumination. Colour constancy would exist as long as an object is not assigned a different colour under different illumination conditions... colour shifts resulting from different illuminations are not so great to prevent object identification by colour, because these shifts remain within colour-category boundaries.'

> If colour constancy is an adaptation, it is better regarded not as an adaptation for object detection via the recovery of surface reflectance, but rather as an adaptation for integrating a physically heterogeneous collection of distal stimuli into a small set of visually salient equivalence classes..., ones that can be deployed in a variety of perceptual conditions.[93]

Contrary to the reflectance discrimination view, then, the representational capacities implicated in colour constancy must be colour *category* representations rather than representations of surface reflectance properties.

Reply: The first point to note is that the view of colour constancy presented by Thompson is extremely revisionary. In a recent survey article on theories of colour constancy, Hurlbert (1998: 284) says that 'almost all varieties of computational solutions to colour constancy assume that the goal of colour constancy is to recover an estimate of surface reflectance'. The onus is thus firmly on Thompson to present overwhelming reason to abandon the reflectance based conception of colour constancy in favour of the colour category based view that he favours.

The second point is that while Thompson is right to highlight the importance of colour vision in aiding object identification and in guiding other ecologically relevant types of behaviour, these broader perceptual and sensorimotor abilities do not aid in understanding the *distinctive* nature of colour vision *qua* psychological kind. Thompson claims that 'the primary role of colour vision is probably to generate a relatively stable set of perceptual categories that can facilitate object identification'. In response, however, surely it is the case that *every* type of visual ability 'facilitates object identification' in some way or another. The ability visually to perceive the shape, size, depth, motion, and texture of objects in the environment all presumably contribute to the broader aim of identifying or recognising

[93] See also Hardin (1992), Hatfield (1992), and Troost & De Weert (1991: 596).

objects as falling under ecologically relevant types such as *food*, *prey*, *predator*, *shelter*, *potential mate*, *offspring*, and so on. As such, Thompson's characterisation under-describes the contribution made by colour vision to an organism's overall cognitive economy and biological success. What is *most* distinctive of colour vision and what distinguishes it from other types of visual ability, I believe, is that it enables organisms to achieve colour constancy, which I understand as the ability to discriminate objects in respect of their surface reflectance or productance properties. The contribution of colour vision to object identification processes can easily be understood from this foundational standpoint, whereas Thompson's proposal that we place object identification centre stage in our account robs it of explanatory power.

The third point is that in the next chapter I shall provide a principled argument that colour constancy *qua* reflectance discrimination, and hence colour vision, should be dissociated philosophically from our ability to perceive colour categories or *categorical perception of colour*. The conclusion of the argument is that the ability for categorical perception is not part of the constitutive nature of colour vision. It is important to note that this conclusion is entirely compatible with the claim that our ordinary visual experience of colour displays the structural or organisational features distinctive of categorical perception. This conclusion is even compatible with the claim that in some actual creatures the ability to achieve colour constancy, and hence visually to perceive colour, is computationally and physiologically entwined with the ability for categorical perception of colour. It does not follow from this, however, that categorical perception is essential to the ability to achieve colour constancy and hence Thompson's objection fails to undermine the present constitutive account of colour vision.

1.4.7. Vision for Action

Objection: The proposed account takes colour constancy to be necessarily equivalent to the ability to *discriminate* between objects in respect of reflectance (or productance). It was argued, moreover, that these reflectance discriminations must be made *on the basis of vision* and not some other sense modality such as touch or audition. We could well imagine an organism, however, that is capable of discriminating differences in reflectance on the *basis* of visual processing but that only ever employs the resulting representations for the purposes of *guiding action*. This organism therefore would meet the proposed standard for reflectance discrimination and hence colour constancy, but would never be visually conscious of colour as a distinctive perceptible quality.[94] If you like, the organism would possess a type of 'vision for action' for reflectance properties while lacking any type of 'vision for perception' for these properties.[95] The proposed discriminatory criterion is thus insufficient for colour vision *qua* vision for perception.

Reply: I believe that the notion of discrimination operating in the proposed constitutive account provides us with the explanatory resources to counter this objection. In the initial clarifications of the knowledge account of discrimination provided above, I noted that discriminating subjects must possess some sort of first personal or whole organism access to, or monitoring or registration of, the relevant representations of environmental differences, such as would allow the acquisition of knowledge of distinctness. The knowledge

[94] Pigeons are known to use visual discrimination in the short-wavelength and ultraviolet range to aid aerial navigation (Thompson et al. 1998: 389) but it is a moot point whether these representations are used in a dedicated action-guiding system or pathway or whether they should be classed as visual perceptual representations. Thanks to David Hilbert and Martin Davies for raising this objection to discrimination-based accounts of colour vision.

[95] This distinction originates with Goodale & Milner (1992) and Milner & Goodale (1995). Very roughly, vision for perception (the so-called 'what' pathway) is taken to be subserved by processing in the ventral stream whereas vision for action (the so-called 'where' or 'how' pathway) is subserved by processing in the dorsal stream. Goodale and Milner (1992: 20) contend that vision for perception, which leads to our familiar qualitative conscious visual experience of the world, is a relatively recent development or acquisition when compared with vision that serves to guide or control action.

account of discrimination therefore is a *personal* or *whole organism* rather than *sub*personal account: it is subjects or whole organisms that discriminate, and not the subsystems, organs, or any other putatively information-sensitive parts of that subject. As noted above, in paradigmatic cases visual discriminations are often subserved by the organism's phenomenal consciousness of the relevant environmental features. In other cases the operative system of monitoring or registration may be more basic than phenomenal consciousness. In any event it follows from this account that any representational processes occurring in a visual subsystem or pathway that are not in principle accessible to the subject, or capable of integration in some type of personal or whole organism level monitoring or registration system, could not subserve reflectance discrimination in the sense at issue. In particular I would argue that reliable representations of differences in reflectance properties tokened in a dedicated 'vision for action' pathway would not suffice for reflectance discrimination on the present view and hence would not suffice for colour vision. This is not to deny, however, that 'vision for action' systems might achieve something *very much like* perceptual constancy with respect to features such as shape, size, colour, and motion. Indeed for the purposes of guiding motor responses it seems essential that an action for vision system should be capable of forming representations of distal features that are relatively stable across changes in perspective and other contextual parameters within the scene. The crucial point is rather that the notion of perceptual constancy that is most relevant to understanding colour vision *qua* vision for perception requires *more* than just the stable representation of distal features within an action guiding system.

1.5. Conclusions

In this chapter I have argued that it is constitutively necessary and sufficient for the possession of colour vision that an organism is able to achieve colour constancy. I first evaluated the dominant conception of colour vision as the ability to discriminate differences

in wavelength. This discussion brought to light several contrasting notions of discrimination and I argued that the best of these options is provided by the knowledge account of discrimination. The wavelength discrimination view of colour vision ultimately failed due to the intuitively evident centrality and importance of colour constancy to colour vision. This argument provided *prima facie* motivation for the alternative view that colour constancy itself is constitutive of colour vision. I then provided several independent arguments in support of this constitutive thesis, drawing on computational, comparative, and evolutionary accounts of colour vision. I then proposed and defended a novel characterisation of colour constancy as the ability to discriminate surface reflectance properties across changes in ambient illumination. In contrast to the standard appearance invariance view, this characterisation avoids making any explicit appeals to features of our conscious experience of colour. I believe that this account is better fitted to the background scientific theories of vision and that it also provides a more secure epistemic basis for doing colour ontology. After briefly outlining the consequences of this constitutive account for colour ontology, I considered a range of objections and offered replies, which show the constitutive account to be remarkably resilient.

2. Colour Constancy and Illumination

In the previous chapter I argued that it is constitutively both necessary and sufficient for the possession of colour vision that an organism possesses colour constancy. The standard *appearance invariance* or *invariantist* view characterises colour constancy as the phenomenon whereby objects phenomenally appear approximately the same in respect of colour across changes in illumination. As part of my constitutive account of colour vision I proposed an alternative *reflectance discrimination* view of colour constancy as the ability to discriminate differences in reflectance (or productance) across changes in the illumination conditions. In this chapter I shall discuss at some length an important objection to the appearance invariance view posed by the *argument from illumination*. The basic thrust of the argument is that in standard cases of colour constancy the phenomenology of apparent colour is partly *illumination-dependent*. This point suggests that invariantism is at best incomplete and at worst outright false as an account of colour constancy. In theory one might think that this objection to invariantism should be grist to my dialectical mill. The situation is more complicated, however, as the objection also seems to carry over to the reflectance discrimination view. The problem is that like invariantism, the reflectance discrimination view seemingly assigns no significant role to information about the illuminant in achieving colour constancy. It is therefore important that we get clear on the nature of this objection and assess the threat posed to the reflectance discrimination view. I shall argue that the reflectance discrimination view is in fact much better placed than invariantism to deal with this objection, which counts as another point in its favour.

2.1. The Argument from Illumination

The appearance invariantist holds that colour constancy is the phenomenon whereby the colours of objects phenomenally appear the same across changes in illumination. As will no doubt be clear, however, in ordinary cases of colour constancy objects certainly do not phenomenally appear *exactly* the same across changes in illumination. On the contrary, while objects viewed under different conditions of illumination do indeed appear relatively similar in respect of colour, there is clearly *some* sense in which objects appear different across such changes. This point is illustrated in figure two. On the one hand, the surface of the pillow and the pavement on which it sits clearly appear relatively uniform in respect of colour: the pillow appears to be more or less the same shade of red all over and the pavement has a relatively uniform gravel grey colour. On the other hand, however, it seems equally clear that the appearance of the pillow on the directly illuminated left-hand side differs from that of the right-hand side, which is in relative shade. By the same token, there seems to be a clear difference in the appearance of the pavement in the region cast in shadow by the pillow compared to that of the surrounding area. The assumption here is that these contrasting aspects of our visual phenomenology are *concurrent* or *synchronic*. Let's call the first type of visual response the 'invariance response' and the second type the 'variance response', and let's refer to this contrasting pattern of visual responses as the *duality of colour constancy*:

> (DUALITY) Colour constancy in normal subjects characteristically involves two contrasting types of visual response to uniform surfaces viewed under differing illumination conditions: one of *invariance* in respect of the appearance of the colour of the surface and one of *variance* in respect of some aspect of appearance.[96]

[96] As Hilbert (2005: 145) cautions, we should not 'overemphasise the degree and kind of change that we actually observe. The character of the illumination varies frequently and substantially as we move about the world and within a single setting. Nevertheless there is a definite perception that objects do not appear to be frequently and substantially changing in colour.'

This pattern of invariance and variance of appearance across changes in illumination is an entirely regular feature of colour constancy in human subjects and indeed Jonathan Cohen (2008: 62) goes so far as to say that these two contrasting types of response are 'the hallmarks of colour constancy'. It is therefore reasonable to expect that all candidate accounts of colour constancy should have some explanation of Duality.



Figure 2: Red Pillow Casting a Shadow (Hilbert, forthcoming)

By itself Duality does not pose a significant threat to the appearance invariance view. The most obvious and straightforward interpretation of the variance response is that objects viewed across different illumination conditions differ precisely in respect of the appearance of the *illumination* in the scene, where this putative aspect of visual appearance is understood to be independent of the appearance of the colours of these objects. Let's call this the *apparent illumination interpretation*. In support of this interpretation Palmer (1999a: 126) says that 'when an object on a table casts a shadow, ... the surface of the table does not appear darker [or different in colour] there, as though it were stained or painted, it simply looks as

though there is less light falling on it because of the object's shadow...'.[97] Brad Thompson (2006: 25) expresses the same thought that 'in describing the experience of highlight and shadow, it would not be correct to say that an object is a different colour in the shadowed or highlighted areas. The object looks different in those areas, but does not look to be a different colour in those areas'. The apparent illumination interpretation is entirely consistent with the claim that objects viewed across changes in illumination appear approximately the same in respect of colour and hence poses no immediate threat to invariantism.

As noted in my response to the objection in section 1.4.3 of the previous chapter, the apparent illumination interpretation is also perfectly consistent with standard computational accounts of colour constancy and with the reflectance discrimination view. One might initially think that this is not the case. This is because standard presentations of computational theories of colour constancy typically claim that colour vision systems 'discard' or 'discount' information about the illuminant in forming a stable representation of the reflectance properties in the scene. Likewise on the reflectance discrimination view, the supposed 'outputs' of colour constancy processes are stable representations of reflectance properties, not of reflectance *and* illumination properties. The reflectance discrimination view therefore seems to assign no significant role to information about the illuminant in achieving colour constancy. To coin some expressions, let's refer to such views of colour constancy as *reflectance focused*, as opposed to *reflectance and illumination focused*. It is clearly mistaken, however, to infer that reflectance focused views of colour constancy imply that information about illumination conditions plays no role *whatsoever* in the visual system at large. As discussed in the previous chapter, an obvious explanation of the putatively illumination-dependent aspects of our visual phenomenology would be that they reflect the

[97] Admittedly, Palmer is idealising somewhat here. Later on, he admits that in some contexts 'we can see the region of a wall on which a shadow is cast as being darker than surrounding regions at the same time as it appears to have the same surface colour (reflectance) as the rest of the wall', (1999a: 313).

upshot of the *luminance* vision system, which is known to be implicated in the perception of shading and shadows cast by objects.

For reasons that will be discussed in detail below, many people have rejected the apparent illumination interpretation and have argued instead that the variance response tracks some illumination-dependent aspect of the appearance of the *colour* of the object, or as I shall say, a difference in the *colour appearance* or *apparent colour* of the object.[98] Let's call this the *colour appearance claim*. This claim does cause serious problems for invariantism, as noted by Cohen (2008: 65),

> [I]f the [colour appearance claim] is correct, then it is at best seriously misleading to say, with the invariantist, that colour constancy involves an invariance of apparent colour across variations in illumination. This claim respects only one of the two reactions normally sighted subjects have to these cases. While it respects one of their reactions—their reaction that the two members of the pair of stimuli are alike in apparent colour, it ignores their reaction to the effect that the members of the pair differ in apparent colour—that there is not an invariance of apparent colour across changes in the illumination... The problem for the invariantist about colour constancy is that her characterisation of the phenomenon is consistent with the invariance reaction, but not the variance reaction.

Cohen argues that if we accept the colour appearance claim, then the invariantist view of colour constancy accommodates only half of the characteristic data regarding colour constancy in normal subjects. This warrants the conclusion that it is at best incomplete and at worst simply incorrect to characterise colour constancy as involving the invariance of colour appearances across changes in illumination. Let's call this the *argument from illumination*. We can state the argument more precisely as follows:

[98] In talking about 'colour appearances' and 'apparent colours' I do not intend to commit to a range of 'appearance properties' as distinct from colour properties themselves. As we shall see below, authors such as Noë (2004) and Cohen (2008) do in fact introduce two distinct types of colour-related property in response to this issue, but this type of view is by no means mandatory.

1. Colour constancy is the phenomenon whereby the colours of objects phenomenally appear the same across changes in illumination. **(Appearance Invariance View)**

2. Normal subjects viewing a uniform surface under differing illumination conditions typically report that the surface appears both the same in respect of colour and yet that it appears different in some respect *r*. **(Duality)**

3. The respect *r* in which the surface appears different under differing illumination conditions is that of *colour*. **(Colour Appearance Claim)**

4. Objects phenomenally appear both the same *and* different in respect of colour across changes in illumination.[99] **(2, 3)**

5. Therefore, the appearance invariance view of colour constancy is either incomplete or incorrect. **(4, 1)**

This argument will be the focus of discussion in this chapter.

Why have I chosen to focus critical attention on the argument from illumination given that I have already rejected the appearance invariance view of colour constancy? Isn't this argument simply grist to my mill in providing further evidence as to the inadequacy of invariantism? Unfortunately the dialectical situation is not so straightforward. It seems that the argument from illumination undermines the reflectance discrimination view just as much as appearance invariantism. If the argument is sound, it follows that there is a highly regular and characteristic aspect of colour appearance that is *illumination-dependent*. The reflectance discrimination view *prima facie* seems to have difficulty in accounting for this putative illumination-dependent aspect of the phenomenology of apparent colour. If colour constancy

[99] At first blush, this claim seems to commit us to attributing representations of inconsistent states of affairs to subjects' visual perceptual states. As Hilbert (2005: 145-146) points out, this would render colour constancy analogous to perceiving an impossible object such as an Escher triangle, which seems an inaccurate description of the phenomenon. As we shall see below, advocates of this argument therefore need to provide a more complex or nuanced account of the contents attributed in colour constancy in order to avoid this outcome.

were constitutively a matter of discriminating reflectance properties across changes in illumination, then why would normal subjects routinely report some illumination-sensitive aspect of the phenomenology of apparent colour in normal cases of colour constancy? More generally, how can *any* such reflectance focused account of colour constancy accommodate this fact? In focusing primarily on our representational capacities with respect to reflectance properties, the reflectance discrimination view seemingly lacks the resources to explain this illumination-dependent feature.

Before launching into discussion of these issues, I want to highlight some analogous phenomena in other areas of vision research that will be important in what follows. Colour constancy is one among many types of perceptual constancy standardly characterised in appearance invariance terms. The standard characterisation of *size constancy*, for example, is that an object ordinarily will appear the same in respect of size across changes in the distance between the object and the subject. It is therefore interesting to note that all of the major types of visual perceptual constancy display patterns of contrasting visual responses analogous to those recorded in Duality, including size, shape, depth, orientation, position, and motion constancy.[100] It is a familiar fact that in size constancy, for example, while there is a clear sense in which two objects of the same physical dimensions viewed at different distances do indeed appear the same in respect of size, there is also a clear sense in which these objects appear different.[101] By analogy with the apparent illumination interpretation of the variance response in colour constancy, perhaps the most obvious and straightforward explanation of this difference would be that these objects differ in respect of the appearance of the *relational*

[100] The general tendency of perceptual representations towards proximal or contextually variant features, such as changes in illumination in colour constancy or changes in distance from the viewer in size constancy, is sometimes referred to as *under-constancy*. Although infrequently discussed, the contrary phenomenon is *over-constancy*, in which perceptual representations 'overshoot' or 'over compensate' for variations in the proximal stimulus. In size constancy this would have the result that objects appear larger than they are.

[101] To pick a more complex example, in standard cases of *motion* constancy while there is a clear sense in which an object travelling at constant velocity from left to right appears to move at the same rate as one moves one's head to the right, there is also a clear sense in which the object may appear to increase in velocity.

distance of the objects from the subject, that is, in how far away these objects appear to be. In parallel with the colour appearance claim, however, many people would claim that objects viewed at different distances in fact differ in respect of *apparent size*. As Cohen (2008: 69, fn. 16) observes, if this claim were sound we could argue against the appearance invariance view of size constancy as follows,

> Assume..., as seems not implausible, that the empirical facts about apparent size are relevantly analogous to those about apparent colour: suppose that we can get subjects to treat two discs of the same diameter, viewed at different depths, as unlike in apparent size. In this case, I think we would be justified in concluding that there's a good sense in which the two discs are unlike in apparent size, hence that the size invariantist's claim that the two discs are alike in apparent size (despite a difference in depth) is inadequate as a characterization of the phenomenon.

As in the argument from illumination against invariantism about colour constancy, this argument implies that the appearance invariance view of size constancy is at best incomplete and at worst straight out false. Generalising on this strategy, we can assume that *any* appearance invariance characterisation of perceptual constancy will be under similar threat.

Now suppose that we sought to adapt and extend the reflectance discrimination view to provide a constitutive account of other visual perceptual constancies such as size constancy. As a rough sketch, the view would be that it is constitutively both necessary and sufficient for the possession of visual size constancy that a subject is able visually to discriminate objects in respect of their physical dimensions across changes in viewing distance. By analogy with the reflectance discrimination view, the supposed 'outputs' of size constancy processes would be stable representations of the object's physical dimensions, *not* of the object's dimensions *and* distance from the viewer. In other words, the size discrimination view of size constancy would be primarily *size focused* as opposed to *size and*

distance focused. If the objection to the reflectance discrimination view of colour constancy posed by the putative illumination-dependent aspect of apparent colour is successful, we might then predict that the size discrimination view of size constancy would face similar problems in accommodating the putative distance-dependent aspect of apparent size. On the other hand, if this objection would be considered unreasonable or implausible in the case of size constancy, then on the grounds of parity we might reject the analogous objection against the reflectance discrimination view of colour constancy. This point will provide some much needed perspective as we progress.

2.2. The Colour Appearance Claim

The key premise in the argument from illumination is premise three, the *colour appearance claim*. This premise states that the respect *r* in which surfaces appear different under differing illumination conditions is that of *colour*. Together with Duality, this premise entails that objects phenomenally appear both the same *and* different in respect of colour across changes in illumination, which undermines the invariantist view and also raises problems for the reflectance discrimination view. In this section I shall review some arguments that have been given in support of the colour appearance claim. We shall see that this claim is far from clear, which casts the argument from illumination in a dubious light.

The colour appearance claim is sometimes motivated by appeals to the ordinary visual phenomenology of perceiving objects under differing illumination conditions. Such arguments are by no means new, as we can see from the following remarks by Russell (1967: 2),

> Although I believe that the table is "really" of the same colour all over, the parts that reflect the light look much brighter than the other parts, and some parts look white

> because of the reflected light. I know that, if I move, the parts that reflect the light will be different, so that the apparent distribution of colours on the table will change.[102]

Russell argues on the basis of naive reflection on visual phenomenology that the table looks 'brighter' in regions in which it reflects most light, and indeed looks 'white' in some such regions. More recently Noë (2004: 127) has made essentially the same claim,

> You also see that the wall is brighter here, where it falls in direct sunlight, than it is there, where it falls in shadow. Differences in brightness, however, mean differences in colour. You see the uniformity of colour despite the evident non-uniformity of different parts of the wall's surface.

Following Russell, Noë appeals to visual phenomenology in arguing that the wall looks 'brighter' in regions that are directly illuminated, from which he infers that there is a 'difference in (apparent) colour' in these regions. Hilbert (2005: 144-145) offers some similar remarks,

> It is also common for the surfaces of objects to be visibly shaded with some areas appearing brighter and others dimmer.... [T]here are also, equally familiar, effects on colour appearance of the character of the illuminant. Although the white (black) areas of the printed page… look white (black) under both the indoor and outdoor illuminations they don't look the same. Outside the white seems whiter and the black blacker than when viewed under the dimmer indoor light.

As with Russell and Noë, Hilbert argues that the differences in appearance that we encounter across changes in illumination conditions are differences in the apparent 'brightness' or 'dimness' of a colour.

[102] As quoted by Hilbert (2005: 144).

What should we make of these arguments for the colour appearance claim? All three of the authors quoted above argue that changes in illumination conditions can result in a difference in the apparent 'brightness' of a surface, which is assumed to be sufficient to establish a difference in the apparent *colour* of a surface. In a critique of Noë (2004), Kalderon (forthcoming) rightly notes that this assumption is problematic because 'brightness' is ambiguous between 'chromatic brightness', which is agreed to be a dimension of colour appearance, and 'luminosity', which is the phenomenological correlate of *luminance*, which as we have seen is a measure of the intensity of the light reflected from a surface, and not a dimension of colour appearance. It is not straightforwardly valid, therefore, to infer a difference in apparent colour from a difference in the apparent 'brightness' of a surface. Kalderon (forthcoming: 7) expands on this point as follows,

> The reasoning, here, can seem to turn on an equivocation. It is true that the wall is brighter where it is directly illuminated where "brightness", here, means the intensity of the light reflected by the white wall in direct sunlight. However, Noë needs to understand "brightness" as a dimension of colour similarity and difference, if he is to argue from a difference in brightness to a difference in colour. Unfortunately these distinct senses of brightness can come apart. Consider viewing a page of black print on white paper, first indoors under artificial illumination, then outside in natural daylight. The intensity of the light reflected by the white area of the page indoors is approximately the same as the intensity of the light reflected by the black print in sunlight... Despite being equally bright, in the sense that the reflected light is equally intense, the apparent colours differ in chromatic brightness—the surrounding white when viewed indoors seems brighter than the black print when viewed in daylight.

As Kalderon makes clear, chromatic brightness and luminosity are distinct psychosensorial or phenomenological variables that frequently come apart in ordinary visual contexts. Given that luminosity is not strictly a dimension of apparent colour, then, there is at least one interpretation of the term 'brightness' on which it would be invalid to infer a difference in apparent colour from a difference in brightness. Indeed if 'brightness' were interpreted as

'luminosity', then presumably the most natural explanation of the variance response would be the apparent illumination interpretation introduced above.

Advocates of the colour appearance claim therefore need some further argument as to why the putative differences in apparent 'brightness' to which they appeal should be interpreted as differences in *chromatic* brightness and not differences in luminosity. It is unclear to me whether naive reflection on visual phenomenology could suffice to establish this point. First, the distinction between chromatic brightness and luminosity is both subtle and technical and accordingly it takes a certain amount of specialist theoretical knowledge to distinguish between these variables. Second, as intimated in the previous chapter, despite being theoretically distinct visual pathways it is predictable that in practice the colour vision system and the luminance vision system will be intertwined computationally and physiologically. It is likewise reasonable to assume that the phenomenology of chromatic brightness and luminosity will be similarly complex. As such I strongly doubt that normal subjects would possess the introspective acuity to distinguish between these two phenomenological variables, at least not without significant further cues or guidance.

Noë and Cohen have provided some further arguments in support of the relevant chromatic interpretation of the data on 'brightness', which might serve to avoid the dangers noted in the previous paragraph. Noë (2004: 128) presents his argument as follows,

> Consider that although we can perceive a wall that is illuminated unevenly as uniform in colour, it is also the case that when a wall is in this way illuminated unevenly, it is also visibly different with respect to colour across its surface. For example to match the colour of different parts of the wall, you would need different colour chips.

Noë's argument is that while we could match the appearance of a white wall that is normally illuminated with the appearance of a white colour chip, we could match the appearance of a shadowed portion of that same white wall with the appearance of a grey colour chip. The key assumption here is that if such changes in illumination produce differences in appearance that can be *matched* by different colour samples, then the differences in appearance themselves must be differences in the appearance of the *colour* of the wall.

Cohen (2008) has provided a very similar argument for the same conclusion. Whereas Noë's argument appeals to intuitive considerations regarding phenomenological matching, however, Cohen appeals to some famous psychophysical experiments by Arend and Reeves (1986) designed to test the degree of colour constancy displayed by subjects under different task instructions. The experiments used the classic asymmetric matching task paradigm described briefly in the previous chapter. To recap, in these tasks subjects are presented with two stimuli, one referred to as the 'standard field', with a uniform reflectance R presented under an illuminant I_1, and another 'test field', with the same reflectance R but presented under a different illuminant I_2. The subject is expected to adjust the 'chromaticity' values of the test field until she achieves a match between the test field and the standard field. The chromaticity value of a stimulus signifies its coordinates in the CIE chromaticity diagram presented in figure three below, a prevalent colorimetric system of describing physical stimuli in which the two axes represent the proportions of carefully selected lights or 'primaries' required to match the stimulus in respect of colour.[103] Colour constancy is characterised operationally such that the smaller the adjustment in the chromaticity of the test field needed to achieve a match with the standard field, the higher the implied degree of

[103] Due to the *trichromacy* of human colour vision, any stimulus can be matched in colour by a mixture of *three* suitably chosen primaries *x*, *y*, and *z*. For obvious reasons of convenience, it is preferable to use a two-dimensional representation of these data as in the CIE chromaticity diagram. Given that the proportions of these primaries must add up to 1, the value of the third primary is always recoverable from a two-dimensional (*x*, *y*) value. For example, for a bluish stimulus mapped to (0.2, 0.2), *z*=1-0.2-0.2=0.6. See Palmer (1999a: 692-693) and MacAdam (1985) for more details.

colour constancy. By the same logic, larger adjustments in chromaticity correspond to lower implied degrees of colour constancy. In effect, the experiments are probing how much effect the difference in illumination between the standard and test fields has on the subject's colour perceptions. Intuitively speaking, the less the subject needs to compensate for this difference in illumination by adjusting the chromaticity of the test field, the lower the effect of this difference in illumination on her perceptions of the colour of the surface and thus the better her colour constancy.

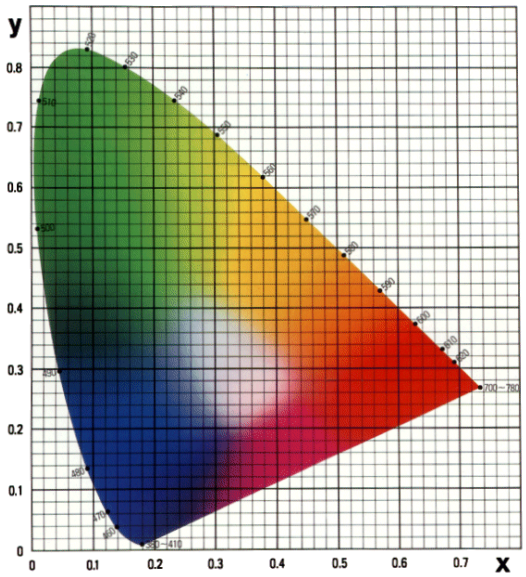


Figure 3: CIE Chromaticity Diagram

The crucial twist in Arend and Reeves' experiments is that prior to making any matches, the subject is given one of two instructions. The first type of instruction is to 'match the hue and saturation of the test field to that of the standard field'. Call this the 'hue/saturation matching' instruction. The second type of instruction is to 'make the test field look as though it is cut from the same piece of paper'. Call this the 'surface matching' instruction. The interesting result is that subjects' performance on the asymmetric matching task differs reliably and significantly depending on the instructions given. Subjects given the hue/saturation matching instruction made fairly large changes in the chromaticity settings of

the test field in order to achieve a match between the two differently illuminated fields. By the assumed operational standard for colour constancy, these subjects displayed relatively low degrees of colour constancy on this task. In contrast, subjects given the surface matching instruction made fairly small changes in the chromaticity settings of the test field in order to achieve a match between the two differently illuminated fields. By the assumed operational standard for colour constancy, these subjects displayed relatively high degrees of colour constancy on this task.

Like any psychophysical experiments, the Arend and Reeves results are of course open to interpretation. Taken at face value, however, the results seem to provide a quantitative basis for the intuitions behind Duality, the claim that colour constancy in normal subjects typically involves two contrasting responses to uniform surfaces viewed under differing illumination conditions: one of sameness or invariance and one of difference or variance. One plausible interpretation of the results is that under the hue/saturation matching instructions subjects attend to some feature of their visual phenomenology that is relatively *variant* across changes in illumination, whereas under the surface matching instructions subjects attend to and seek to match some feature that is relatively *in*variant across changes in illumination. This pattern of responses coheres extremely well with the intuitive observation that the phenomenology of colour constancy has both illumination invariant and illumination variant aspects.[104]

Of more immediate relevance to the colour appearance claim is the fact that subjects under the first set of instructions were asked to match the *hue and saturation* of the test field to that of the standard field. This provides *prima facie* reason to think that the relatively

[104] The instructional effects reported by Arend and Reeves are by no means original in the wider context of perceptual constancy research. Experiments by Carlson (1960) found that subjects' performance on an asymmetric matching task for *size* constancy was similarly sensitive to instructions. Under so-called 'analytic' matching instructions, Carlson found that subjects' size constancy was relatively poor, with the converse being true under 'objective' instructions. See also Singer (1952).

variant, illumination-dependent, aspect of phenomenology that subjects attend to under these instructions is indeed an aspect of apparent colour. Cohen (2008: 67-68) presents this reasoning as follows,

> When subjects make appearance matches [i.e. hue/saturation matches]... they make the regions cease to be discriminable (along whatever dimension they were previously discriminable) by adjusting the hue and saturation of one of them. Now, it is a standard assumption in visual psychophysics that the hue and saturation of a patch are dimensions of its apparent colour; if so, then adjusting the hue and saturation of the test patch just is adjusting the patch's apparent colour. Therefore, whatever the difference was in virtue of which the patches were initially visually discriminable, that difference can be offset by a difference in apparent colour. And this, in turn, might lead us to suspect that the difference revealed in the variance reaction is a difference in apparent colour...

Cohen claims that the difference in appearance induced by the difference in illumination conditions between the two fields can be 'offset' or cancelled out by adjusting the hue and saturation of the test field, where 'hue' and 'saturation' are taken to refer to psychosensorial or phenomenological dimensions of apparent colour. Cohen then argues that if the difference in appearance between the standard and test fields can be *offset* by adjusting dimensions of apparent colour, it is reasonable to conclude that the standard and test fields *differed* in respect of apparent colour. Cohen is not alone in drawing this conclusion, which is echoed by Arend and Reeves themselves (1986: 1747, 1749) and Karl-Heinz Bäuml (1999: 1532) who agrees that 'this pattern of results demonstrates that the same surface can have quite a different colour appearance under two different illuminants but, at the same time, may be perceived as roughly the same surface rendered under two different illuminants'.

How successful are these further arguments for the colour appearance claim? I shall focus on Cohen's argument as this is the more detailed and sophisticated. One immediate

issue with this argument concerns the design of Arend and Reeves' psychophysical experiments.[105] Those familiar with the asymmetric matching paradigm will be aware that in most experiments the standard field and the test field are computer generated stimuli and thus not actually presented under different illuminants. This is certainly the case with Arend and Reeves (1986). It is therefore not entirely straightforward to draw conclusions from such experiments as to the putative effects of illumination on colour appearance. To his credit, Cohen (2008: 65, fn. 10) acknowledges this complication,

> [T]he typical experimental display consists of a test and a standard patch set against circular surrounds of differing chromaticity, but with the whole scene under uniform illumination (indeed, usually surface illumination is negligible, as such displays are produced on luminous computer monitors). Subjects are then asked to change the chromaticity of the test patch against its surround until it perceptually matches the standard patch against its surround. On its face, this task seems to use chromaticity changes to measure the effect of a difference in surround, rather than the effect of a difference in illumination.

As Cohen notes, the so-called 'difference in illumination' between the two fields is in fact simulated by a difference in the chromaticity of the surrounds of the fields. These experiments are thus only indirectly studying the effects of illumination on colour appearance and are instead using 'colour contrast as a proxy for changes in illumination'.

[105] Arend and Reeves performed their experiments on only three subjects (1986: 1744), two of which were themselves. Their results are thus open to the objection that specialist or expert knowledge was required in order to perform the two matching tasks. Fortunately others have since replicated their findings with larger sample sizes of naive subjects. See Bäuml (1999: 1548), and the references in Cohen (2008: 66, fn. 11). Hilbert (2005: fn. 7, 156) objects that the 'task described in [Arend & Reeves 1986] is somewhat unusual and the experimental literature on instruction effects does not seem to me to support any particular clear conclusion'. In support of this point Hilbert cites Delahunt & Brainard (2004), who found only fairly small, albeit consistent, instructional effects on the amount of variance reported by subjects across changes in illumination in colour constancy tasks. Brad Thompson (2006: fn 21, p19) notes that the experiments in Delahunt & Brainard (2004) are testing *successive* colour constancy, as opposed to *simultaneous* colour constancy, which was studied by Arend & Reeves (1986). Subjects were allowed time to adapt to the illumination changes in the new scene, which might explain the reduced effect of illumination under the hue/saturation matching instructions.

How serious is this issue for Cohen? I suspect that it is not too serious. First, although it would be obviously preferable to study directly the effects of illumination on appearance in assessing colour constancy, many researchers now believe that colour contrast and colour constancy are very closely related phenomena and indeed it seems increasingly likely that they reflect a single underlying perceptual ability.[106] The use of colour contrast as a proxy in studies of colour constancy therefore is quite likely to be benign. Second, Cohen (2008: 65, fn. 10) notes that researchers in any case have now 'begun to measure the effects of illumination differences more directly', and the early signs are that the results of these experiments follow the pattern found by Arend and Reeves (1986).[107] These points lead me to think that this issue is unlikely to be decisive against Cohen's argument for the colour appearance claim.

A more serious problem for Cohen's argument relates to his claim that subjects under the first set of instructions are 'adjusting the hue and saturation of the test patch', from which he infers that they are 'adjusting the patch's apparent colour' in achieving a match with the standard field. This description rather subtly misrepresents the actual situation in these experiments. Strictly speaking, while subjects are asked to *match* the two patches in respect of hue and saturation, the variable that is being *adjusted* by the subject is the chromaticity of the test field, which as we have noted is a physical colorimetric variable yielding the coordinates of the stimulus in the CIE chromaticity diagram.[108] In terms of experimental design, the illumination on the test field is the *independent* variable and the subject's adjustment of the chromaticity of the test field is the *dependent* behavioural variable. It is

[106] The seminal article is Lotto and Purves (2000).
[107] Although not directly concerned with the patterns observed by Arend and Reeves, Brainard et al. (2003: 314-315) acknowledge that computer simulated displays are 'somewhat artificial', and set out to assess colour constancy in 'nearly natural' environments, which involves 'using stimuli consisting of actual illuminated surfaces, configured in three dimensions'.
[108] Arend and Reeves (1986: 1744) say that 'subjects controlled the chromaticity of the test patch with a high-resolution graphics tablet. The computer interpreted the hand-held cursor's position on the tablet as a position within an RGB colour triangle and set the test patch to the corresponding chromaticity. Vertical cursor movement controlled yellow-blue content, and horizontal movement controlled red-green content.'

therefore strictly inaccurate to claim that subjects are literally 'adjusting' the phenomenological dimensions of hue and saturation in achieving a match between the two fields. The relationship between the subject's adjustments in chromaticity and her visual phenomenology is rather more indirect. To be clear, the key psychophysical assumption here is that the subject's adjustments in chromaticity serve as a *metre* or *measure* of her visual phenomenology; more specifically in this case, the phenomenological dimensions of hue and saturation. The relationship between the chromaticity values of the test field and the subject's phenomenology is thus analogous to the relationship between the value displayed by a set of weighing scales and the actual weight of the object that is placed on them. The subject's role in such experiments is likewise akin to that of the weighing scales or measuring device itself: the subject provides a measurement of the variations in her own phenomenology *via* her adjustments in the chromaticity values.

This point highlights a significant lacuna in Cohen's argument for the colour appearance claim. Cohen claims that subjects are 'adjusting the hue and saturation' of the test field and thus 'adjusting its apparent colour' in achieving a match with the standard field. As we have just seen, however, subjects are in fact adjusting the *chromaticity* of the test field under the experimental assumption that these adjustments provide a *measure* of the change in the phenomenological dimensions of hue and saturation. This assumption only holds, however, given the further assumption that subjects are sufficiently accurate in introspecting on their own visual phenomenology that they can reliably track these specific dimensions of colour appearance. There is clearly room for scepticism on this point. It seems entirely possible, for example, that despite being instructed to match the two fields in respect of hue and saturation and despite their best efforts to follow these instructions, subjects under the first set of instructions are in fact attempting to match these stimuli in respect of some *other*

phenomenological dimension.[109] In motivating this possibility it is significant to note that the standard CIE chromaticity diagram says nothing about the dimension of luminosity, the phenomenological correlate of luminance or light intensity.[110] It is therefore possible that two objects mapped to the same point in the chromaticity diagram could differ significantly in respect of luminosity. This point is extremely pertinent in the present context. As noted above, advocates of the colour appearance claim argue that changes in illumination often result in changes in the 'apparent brightness' of a stimulus, where this is interpreted as a change in the dimension of *chromatic* brightness rather than luminosity. The problem is that the chromaticity of a stimulus does not control for variations in luminosity between stimuli and it is therefore difficult to rule out the possibility that subjects are in fact attending to this distinct phenomenological dimension rather than that of hue and saturation. Even if the standard and test field had exactly equal chromaticity values, then, there might still be a difference in luminosity between them that the subject attempts to cancel out. Given that the only variable under their control is that of chromaticity, their only option would be to attempt to cancel out this apparent difference *via* adjustments in chromaticity. A deflationary explanation of this sort would offer a plausible and conservative explanation of the two contrasting sets of results encountered in the Arend and Reeves experiments and thereby would undercut any support that these experiments might provide for the colour appearance claim.

[109] More generally, several authors have recently highlighted the frailty of introspection as a source of knowledge about the exact character of our own conscious experience, even under apparently optimal epistemic conditions. Schwitzgebel (2008: 247), for example, has argued that 'we make gross, enduring mistakes about even the most basic features of our currently ongoing conscious experience (or "phenomenology"), even in favourable circumstances of careful reflection, with distressing regularity'. See also Bayne & Spener (2010). I do not want to place too much emphasis on such global scepticism about introspection, however, given the crucial and ineliminable role that introspection plays in experimental psychophysics. Indeed absent the assumption that introspection is at least *sometimes* a reliable source of knowledge about the character of conscious experience, the entire field of psychophysics would begin to look rather suspect, a consequence that I for one certainly do not want to entertain.

[110] See Gordon and Abramov (2001: 98) and Broackes (2010: 308-309).

Changing tack slightly, it is significant also to note that colorimetric systems such as the CIE chromaticity diagram are designed only to allow reliable predictions of *matches* between stimuli and not their total colour appearance. This point is frequently made in the literature, as Byrne and Hilbert (1997b: xii) point out,

> Most colorimetric systems and methods only allow predictions of perceptual matches rather than perceived colour. Since two stimuli may change substantially in their colour appearance while remaining matched, this constitutes a significant limitation. This caveat applies in particular to the most commonly used colorimetric system – the *CIE standard observer*...[111]

This limitation implies that two stimuli assigned the same coordinates in the CIE chromaticity diagram could differ significantly in respect of their overall colour appearance. Broackes (2010: 309) provides the following detailed discussion of this point,

> A blue card in yellow light may be reflecting light to the eye that matches a certain mixture of our primaries – a mixture that is (we may suppose) in fact white, more or less equivalent to CIE standard illuminant C. If so, it will be mapped at position x = 0.310, y = 0.316. Similarly, a yellow card in blue light may also be reflecting light that matches that same combination of primaries (and, similarly, illuminant C); it too will be mapped at x = 0.310, y = 0.316. And yet, and it may be (if a viewer adapts to the two different illumination conditions) that the first card *looks blue* (as indeed it is); and the second card *looks yellow* (as it too is). They are mapped at the same position in a CIE diagram because the light reaching the eye from the two of them would be matched with the same combination of reflectance lights. But that fact does not fix the things' respective appearance: one card looks blue, and the other looks yellow.

As Broackes points out, the colour appearance of a stimulus of constant chromaticity can 'switch' depending on real or implied features of viewing context and the resultant interpretation of the relative contribution of the illuminant to the proximal stimulus. This

[111] See also Kuehni (2005: 82, 85).

scenario presents an intriguing possible explanation of the two contrasting responses recorded in the asymmetric matching tasks. Suppose that a subject is presented with the test field set at chromaticity $x = 0.310$, $y = 0.316$ and asked to match this to the standard field which appears bluish to her. Depending on her visual system's interpretation of the visual context and the contribution of the illuminant to the proximal stimulus, the test field could appear either bluish or yellowish to her. The subject is thus able to 'switch' or 'aspect shift' the colour appearance of the test field depending on her visual system's interpretation of the illumination in the scene. In the former case, clearly the subject would make relatively small adjustments in chromaticity in order to match the standard field, as the two fields appear very similar in respect of colour. In the latter case, however, the subject would have to make relatively large adjustments to shift the yellowish appearance of the test field to match the bluish appearance of the standard field. The proposal, then, is that rather than indicating two concurrent and contrasting aspects of the phenomenology of apparent colour, the invariance and variance responses in fact track two 'switching' interpretations of the test field stimulus.[112] The case outlined by Broackes provides independent reason to think that such switches in appearance are possible for stimuli that are equated for chromaticity. This alternative explanation again offers a plausible and conservative explanation of the two contrasting sets of results recorded by Arend and Reeves and undercuts any support that these experiments might provide for the colour appearance claim.

To recap, the preceding proposal is that the variance and invariance responses track two different visual interpretations as to the surface properties of the test field. A

[112] By analogy, consider an asymmetric matching task for *size* or *length constancy*. Here two lines – the standard and test lines – of the same physical dimensions are presented at varying distances from the subject and the subject is asked to match the two lines in respect of how long they appear to her by adjusting the length of the test line. As in the tests for colour constancy, suppose that the standard and test lines are in fact represented on two-dimensional fields and that the distance or depth of the line from the viewer is simulated or implied in some way. It is entirely predictable that two-dimensional representations of lines at differing distances from the subject could be open to numerous possible visual interpretations as to their actual physical length. The Necker cube provides a well known example of such ambiguity. The Müller-Lyer illusion also indicates how different cues as to the relative depth of a line can influence its apparent length.

consequence of this proposal is that at least one of these visual responses must be *illusory*. It might initially seem somewhat unattractive to assume that changes in the illumination of a surface could induce an illusory visual representation of its surface properties. There are in fact good reasons, however, to find this interpretation plausible. Perhaps most notably, subjects in the asymmetric matching experiments are operating in highly restricted and artificial visual contexts in which they are denied any of the standard clues as to the illumination conditions in the scene. It is well known that such visual contexts greatly increase the ambiguity of the stimulus and hence the likelihood of misinterpreting the relative contribution of the illumination to the proximal stimulus, which could result in the mistaken perceptual attribution of a significant change in surface colour under a change in illumination conditions. It is well established that by removing cues as to the presence of a shadow falling across a surface, for instance, it is possible to induce the illusion that the shadowed region of the surface has different reflectance properties. This point is illustrated in Hering's famous ringed shadow experiment as replicated in figure four.[113] In the right hand image the blurred edges of the shadow cast by the teacup are replaced with a thick black line, which cues the visual system to interpret the shaded region as possessing different surface reflectance properties as opposed to being under different conditions of illumination. It seems entirely plausible that something similar could be occurring in the asymmetric matching experiments for colour constancy. In the absence of cues as to the differing illumination conditions between the standard and test fields, the resulting difference in the proximal stimulus could well be interpreted by the visual system (at least under one 'aspect switch') as a difference in reflectance properties.[114]

[113] For further illustrations and discussions of this effect, see Palmer (1999a: 126) and Cohen (2008: 74).
[114] Hilbert (2005) objects to this visual illusion response that if changes in illumination across a uniform surface resulted in misperceptions of its surface reflectance, it should appear as though the surface material of an object changes in shaded regions, whereas this does not seem to be the case in normal visual contexts. What I say above is entirely compatible with Hilbert's point. The proposal is that subjects *in the asymmetric matching*

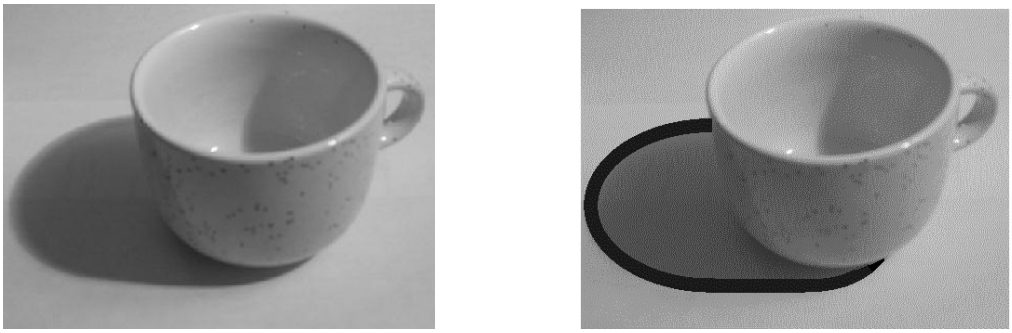


Figure 4: Hering's Ringed Shadow Experiment

The following analogy will help to motivate this idea. It is a well known fact about visual size constancy that the less information the subject receives about the distance between herself and some object, the more likely it is that the subject will confuse a change in the distance of that object with a change in its size. Depth or distance information can be degraded by forcing the subject to view the stimulus only through one eye, for instance, thus reducing depth cues from convergence and binocular disparity. Depth information can be further degraded by eliminating distance cues derived from the position of an object relative to the horizon, texture gradients, and the partial occlusion of farther objects by closer objects in the scene.[115] Absent normal depth or distance information, changes in both object size *and* distance sometimes can be perceived as changes in size. Under such conditions, subjects presumably could 'match' two objects of the same size viewed at different distances by adjusting one of them in size. It would be a mistake in this case, however, to infer that the 'apparent size' of the object is dependent not only on its actual size but also on its distance from the subject. The more sensible conclusion in this case rather is that the distance to an object *can* have an effect on its apparent size under certain impoverished conditions in which subjects are denied many of the contextual cues that normally would aid them in

experiments are subject to such visual illusions because they are viewing stimuli under restricted and artificial conditions in which ordinary cues about illumination conditions are absent. As Hilbert points out, subjects in normal visual contexts typically do not make such errors. In such cases, changes in illumination conditions such as those occurring in shaded regions are themselves visible. This latter fact can be accommodated by the apparent illumination interpretation, however, and does not require us to endorse the colour appearance claim.

[115] See Palmer (1999a: 199-249).

disentangling distance and size information contained in the proximal stimulus. In normal visual contexts in which such cues are present, in contrast, it is eminently plausible that distance or depth information is represented in its own right. Carrying this point over to the case of colour and illumination, we might conclude from the asymmetric matching experiments that the illumination conditions in a scene *can* have an effect on apparent colour under conditions in which subjects are denied many of the contextual cues that normally would aid them in disentangling illumination information and reflectance information. In normal visual contexts, however, it remains eminently plausible that properties of the illuminant are represented in their own right as proposed by the apparent illumination interpretation.

In this section I have discussed a number of arguments in support of the colour appearance claim. Naive appeals to visual phenomenology seem to me inconclusive on this point. Cohen's appeals to psychophysical asymmetric matching tasks also face a number of problems of interpretation. In particular, there are various alternative explanations of these data that seem plausible in their own right and that do not support the colour appearance claim. As things stand, then, the colour appearance claim lacks overwhelming support. This leaves the argument from illumination in fairly dubious standing. By way of a rearguard action, however, in the remainder of this chapter I shall consider what the consequences would be for the reflectance discrimination view if the colour appearance claim turned out to be true. In seeing how the reflectance discrimination view could adapt to this claim, several important explanatory features of the account will come more clearly into view. I shall argue that despite initial appearances to the contrary, the reflectance discrimination view is in fact entirely consistent with the colour appearance claim. In this respect it is much better placed than invariantism in accounting for the putatively variegated nature of colour constancy.

2.3. A Rearguard Action Against the Colour Appearance Claim

The colour appearance claim states that the respect *r* in which surfaces appear different under differing illumination conditions is that of *colour*. As we saw above, the arguments that have been presented to date in support of this claim are inconclusive. In this section, I want to assume that the colour appearance claim is true and explore the consequences. Although this outcome undoubtedly would be extremely bleak for the appearance invariance view of colour constancy, I shall argue that the reflectance discrimination view could accommodate this result. In order to build up to this argument, I shall first discuss some recent attempts by Noë and Cohen to accommodate the colour appearance claim.

2.3.1. Noë and Cohen

When combined with Duality in the argument from illumination, the colour appearance claim has a seemingly puzzling consequence. As we saw in the conclusion of this argument, these premises imply that objects phenomenally appear both the same *and* different in respect of colour across changes in illumination. Let's call this the *two faced view of colour constancy*. At first blush this claim might seem to imply that colour constancy involves visually representing inconsistent or impossible states of affairs. As Hilbert (2005: 146) notes, however, if colour constancy involved representing inconsistent states of affairs,

> Perceiving a red ball with one side brightly lit and the other shaded would then be like perceiving an impossible triangle or an Escher drawing. But this kind of paradox, although possible with colour, is not a feature of our experience of common cases of colour constancy.[116]

Hilbert (2005: 145) thus cautions that in interpreting the two faced view of colour constancy,

[116] Kalderon (forthcoming: 4) makes the same point.

> [It is] important to emphasise that it's not a sense in which objects look to have two colours simultaneously. Rather the printed page looks both similar and different when viewed under the different illuminations but not because there are multiple apparent colours associated with it.

The apparent inconsistency implied by the two faced view therefore must be merely *prima facie*. Our interpretation of this view of colour constancy accordingly requires a more subtle or nuanced account of visual perceptual content.

Historically speaking the most popular way to interpret the two faced view of colour constancy has been via *inferential* or broadly *Helmholtzian* theories of colour vision. Hilbert (2005: 145) describes this type of account as follows,

> One possible response to this combination of change and constancy is to claim that the changing aspect is a purely sensory one and the constancy is the result of inference or judgement… The apparent colour is constantly changing but we are able to infer what it would be under some standard viewing condition and it is this inferred colour that is stable across changes in illuminant and other viewing conditions. This type of view is, of course, the familiar approach to perception that has dominated modern philosophy since the empiricists and still has a powerful grip on philosophical thinking about perception.

The rough thought is that our visual perception of colour is *multi-layered* or *multi-staged* in that there exists a so-called 'sensory core' that represents a range of illumination-dependent properties, from which the visual system unconsciously or subconsciously 'infers' or 'judges' the presence of stable surface properties of objects in the environment.[117] Hilbert

[117] There are forms of the inferential theory on which A) illumination-dependent sensory properties are non-conscious and B) inference or judgement is conscious. Theories which hold A, however, are unsuited to explaining the phenomenon at issue which is after all a matter of an illumination-dependent aspect of conscious

(forthcoming) therefore refers to such views as 'sensory core plus judgement' views of colour vision.

How do inferential theories of colour vision help in accommodating the putatively two faced view of colour constancy? There are at least two ways that one could go here. The first route would be to hold that the sensory core and the inferential or judgemental component of colour vision do indeed represent or attribute inconsistent or contrary colours but that these colours are attributed to *different objects*, thus avoiding outright inconsistency in visual perceptual content. Sense data theorists of a certain stripe, for example, take the sensory core to consist of sense data of which we are directly aware in visual experience, which themselves instantiate colour properties.[118] When we visually perceive a white wall in shade, then, we are thus supposed to be directly aware of a sense datum that instantiates the property *grey* from which our visual systems infer or judge the presence of a physical surface which instantiates the colour property *white*. I do not intend to discuss such views further here.

The second and more plausible route would be to hold that the sensory core and the inferential or judgemental component of colour constancy attribute two or more *different types of property*.[119] Noë (2004) and Cohen (2008) recently have adopted this approach in explaining the two faced nature of colour constancy. Noë distinguishes between 'apparent colours*', which are illumination-dependent properties of which we are phenomenological

---

appearances. Theories which hold B, on the other hand, seem to over-intellectualise visual perception and are therefore inherently implausible. I shall ignore both such views here.

[118] Such changes in visual appearance across differing perspectives or viewing conditions have served as *the* canonical motivation for the sense data theory at least as far back as the British empiricists. Thus Hume (1975: 152) argues that 'the table which we see seems to diminish as we remove farther from it. But the real table, which exists independent of us, suffers no alteration. It was, therefore, nothing but its image which was present to the mind. These are the obvious dictates of reason, and no man who reflects ever doubted that the existences which we consider when we say *this house* and *that tree* are nothing but perceptions in the mind, and fleeting copies or representations of other existences which remain uniform and independent.'

[119] Peacocke's (1983) view seems to combine both the 'different objects' and 'different properties' approaches. Peacocke distinguishes between the ordinary colour properties *red, blue*, and the 'sensational properties', which he labels *red', blue'*. Whereas *red* and *blue* are properties of ordinary material objects, however, *red'* and *blue'* are properties of *experiences*, more accurately, of the *visual field*.

aware, and colours *proper* or 'colour aspect profiles', which are complex dispositions of objects to present various changing apparent colours* when viewed under different conditions and from differing perspectives.[120] Whereas apparent colours* are strictly speaking sensed or seen, the colour aspect profile of an object is merely 'grasped' via its apparent colour*. Noë (2004: 132) explains this thought as follows,

> Each colour can be thought of as corresponding to a unique *colour aspect profile*, a unique range of ways its colour aspect [i.e. its apparent colour] transforms as the relevant kinds of movements (colour-critical changes) occur… To see the colour of an object – to experience which colour it has – is to discover its visual potential or colour aspect profile; it is to grasp how its appearance changes or would change as colour-critical conditions change. To experience something as red, then, is to experience not merely how it looks here and now, but how it would look as colour-critical conditions vary.

One of the more interesting aspects of Noë's (2004: 129) view is that one's ability to 'grasp' an object's colour is dependent on possessing certain sensorimotor abilities,

> Our ability to perceive the wall's colour depends on our implicit understanding of the ways its apparent colour varies as colour-critical conditions vary. At ground, our grasp of these dependencies is a kind of sensorimotor knowledge.

Although colour aspect profiles are not strictly speaking sensed, on Noë's view they are *experienced* as 'present in absence', where the 'presence' of some feature for a subject is closely related to the potential for that subject to interact with the feature via sensorimotor activity (2004: 129),

[120] It is important to note that Noë uses the term 'apparent colour' in a technical sense to denote a distinct type of sensory property, and not in the ordinary sense as in 'the colour that an object phenomenally appears to have'. Accordingly I shall refer to these properties as 'apparent colours*' in order to distinguish them from colour appearances in the latter sense.

> We can experience the wall as uniform in colour *and* as differently coloured across its surface. Just as we see the circularity in the elliptical appearance, so we see the invariant colour *in* the apparent variability. The colour of the wall is present in absence; it is implicitly present.

In much the same way, the rear occluded side of an object is 'present in absence' because we possess the sensorimotor ability to rotate the object and thus interact with the occluded side. To summarise, on Noë's view a white wall in shadow is visually represented both as *apparent grey** and as *white*, the former being an apparent colour* of which we are strictly speaking aware in visual consciousness, the latter being a colour aspect profile that is 'grasped' from the wall's apparent colour* *via* some implicit or unconscious sensorimotor knowledge. Crucially *apparent grey** and *white* are different types of property and hence they are not strictly contraries. Noë's account thus seemingly accommodates the two faced view of colour constancy while avoiding the attribution of inconsistencies in visual perceptual content.

Cohen's (2008) view bears a number of similarities to that of Noë (2004), although his account lacks the distinctive association between our visual abilities and sensorimotor knowledge. Cohen holds that the illumination-dependent variance responses track differences in 'occurrent apparent colours', which are glossed as 'the colours that objects are represented by subjects (/visual systems of subjects) as having'. For Cohen (2008: 81), however, the *in*variance response to a surface presented under differing illuminants reflects a similarity in respect of the *counterfactual property* of being such that these surfaces 'would share an apparent colour if, contrary to fact, both regions were presented under the same illumination'. Cohen calls these properties 'counterfactual apparent colours'. Cohen (2008: 82) thus claims that in the asymmetric matching task described above, the two fields 'differ in apparent

colour but meet the further condition that if, counterfactually, the test patch and the standard patch were presented under the same illumination, then the two would match in apparent colour'.

Before moving on to discuss the reflectance discrimination view's approach to handling these issues, I want to raise a few problems with Noë and Cohen's accounts. An unwelcome consequence of both views is that illumination-independent, stable surface colours are not strictly sensed or seen. On Noë's view, colour aspect profiles are experienced only as 'present in absence'. Cohen holds that stable surface colours are *counterfactual* properties, which by their nature presumably can only be inferred and not sensed. Hilbert (2005: 154) presents the problem for these views as follows,

> Both theories maintain that the constant element in colour experience is not itself sensory. Noë seems convinced that this is phenomenologically correct. His view entails that experience of shaded area of a white surface and the experience of grey surface will be identical in their sensory properties. This claim is, at the very least, not obviously true.

Hilbert's last sentence smacks of understatement. If anything, Noë and Cohen's claim that stable surface colours are not strictly speaking sensed surely *conflicts* with our intuitive interpretation of the phenomenology of colour constancy. This seems to be something that both Noë and Cohen ought to acknowledge. After all, the entire discussion of the two faced nature of colour constancy is premised on Duality, which is supposed to report how objects actually and presently phenomenally appear across changes in illumination, not how objects counterfactually *could* appear.

Another problem with the approach taken by Noë and Cohen is that it is unclear what adaptive biological advantage would be conferred on organisms by the ability to represent

colour aspect profiles or counterfactual appearance properties. As such it is difficult to square this approach with the biological rationale for the evolution of colour constancy discussed in chapter one. Colour aspect profiles and counterfactual apparent colours are properties that determine how objects would appear under certain circumstances. We might wonder, however, what relevance the *appearance* of an object under some set of counterfactual circumstances has to an organism's biological needs in respect of mating, seeking shelter, foraging, predating or avoiding predation. To put the point somewhat glibly, *appearances* could not possibly nourish, threaten, mate, or kill an animal, but *objects and their properties* certainly could! In contrast, the disposition of a surface (or gas or liquid) to interact with incident light provides highly significant information about its chemical constitution and hence about states of sexual arousal, blood flow, surface material, ripeness, and so on. Adopting this biological standpoint, it is natural to suppose that colour constancy could have evolved to enable the discrimination of these stable, illumination-independent, surface properties. From the same naturalistic standpoint, however, it is extremely difficult to see why colour constancy would have evolved to provide information about the counterfactual appearances of objects in the environment.

2.3.2. The Reflectance Discrimination View

How does the reflectance discrimination view fare in accommodating the putatively two faced view of colour constancy? As noted above, it might seem as though the reflectance discrimination view would struggle to explain this feature. On the view that colour constancy constitutively involves only the ability to discriminate reflectance properties across changes in illumination, it is *prima facie* difficult to explain why the phenomenology of apparent colour should be partially illumination-dependent. To put the point another way, the reflectance discrimination view is primarily *reflectance focused* and thus seems to assign no

significant role to information about the illumination conditions in understanding colour constancy. In this section I shall first argue that the reflectance discrimination view is in fact *perfectly consistent* with the two faced view of colour constancy. Nothing in the reflectance discrimination view rules out the possibility that colour appearances in normal human subjects should be partly illumination-dependent. Further to this point, I shall then argue that the reflectance discrimination view provides a *rationale* or *explanation* as to why representations of illumination conditions should sometimes be phenomenally conscious for an organism. This does *not* mean, however, that the view provides a rationale for the colour appearance claim. Although the reflectance discrimination view predicts that representations of illumination conditions could sometimes be phenomenally conscious, it is officially neutral as to *how* these representations might be presented in our visual phenomenology, if indeed at all. In particular the view is officially neutral as to whether these representations are presented phenomenologically as an aspect of apparent colour or some other phenomenological dimension such as luminosity and therefore neutral as between the colour appearance claim and the apparent illumination interpretation of the variance response in colour constancy. In seeing how the reflectance discrimination view can adapt to accommodate such possibilities, various interesting features of the account therefore will be brought more clearly into focus.

Having presented these further arguments in support of the reflectance discrimination view I shall then consider an objection to this account. As noted in the previous chapter, the motivation for reflectance focused accounts of colour constancy derives partly from the supposedly illumination-invariant nature of apparent colour. If the colour appearance claim turns out to be true, however, it might seem once again that some of the initial phenomenological motivation for my view will be undercut. My final aim in this section is therefore to argue that, even granting the truth of the colour appearance claim, it remains the

case that reflectance focused views provide the best model of colour constancy that we have. The argument appeals to considerations of parity between colour constancy and other major types of visual perceptual constancy. The argument crucially provides independent reason to privilege the relatively illumination-independent aspect of the phenomenology of apparent colour in motivating a reflectance focused account of colour constancy.

My first aim is to argue that the reflectance discrimination view is perfectly *consistent* with the two faced view of colour constancy. In the previous chapter, the reflectance discrimination view of colour constancy was presented as follows:

> (DISC) Necessarily, for any organism *x*, *x* possesses colour constancy if and only if *x* has the ability visually to discriminate between the surface spectral reflectances of objects across changes in ambient illumination.

A key feature of this thesis is that the ability to represent differences in reflectance must be reliable across worlds in which the illumination conditions vary, as in the actual world. As we saw in the previous chapter, in worlds in which the illumination conditions are invariant an organism might be able to detect differences in surface reflectance merely by detecting changes in the wavelength composition of the proximal stimulus. This ability crucially would not be sufficient, however, for the possession of reflectance discrimination in the sense at issue. Placed in the actual world, this organism would always mistake a change in illumination for a change in reflectance and therefore would be totally unreliable in representing differences in surface reflectance across changes in illumination.

An important feature of the reflectance discrimination view of colour constancy, then, is that it requires *not only* the ability to detect changes in surface reflectance but also *some independent type of informational sensitivity* to the illumination conditions within the scene, such as would sustain the counterfactual reliability of the organism's representations of

constant reflectance properties across changes in illumination. There are many ways in which an organism might achieve the relevant informational sensitivity to illumination conditions. As has been noted already, the standard computational view is that colour vision systems employ mechanisms or processes to 'disregard' or 'filter out' information relating to illumination conditions *en route* to forming a representation of reflectance. The standard computational view therefore is that the informational sensitivity to illumination conditions displayed by the colour vision system is non-representational, non-perceptual, and non-conscious. This view clearly lacks the resources to explain the putatively illumination-dependent aspect of colour appearances. The reflectance discrimination view might thus seem to inherit these difficulties. As Hilbert (2005: 157) has recently pointed out, however, there 'are no technical reasons that I am aware of that stand in the way of constructing [computational] theories that deliver information about the illuminant as well as the surface reflectance'. Hilbert (2005: 150) thus suggests that 'rather than discarding the illuminant estimate... the current proposal is that it is retained and is part of the visual representation of the scene'. Similar modifications to the standard computational model have been proposed in recent vision science. In a survey article on colour constancy, for instance, Smithson (2005: 1331) suggests that,

> Rather than discounting the illuminant, would it not be more desirable to recognise that surfaces were being viewed under different illuminants, to infer the relative properties of different illuminants, and the identity of surfaces across illuminant changes...?[121]

[121] MacLeod (2003: 99) makes effectively the same proposal. Adelson (2001: 6) proposes that in shadow 'a new perceptual variable, a kind of haziness, has become apparent. Thus at least two values will have to be used to describe each point in the scene'. See also Arend (1994) and Mausfeld (1998).

Zaidi (1998: 1767) likewise suggests that 'the problem for the visual system to solve is not to bring about stable colour appearance under different illuminants by discounting them but to recognize that objects are indeed being viewed under different illuminants and to discover what the illuminant properties are.'[122] Rather than interpreting the variance in apparent colour with changes in the illumination as a flaw in the colour vision system, Zaidi (1998: 1767) wonders why we do not think of this as 'a design feature that allows the observer to extract information about illuminants as well as objects?'

The reflectance discrimination view could adopt a similar strategy in allowing the possibility that organisms form representations of both the reflectance properties and illumination conditions in the scene as part of fulfilling the constitutive function proposed in DISC. It is moreover consistent with the reflectance discrimination view to hold that representations of illumination conditions can be *phenomenally conscious* for an organism. As noted in the previous chapter, the reflectance discrimination view of colour constancy makes no explicit appeal to the visual phenomenology of colour appearances. In particular, the reflectance discrimination view does not impose the constraint on possessing colour constancy that objects must phenomenally appear the same in respect of colour across changes in illumination. As such nothing in the reflectance discrimination view rules out the possibility that some subjects capable of achieving colour constancy perform this discriminatory ability via phenomenally conscious representations of both reflectance *and* illumination. There is moreover nothing in the view that rules out the possibility that reflectance and illumination are represented by distinct aspects or dimensions of phenomenally apparent colour. By disengaging the constitutive account of colour constancy from features of visual phenomenology, the reflectance discrimination view leaves it open as to how information about reflectance properties is made accessible to the organism in making

[122] Zaidi (2001) argues that the standard for colour constancy should be to identify similar materials across illuminant changes while simultaneously extracting 'illuminant colours'.

its visual discriminations. If some organisms happen to make this information available together with information about illumination conditions as aspects of apparent colour, then so be it! The reflectance discrimination view therefore is considerably better placed than appearance invariantism to accommodate the putatively illumination-dependent aspect of phenomenally apparent colour.

I have argued that the reflectance discrimination view is *consistent* with the two faced view of colour constancy and in particular with the claim that apparent colour is partly illumination-dependent. My second aim is to argue that the reflectance discrimination view also provides a partial *rationale* or *explanation* as to why it could be beneficial to an organism in achieving colour constancy also to token phenomenally conscious representations of the illumination conditions in the scene. As noted above, the claim is *not* that the reflectance discrimination view provides a rationale specifically for the colour appearance claim. The proposed rationale is thus entirely neutral as to whether illumination conditions are represented as an aspect of apparent colour or some completely different dimension of visual phenomenology such as luminosity, if indeed they are consciously represented *at all*. Now as we noted above, the constitutive thesis DISC holds that reflectance discrimination requires not only the ability to detect changes in surface reflectance but also some type of informational sensitivity to the illumination conditions within the scene, so as to sustain the counterfactual reliability of the organism's representations of reflectance properties across changes in illumination.[123] We should also recall that discriminating differences in reflectance requires some sort of first personal or whole organism level registration, monitoring, or access to the relevant representations, which in paradigmatic cases is achieved by way of these representations becoming phenomenally conscious. Given

[123] To recap, the ability to represent a constant reflectance property across changes in illumination is *part* of the ability reliably to discriminate differences in reflectance, because in order to be reliable in representing differences in reflectance one must minimise the tendency for false positives.

these features, it is entirely natural to predict that in some cases an organism's discriminations of reflectance properties across changes in illumination will involve phenomenally conscious representations of the illumination conditions *as well as* the reflectance properties in the scene. This is predictable because in making representations of the illumination conditions in principle accessible to the organism in this way, the organism will be *more likely* to avoid mistaking a change in illumination for a change in reflectance and thus *more reliable* in respect of its ability to discriminate differences in reflectance properties. By representing the variant illumination conditions in a way that is accessible to the subject of these discriminations or whole organism, a development or adaptation of this sort would help to sustain the requisite type of counterfactual reliability in the organism's representations of reflectance properties across changes in illumination.

This proposal is closely related to a characterisation of colour constancy offered by Craven and Foster (1992: 1360) who present,

> an alternative and complementary property of colour constancy, as follows: The ability of a subject to correctly attribute changes in the colour appearance of a scene either to changes in the spectral composition of the illuminant or to changes in the reflecting properties of that scene.

In Craven and Foster's view 'it may be irrelevant whether the colour appearance of an object does or does not change under a change in illuminant, providing that any changes that do occur can be correctly attributed'. The reflectance discrimination view is likewise neutral as to whether the phenomenology of apparent colour is partly illumination-dependent. The proposal is rather that consciously representing the illumination conditions *in some way or other* could allow subjects more easily to distinguish a change in illumination from a change

in reflectance, and thereby enable more reliable reflectance discrimination. The key observation made by Craven and Foster is that it does not matter *how* information about the illuminant is made available, so long as the subject is able to use this information to distinguish changes in the illumination conditions from changes in reflectance. It is crucial to note, then, that the putative conscious representation of illumination is not being proposed as an obligatory, necessary, constitutive condition on achieving reflectance discrimination and hence colour constancy. I assume that an organism could achieve high degrees of reflectance discrimination via some non-conscious or non-representational type of informational sensitivity to the illumination conditions. The point is rather that given the need for such informational sensitivity, it *could* be advantageous to an organism in its capacity for reflectance discrimination that information about the illumination conditions is made conscious for that organism. This crucial point allows the reflectance discrimination view to make sense of the putatively illumination-dependent aspect of the phenomenology of colour constancy, while staying true to the view that colour constancy is a *constitutively* reflectance focused visual ability.

My third and final aim in this section is to discuss a potentially damaging objection to the foregoing proposal and offer some replies. In this section I have been exploring the consequences of the colour appearance claim, which holds that apparent colour is partly illumination-dependent. I have argued that the reflectance discrimination view is *consistent* with this claim. In granting the truth of the colour appearance claim, however, it might seem that we have undercut some of the initial motivation for the reflectance discrimination view.[124] The reflectance discrimination view holds that colour constancy is a *reflectance*

[124] A somewhat similar objection to the constitutive account was presented in section 1.2 in the previous chapter. The recurring theme is that the constitutive approach treads a very thin line in appealing to visual phenomenology in motivating parts of the view, while subsequently disregarding features of phenomenology in accounting for the nature of colour vision. The response provided in the previous chapter should begin to show how this approach can be carried out consistently.

focused rather than a *reflectance and illumination focused* visual ability. This view of colour constancy was motivated in the previous chapter by appealing to the supposedly illumination-invariant character of colour appearances across changes in illumination. If the colour appearance claim is true, however, then this characterisation of the phenomenology of colour appearances is mistaken and hence cannot be assumed as part of the motivation for a reflectance focused account of colour constancy. In response to this problem I shall argue that even if the colour appearance claim is granted, it remains the case that reflectance focused accounts provide the most plausible models of colour constancy that we have.

My argument is based on considerations of *parity* between colour constancy and other major forms of perceptual constancy. By way of background to this argument, from the previous chapter the reader should recall that while the reflectance discrimination view is not constitutively associated with any particular features of visual phenomenology, our conscious experience of colour nonetheless plays a crucial role in providing an *initial means of identification* of colour vision *qua* distinctive psychological ability or kind. To put the point another way, the visual phenomenology of apparent colour serves to *fix the reference* of 'colour vision' and thus to enable an initial characterisation of this kind in terms of its easily identifiable features. I assume that the same is true with respect to other visual abilities. The phenomenology of our conscious visual experience provides our initial means of identification for all major types of visual ability including size, shape, depth, orientation, position, and motion vision. As noted earlier in this chapter, Duality-type phenomena occur in all such types of vision. In size constancy, for example, while two objects of the same physical dimensions viewed at different distances clearly appear relatively similar in respect of size there is also a clear sense in which these objects appear different. Just as in this section we have assumed the truth of the colour appearance claim, let's suppose for the sake of argument that the respect in which such objects differ in appearance across changes in

distance is that of *apparent size*. If sound, this claim would entail that there is a highly regular and characteristic aspect of the phenomenology of apparent size that is *distance dependent*. It seems likely that this claim would fatally undermine the appearance invariance view of size constancy. It is important to ask, however, whether this claim about apparent size would also undercut any putative phenomenological motivation or support for the view that size constancy is a *size focused* visual ability, as opposed to a *size and distance focused* ability. For example, would this claim undercut any intuitive phenomenological support for the view briefly sketched above on which size constancy is the ability to discriminate changes in the physical dimensions of objects across changes in their distance from the subject? In this case, would we take the putatively distance-dependent aspect of apparent size to undermine any account of the *nature* of size constancy as being an ability reliably to represent the *size* of an object across changes in viewing perspective? This seems extremely implausible. Whatever the facts regarding the phenomenology of apparent size, I doubt that anyone would contemplate denying that size constancy is *size* focused. It would seem exceptionally rash to argue on the basis of this putative distance-dependent aspect of phenomenally apparent size that size constancy should be reconceived as a partly *distance* focused visual ability. On the contrary it remains by far the most plausible view of size constancy that it constitutively involves the ability stably or reliably to represent the physical dimensions of an object across changes in the distance of that object from the subject. This case therefore presents a strong rationale for downplaying the explanatory significance of the putative distance-dependent aspect of the phenomenology of apparent size in understanding the nature of size constancy.

The foregoing argument regarding size constancy provides an inductive basis for inferring that colour constancy is a reflectance focused perceptual ability, even *granting* that there exists some putatively illumination-dependent aspect of the phenomenology of apparent colour. Generalising on the case of size constancy, it is thus standard in the study of vision

that we gain an initial fix on the phenomenon of *F* constancy, for some visually perceptible feature *F*, via the visual phenomenology of *F* perception. Intuitively it seems that while there is an important sense in which objects phenomenally appear the same in respect of *F* across changes in some set of contextual features *C*, there is also some aspect of the phenomenology of apparent *F-ness* that is *C*-dependent. Despite this putatively *C*-dependent aspect of the phenomenology of apparent *F-ness*, however, the most plausible view as to the physical properties *P* that *F* constancy functions to represent is that they should be properties that are invariant across changes in *C*. Complexities in the visual phenomenology notwithstanding, therefore, *F* constancy is best modelled as a *P focused* rather than *P and C focused* perceptual ability. Extending this model to the case of colour constancy, we normally gain an initial fix on the phenomenon of colour constancy via the visual phenomenology of colour perception. It seems that while there is an important sense in which objects phenomenally appear the same in respect of colour across changes in illumination, there is also some aspect of the phenomenology of apparent colour that is illumination-dependent. Nevertheless the most plausible account as to the physical properties that colour constancy functions to represent is that they should be properties that are invariant across changes in illumination. The properties that most naturally present themselves are surface reflectance properties. Complexities in the visual phenomenology notwithstanding, therefore, colour constancy should be modelled as a *reflectance focused* rather than a *reflectance and illumination focused* perceptual ability. This argument provides strong independent reason to privilege the relatively illumination-independent, invariant, aspect of the phenomenology of apparent colour in providing a constitutive account of this visual ability.

It might be objected that the demand for parity between colour constancy and other types of visual perceptual constancy is illegitimate because colour constancy is *anomalous or exceptional* in a crucial respect. In the case of size constancy we have *antecedent knowledge*

of the physical properties that size constancy functions to represent, namely the physical dimensions of objects, and we can antecedently distinguish these properties from other contextually variant properties such as distance. Given this antecedent understanding of the ontology of size properties, we are justified in privileging the distance-independent aspect of phenomenally apparent size in theorising about the distinctive nature of size constancy. The case of colour constancy, however, is unlike that of size constancy. In the case of colour constancy we have no antecedent knowledge as to the properties that colour constancy functions to represent, precisely because we have no antecedent knowledge as to the ontology of colour. As noted at the very beginning of this thesis, colour vision is assumed to have epistemic and explanatory priority with respect to colour ontology. Our only route to knowledge about colour ontology therefore goes via our understanding of colour vision. Given that we have no antecedent knowledge as to the illumination-independence of colour properties, in this case we therefore have no basis for privileging the illumination-independent aspect of phenomenally apparent colour in theorising about the distinctive nature of colour constancy.

The following remarks by Cohen (2008: 69, fn. 16) help to illuminate this objection to the parity argument,

> [W]e don't need the phenomenon of size constancy to serve as evidence for the conclusion that size is independent of viewing depth. In particular, we have agreed upon standards for size that are constituted independently of viewing depth (these standards are called 'rulers'), and that give us reasons for believing that size is independent of viewing depth. Thus, while rejecting invariantism about size constancy might require a more careful characterization of the phenomenon, it would not undercut our reasons for taking size to be independent of viewing depth. In contrast, invariantism about colour constancy is the main reason for thinking that colours are independent of illumination (and is presented as such by proponents of that view)... Thus, undercutting invariantism about colour constancy leaves the claim that colours are illumination independent wholly unsupported.

Maloney (2003: 329) presents some very similar points regarding the difference between colour constancy and size constancy,

> First of all, when we study how well human observers judge properties of the environment, including shape, or depth, and separation (length), we usually know, or can determine, the correct answer to any question that we pose to the observer. Put a bit more formally, we have agreed-upon measurement procedures that allow us to determine which of two lengths is longer, or which of two objects is further away (e.g. a ruler). We do not expect to disagree with other observers concerning judgements of length... and we readily resolve differences between what we see and what we measure in favour of the measurements, explaining the discrepancy as due to a visual illusion... For colour perception, we typically don't know what counts as the right answer. We don't have measuring devices to tell us the (true) colour of an object; some researchers... even reject the possibility that we could ever identify measurable, physical properties of objects that correspond to the subjective experience of colour. Consequently, the study of colour is typically framed in terms of invariances or constancies: the experimenter doesn't know what colour a homogeneous object "should" be, but has the intuition that whatever it might be, it should remain the same under changes of illumination in the scene. If we were challenged to justify the claim that colour is invariant under a scene transformation, we could not do so. We would reply to the same challenge in the case of length by simply verifying that measured length remained invariant under the specified transformation.

To recap, then, the objection to the above parity argument is that colour constancy and size constancy are disanalogous in the crucial sense that whereas we have antecedent knowledge as to the distance-independence of size properties, we simply have no antecedent knowledge as to the illumination-independence of colour properties. As such, whereas in the case of size constancy we have independent reason to downplay the distance-dependent aspect of phenomenally apparent size in understanding the nature of this visual ability, in the case of colour constancy the putative illumination-dependent aspect of phenomenally apparent colour cannot be dismissed as explanatorily irrelevant. This is because our only assumed source of

evidence as to the ontology of colour derives from colour vision itself, in particular from the phenomenology of apparent colour.

My first response is that we do not need antecedently to assume that colours are illumination-independent properties in order to expect or demand explanatory parity between colour constancy and size constancy. As noted in the previous chapter, the demand for parity between colour constancy and the other main types of visual perceptual constancy largely derives from the emerging scientific view that the colour vision pathway is entirely *unexceptional* in the sense that it is simply one of several visual information processing systems dedicated to recovering representations of a specific perceptible feature in the environment. The request for parity for colour constancy is therefore independent of the view that colours are illumination-independent properties and derives only from a desire to bring the philosophical study of colour vision into line with the broader scientific view of our visual perceptual abilities.

My second response concerns the assumption that our primary source of evidence as to the ontology of colour derives from colour vision itself, in particular the phenomenology of apparent colour. This assumption implies that the study of colour ontology is completely beholden to features of the phenomenology of apparent colour. Operating under this assumption, if the colour appearance claim did turn out to be true, then this would undercut any support for the view that colours are illumination-dependent properties. It is crucial to note, however, that this assumption engenders the mainstream *experiential view* of colour vision introduced earlier. This orthodox philosophical view holds that the phenomenology of apparent colour trumps all other sources of evidence regarding the nature of colour vision, and by association colour ontology. As I have stated at various points already, I believe that this view provides an unnecessarily impoverished basis for understanding the nature of colour vision and an unnecessarily anaemic evidential basis for doing colour ontology. The

constitutive approach developed in this thesis places significant emphasis on the wider scientific view of our visual perceptual abilities. I have argued that this emerging scientific view warrants the demand for parity between colour constancy and other types of visual perceptual constancy such as size constancy. The assumption of parity in itself thus serves as an important source of evidence and constraints on our account of colour vision and hence colour ontology. From this perspective it is entirely justified and indeed explanatory to view the putatively illumination-dependent aspect of apparent colour in the same light as the putatively distance-dependent aspect of apparent size. In the latter case this phenomenological feature certainly would not undermine the view that size constancy is *size focused*. By parity of reasoning I have argued that the partly illumination-dependent phenomenology of apparent colour provides no reason to reject the view that colour constancy is a *reflectance focused* visual ability. The constitutive approach to colour vision of course takes note of these phenomenological data and seeks to explain them where possible, but it is not beholden to them. The key desideratum in my account is rather to present a picture of the nature of colour vision *qua* distinctive psychological kind that coheres with the broader evidence base provided by vision science.

2.4. Conclusions

In this chapter I have discussed an important and well known objection to the appearance invariance view of colour constancy posed by the *argument from illumination*. A key premise in this argument was the colour appearance claim, which holds that in normal cases of colour constancy the phenomenology of apparent colour is partly illumination-dependent. This is clearly at odds with the standard appearance invariance view, which defines colour constancy as the invariance of apparent colour across changes in illumination. This objection also threatens to undermine the reflectance discrimination view, which has *prima facie* difficulty explaining why there should be an illumination-dependent aspect of apparent colour. I

reviewed several arguments in support of the colour appearance claim and found them to be inconclusive. As a pre-emptive defensive measure, I then assumed the truth of the colour appearance claim and explored the consequences for the reflectance discrimination view. After reviewing some recent attempts by Noë and Cohen to accommodate the colour appearance claim, I argued that the reflectance discrimination view is entirely consistent with the colour appearance claim. In addition I argued that the view provides a partial explanation or rationale as to why it could be advantageous for an organism consciously to represent the illumination conditions in the scene in the service of achieving colour constancy *qua* reflectance discrimination. I then turned to consider an important objection to my account. It might seem that if we were to admit the colour appearance claim, this would undercut some of the initial phenomenological motivation for the reflectance focused view of colour constancy. I argued on the basis of parity with other visual perceptual constancies that this is not the case and that we have independent reason to privilege the illumination-independent aspect of apparent colour in theorising about the nature of colour constancy, colour vision, and by association colour ontology.

3. Colour Vision and Seeing Colours

What does colour vision, as a constitutive matter, enable us to see? A straightforward answer is that colour vision enables us to *see colours*. What does this mean? Another straightforward answer is that 'seeing colours', at least in humans, means *seeing red, orange, yellow, green, blue,* and so on. Presumably something analogous is equally plausible in the case of shape vision: shape vision enables us to see *shapes*, that is, to see *squareness, triangularity, circularity,* and so on. So far, so simple.

Things are not quite so simple, however. The most natural reading of 'seeing red, orange, and green' is *categorically committed*, in the sense that it implies the ability to see, in a sense not yet explained, the colour categories *red*, *orange*, *yellow*, and so on. That is to say, on its most natural reading, the ability to 'see colours' implies the possession of *categorical perception of colour (CP).* CP is a vivid feature of our ordinary experience of colour, particularly apparent in the banded or segmented appearance of a rainbow, for example. Colour category terms, such as 'red', 'reddish', 'shade of red', also provide our primary means of communicating about colour and colour vision, both in ordinary language and more scientific contexts. In spite of the experiential vividness and communicative pervasiveness of colour categories, however, I believe that colour vision should be philosophically and conceptually dissociated from CP for colour. CP, I shall argue, is not part of the constitutive nature of colour vision. To put the claim in its most provocative form, *colour vision does not, by its nature, confer the ability to see colours*, in the natural, categorically committed, sense.

The chapter is organised as follows. In the first and second sections I outline the phenomenon of categorical perception of colour, and the dual notion of non-categorical perception of colour. The third section presents the central argument for the claim that

possessing colour vision does not constitutively involve CP. The argument presupposes the account developed in the previous chapters on which colour vision constitutively involves the ability to achieve *colour constancy*, which I take to be equivalent to the ability to discriminate the surface reflectance properties of objects across changes in illumination. The argument proceeds via a thought experiment in which two subjects, exactly alike in respect of these constitutive discriminatory capacities, nonetheless differ in respect of whether they have CP for colour. The putative dissociation between these constitutive abilities and CP provides the primary basis for my claim that possessing colour vision does not constitutively involve the ability to see colours in the natural categorically committed sense.

Although colour CP has not featured prominently in philosophical discussions about colour, the psychological origin of this ability has been a matter of considerable debate in cognitive psychology. There are interesting connections between my argument and recent *neo-Whorfian* theories, which hold that the colour categories apparent in ordinary visual experience are determined or influenced in part by the language that we speak. Such theories complement, and indirectly support, the view that colour CP is not part of the constitutive nature of colour vision, and provide an intriguing alternative explanation as to the origins of our ability to see colour categories. In the fourth section I discuss some fascinating cognitive neuropsychological evidence for the neo-Whorfian view, which also provides support for my central thesis. In cases of *optic aphasia for colours*, subjects retain normal visual discriminatory abilities with respect to colour while becoming severely impaired on explicit colour categorisation tasks. In the evocative phrase of Goldstein and Gelb (1918), such subjects have an 'impairment of the categorical attitude' towards colour. These cases suggest that the possibility of a total dissociation between our constitutive colour vision abilities and CP, on which my argument rests, is not as distant as one initially might think.

3.1. Categorical Perception of Colour

Categorical perception of colour is a vivid and pervasive feature of our ordinary visual experience. The effects of CP are particularly apparent in our visual experience of rainbows, for example, in which continua of light wavelengths appear as a set of bounded, or banded, regions of colour, corresponding to the ordinary colour categories *red*, *orange*, *yellow*, *green*, *blue*, *indigo*, and *violet*. Colour categories also provide the predominant basis for linguistic communication about colour. If asked to name some colours, most of us would of course name *red*, *orange*, *yellow*, and so on. If asked to name some colours that we *see*, most of us would say that we see *red*, *orange*, *yellow*, or *shades* of these categories, various *reds*, *oranges*, and *yellows*, and so on. It is undoubtedly central to our ordinary understanding of colour vision that it involves *seeing colours* in this categorical sense.

How should we characterise colour CP? In the empirical literature, the operational standard for CP is as follows. Take three shades of colour, *a*, *b*, and *c*, such that the pairs *a* and *b*, and *b* and *c*, are separated by exactly the same discriminable distance, (*a* and *c* will be separated by twice that distance). Moreover, *a* and *b* are 'within-category' shades – both shades are within the category *green*, say – and *b* and *c* is an 'across-category' pair – *c* lies within the category *blue*. Now a subject classes as possessing CP for *green* just in case the within-category shades *a* and *b* appear more similar to her than the across-category shades *b* and *c*, despite the fact that they are separated by the same discriminable distance.[125] I shall call these sorts of effects 'differential similarity effects'. The commonest way of assessing CP in an experimental setting is thus to measure subjects' similarity judgements with respect to triads of shades spanning a single category boundary. Figure five illustrates a range of colour chips chosen for this purpose. G1 and G2 are both members of the category *green*, whereas

[125] The reader should recall the discussion of the similarity effects implied by the barometer categories *fair*, *dry*, and so on, in the introduction to this thesis.

B1 and B2 are both members of the category *blue*. Subjects with CP for *green* reliably judge that G2 is more similar to G1 than to B1, and subjects with CP for *blue* reliably judge that B1 is more similar to B2 than to G2, despite the fact that G1 and G2, G2 and B1, and B1 and B2, are separated by the same discriminable distance.

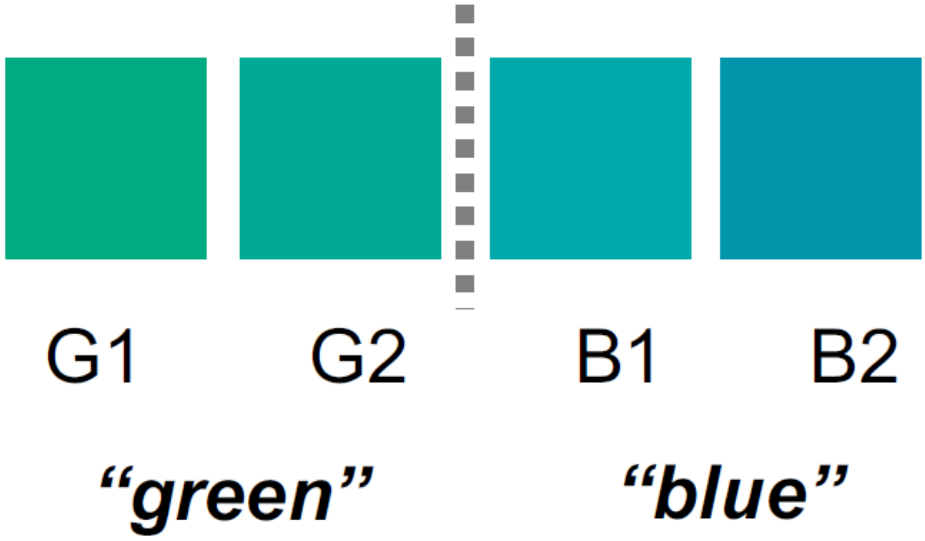


Figure 5: Siok et al. (2009)

Testing subjects' similarity judgements in this way is an example of a so-called '*implicit*' test for CP, a test that does not require the subject to attend to, or report on, the category membership of the colour stimuli. As it provides the simplest and most precise way of assessing whether a subject possesses CP for a given category, this standard will be assumed in what follows.[126] There are other, so-called '*explicit'*, tests for CP, which require either the comprehension or production of a colour category name, (or indeed the word 'category/ies'), in relation to a stimulus or stimuli; the selection of the 'best' or 'focal instance' of a given colour category, given some suitable range of stimuli; or the active sorting of colour chips into categories or groups. It seems entirely plausible that these explicit

[126] Another implicit test for CP is a 'recognition memory' task, in which subjects are first shown a colour stimulus *A*, which is followed after some delay by two further stimuli *B* and *C*, and then are asked to judge which of *B* and *C* was the colour previously shown. The expectation is that subjects will be more reliable in selecting which colour was previously seen, if *B* and *C* straddle a category boundary, than if *B* and *C* are within a category.

tests pick up on important and distinctive features of colour CP, which are not captured by the implicit tests for differential similarity effects. The distinction between implicit and explicit ways of assessing colour CP will reappear in section four.

Several clarifications are now in order. First, the term 'category' is standardly used in psychology in an objective sense: categories are simply groupings of objects, or extensions. Markman and Ross (2003: 592-93), for example, say that the term 'category' is 'used to denote [a] set of items in the world', which I take to mean that an expression such as 'the category *red*' denotes a set of objects – presumably the extension of the English predicate 'red'.[127] A simple extensional view of colour categories cannot be quite right, however. Objects can change their colour over time, and across worlds, so perhaps the category *red* is perhaps better associated with the intension of 'red', that is, triples of worlds, times, and extensions. In saying, at world *w* and time *t*, that an object *a* is a member of the category *red*, then, I mean that *a* is in the extension of 'red' in *w*, at *t*.[128] We could generalise to say that, for any category denoted via the canonical form 'the category *C*', where 'C' is a predicate of a language *L*, for worlds *w* and times *t*, we have the following schema: an object *a* is a member of the category *C* in *w* at *t* iff 'C' is true of *a* in *L* in *w* at *t*.[129]

Second, the foregoing standard for CP leaves open many questions regarding the nature or constitutive basis of this phenomenon. This is largely a matter of dialectical propriety, for the conclusion of this chapter will be that CP is not part of the constitutive nature of colour vision. It is relevant now to ask, however, whether colour CP is a

[127] In some experimental contexts, categories are derived from performance on categorisation tasks. Subjects might be asked, for instance, to *form* categories, or groupings, by placing certain colour chips together, in accordance with certain rules. The resulting groupings are generally taken to indicate the nature of the subjects' representations of superordinate colour categories.
[128] Even this cannot be quite right, however, for objects needn't be red all over. Perhaps the relevant extensions therefore will need to be object parts, or property instances of some sort. I shall ignore this complication.
[129] Note that this is compatible with the claim that translational equivalents, such as *red* and *rouge*, are in fact different categories. The issue of linguistic relativity in colour categorisation will arise in section three.

representational or *non-representational* phenomenon.[130] That is, does possessing CP for *red*, say, imply that one is able to represent the category *red* in one's visual experiences? Much of the philosophical literature seems to assume a positive verdict on this matter. Boghossian and Velleman, for instance, assume that 'ordinary discourse about colour reports the *contents* of visual experience', (1989: 98, italics added). Given that ordinary discourse about colour predominantly involves category terms such as 'red', 'reddish', 'shade of red', and so on, this assumption implies that visual experience represents colour categories such as *red*. Jackson and Pargetter (1987: 68) are also explicit on this point: 'what is most obvious about the colour terms, "red", "yellow", "blue", and so on is that we use them to denote properties that we take to be presented to us in visual experience'.[131] In what follows, then, I shall assume that it is a necessary condition on possessing CP for some colour category *C* that one is able to represent the category *C* in visual experience.[132]

It might be objected that a subject could see *red*, for example, without being able to represent *red* in visual experience, so long as she is able to register or represent *in some way* the chromatic properties of an object which *in fact* belongs to the category *red*, (at *t*, in *w*). To put the point another way, perhaps a subject could see the colour of a red object, without seeing, or being able to represent, that colour *as* red.[133] By analogy, it is plausible that one could see the shape of an icosagon, (a twenty-sided shape), in some sense, without seeing, or being able to represent, that shape *as* an icosagon. Although I have no problem with some such deflationary notion of 'seeing red', it should be clear that a subject could class as 'seeing red' in this sense even if they lacked any of the distinctive warping or banding effects

[130] This distinction is briefly discussed in Palmer (1999b: 929).
[131] See also Tye (2000: 165) and Byrne & Hilbert (2003: 5).
[132] Necessary, but not sufficient: it is conceivable, for example, that one could represent the category *red* in visual experience always as a constituent of some complex content, such as *not-red*. Intuitively, however, a subject representing all objects as *not-red* would not class as possessing CP for *red*.
[133] Although the locution 'seeing *as* red' is useful to fix ideas, I do not want to commit to the view that colour CP is a type of 'seeing *as*', in the familiar sense in which we can see an object *as* a duck, or *as* a rabbit. On this reading, 'seeing as' probably implies possession of concepts for *ducks* and *rabbits*. In contrast, I think that nonconceptual creatures can possess colour CP.

distinctive of CP. The deflationary notion therefore should be distinguished from the categorically committed notion of 'seeing colours' associated with possessing CP.

One possible objection to the very notion of CP is that surely *all* perception is as of some category, type, or kind, and hence surely *all* perception is categorical. One thought underlying this objection might be that perceiving an object necessarily involves perceiving that object under some aspect, mode of presentation, or via the application of some perceptual attributive.[134] Ergo, all perception must involve categorising an object as being one way or another. Another thought might be that perceptual systems 'go digital' at a very early stage in processing, categorising certain light gradations as *edges*, or certain spatial configurations as *blobs*, and so on. Pylyshyn (1999: 361), for example, claims that 'the very fact that the mapping from the distal environment to a percept is many-to-one means that the visual system induces a partition of the visual world into equivalence classes... This is another way of saying that vision divides the perceptual world into some kinds of categories'. Either way, the 'categorical' in 'categorical perception' would appear redundant, and the notion of CP devoid of independent interest.[135]

In response, let's assume that Pylyshyn is correct that all visual systems 'categorise' in the minimal sense that they respond in the same way, or enter into the very same type of state, given a wide range of worldly stimuli. This is a simple corollary of the fact that we have limited visual acuity, and accordingly limited powers of visual discrimination. For example, lines differing by only 1nm in length typically illicit exactly the same visual responses in us, *ceteris paribus*, and will therefore be visually indiscriminable. We could

[134] Perhaps we could have said that we cannot perceive objects as 'bare particulars'. On the role of perceptual attributives in objective perception, see Burge (2009).

[135] Livingston et al. (1998: 732-33) describe CP as an effect 'in which a continuous physical dimension such as intensity (of mechanical or electromagnetic energy) or frequency (of electromagnetic radiation or of acoustic vibration) is partitioned into discontinuous or bounded regions by some kind of perceptual mechanism'. This encourages the interpretation of CP as simply involving many-to-one mappings from stimuli to perceptual states.

accordingly say that our visual systems 'categorise' these lines as being of the same length. But this is not the sense of 'categorisation' that is at issue in CP. As should seem obvious, our visual perception of length does not display the characteristic banding, warping, segmentation, or grouping effects, which are so distinctive of our experience of colour. Presented with a series of one hundred lines steadily increasing in length, for example, there would not be any visually apparent segmentation of this series into regions corresponding to categories of length such as *short*, *medium*, or *long*.[136] Pylyshyn's minimal notion of 'categorisation' is thus insufficient for CP in the sense at issue.

This last point can be made more concrete by considering the fact that, by our assumed standard, a subject possesses CP for some colour category *C* just in case within-category colours *a* and *b* appear more similar than across-category colours *b* and *c*, *despite the fact that these pairs are separated by the same discriminable distance*. It follows from this characterisation that *C* cannot be a category of indiscriminable colours. For if it were, *a* and *b* would be indiscriminable whereas *b* and *c* would be discriminable, which would mean that these pairs were not separated by the same discriminable distance.[137] The moral here is that CP always involves the categorisation of objects or features that are discriminated by some other, typically more basic, perceptual process. Bornstein (1981: 140) makes effectively the same point by saying that in this context the term 'category' signifies not '*identity categories*', but '*similarity categories*', on which 'groupings of nonidentical wavelengths which are perceptually discriminable… are treated as similar'.[138] Possessing CP with respect

[136] It is irrelevant whether we would *verbally* categorise these lines as 'short', 'medium', or 'long' in length.
[137] It would still be true, of course, that *b* appears more similar to *a* than *c*, given that *a* and *b* are indiscriminable whereas *b* and *c* are discriminable. There are of course serious questions as to whether such equivalence classes of indiscriminable colours exist given the non-transitivity of indiscriminability. For extended discussion of this issue see Williamson (1990).
[138] As Bornstein (1981: 142) notes, some early models of CP for speech assumed that within-category speech sounds were indiscriminable. This assumption has since been rejected.

to some category *C* therefore entails that one is able to represent determinates of *C*.[139] Equivalently, possessing CP for some general feature *F* (e.g. colour, shape) entails the ability to represent properties that are determinate relative to *F* (*red*, *square*), but determinable relative to the finest grain determinates of *F* (*particular shades of colour*, *particular shapes*) that one is able to discriminate visually.[140]

Does this mean that any subject capable of representing both determinates and determinables of some feature *F* in visual experience qualifies as possessing CP for that feature? It seems not. For example, some would propose that as well as representing particular determinate lines, visual perception also represents lines under determinables such as *straight* and *bent*. In addition, perhaps in depth or distance perception objects perceived on or near the horizon are represented as *far*, whereas objects perceived in the immediate vicinity are represented as *near*. Nonetheless, we do not possess CP for the categories *straight* and *bent*, or *near* and *far*. More precisely, there are no cases in which given three lines (objects) that are pairwise equally discriminable in respect of curvature (distance), the first pair appear more similar in respect of their curvature (distance) than the second pair. What is distinctive of CP, then, if not the ability for both determinate and determinable representation of some feature? It seems absolutely central to possessing CP for *F* that some determinates of *F* appear, in some metaphorical sense, to 'go together' in visual experience. In the case of CP for colour, determinate hues in certain regions of the spectrum appear *grouped* or *banded* together. A useful analogy here is the Gestalt notion of *perceptual organisation* or *grouping*. Gestalt grouping principles relate to which objects, (broadly construed so as to include lines, dots, regions, etc.), appear grouped together in visual

[139] Of course, these discriminated properties will almost certainly not be *super*-determinates, in the sense of Funkhouser (2006). Again, this is a corollary of our limited visual acuity.
[140] Someone representing *only* the categories *red*, *yellow*, *green*, but unable to make any discriminations within these categories, therefore wouldn't class as possessing CP, and thus somewhat surprisingly would not class as 'seeing colours' in the intuitive categorically committed sense at issue. Again, this shows that being able to represent *red* is insufficient for possessing CP for *red*.

experience. Figure six provides a well known example, in which an array of dots can be seen either as *columns*, or *rows*, depending on factors such as the relative horizontal and vertical distance between each dot. While I do not claim that CP is a Gestalt effect, the analogy highlights something extremely distinctive about the phenomenon that is otherwise left out by the characterisation in terms of determinate and determinable representation.[141]

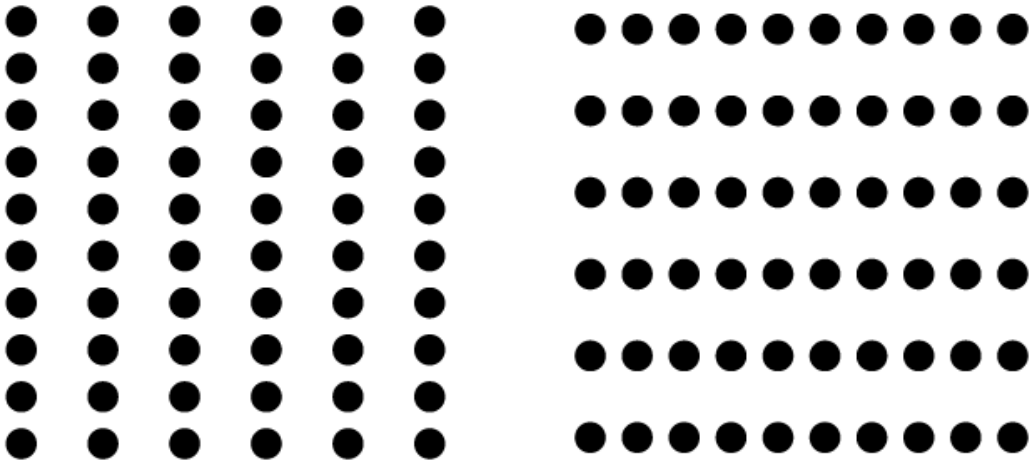


Figure 6: Columns and Rows

3.2. Non-Categorical Perception of Colour

In the remainder of the chapter, the notion of *non*-categorical perception of colour plays a crucial role. I shall therefore offer an explication of this notion, both directly and by way of

[141] For wide ranging discussion of organisational principles in visual psychology, see Palmer (1999a: Chapter 6). Peacocke (1992) is one of few people to focus attention on such effects in the context of philosophy of perception. These phenomena lead Peacocke to distinguish between two levels or types of perceptual content. *Scenario content* provides 'a way of locating surfaces, features and the rest in relation to' a labelled origin and set of axes, together with 'an assignment to the labelled axes and origins of the scenario of real directions and places in the world which fall under the labels' and an assigned time, (1992: 107-8). For each point in the scene, the scenario content specifies 'whether there is a surface there, and if so what texture, hue, saturation, brightness and temperature it has at that point, together with its degree of solidity', (1992: 107). In contrast, *protopropositional content* attributes properties such as 'SQUARE, PARALLEL TO, EQUIDISTANT FROM, SAME SHAPE AS, and SYMMETRICAL ABOUT', (1992: 119). Crucially two perceptual experiences can possess the same scenario content but different protopropositional contents. The phenomenal differences related to perceptual organisation, as in the *row*/*column* or *square*/*diamond* cases, are explained by appealing to differences at the level of protopropositional content. Although Peacocke does not extend these notions to the case of colour, they might ultimately prove useful in understanding the differences between the representational contents associated with the constitutive discriminatory function of colour vision and those associated with putatively organisational or structural phenomena such as categorical perception of colour. These extensions will be pursued in future work.

analogy with the non-categorical perception of speech. From the previous section, the assumed standard for CP for a colour category *C* was that shades *a* and *b*, which lie within category *C*, appear more similar than shades *b* and *c*, which straddle the boundary between category *C* and some other category *C'*, despite the fact that *a* and *b*, and *b* and *c*, are separated by the same discriminable distance. Non-categorical perception (NCP) with respect to a colour category *C*, is defined as the failure to meet this standard. That is to say, for a subject with NCP with respect to *C*, the shades *a* and *b*, which lie within *C*, and the shades *b* and *c*, which straddle the boundary between *C* and *C'*, would appear *equally similar*. That is, pairs of neighbouring shades that are homogeneous in respect of their discriminability from one another would consciously appear homogeneous in respect of their similarity to one another. To illustrate this point, consider figure seven. The square at the apex, *a*, and bottom left, *b*, of the triangle both fall within the category *dark blue*, whereas the square at the bottom right, *c*, falls within the category *light blue*. English speakers fail to meet the operational standard for CP with respect to either *dark blue* or *light blue*. It follows that English speakers are non-categorical perceivers with respect to the categories *dark blue* and *light blue*.[142]

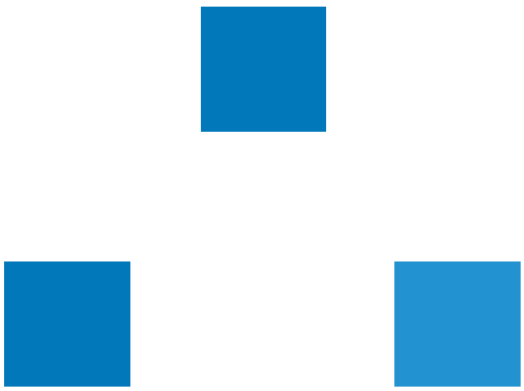


Figure 7: Winawer et al. (2007)

[142] Winawer et al. (2007: 7781) explain that category boundaries for *light blue* and *dark blue* were determined by a linguistic colour classification task for English speakers. As it happens, speakers of Russian, in which the light blue/dark blue distinction is marked by the basic colour terms 'goluboy' and 'siniy', display CP effects for these shades. I discuss the significance of this finding in relation to the Whorfian view of colour categorisation below.

The notion of NCP should not seem contentious or surprising: there are obviously far more categories for which we are non-categorical perceivers than for which we are categorical perceivers. Extending this notion, however, it seems perfectly conceivable that a subject could fail to exhibit CP with respect to *any* superordinate colour categories. Let's call such a global lack of CP *total non-categorical perception*, or *total NCP*. A subject with total NCP would experience the colours of the rainbow in a way which is quite unfamiliar, and which might be somewhat difficult for us to grasp. A subject with total NCP would visually experience the physical continuum of light wavelengths in a rainbow as a non-banded, non-segmented, continuum of hues. The spectrum would appear as a *pure flux*, altogether lacking discrete bands of colour. This is not to say that someone with NCP for colour would see all regions of the spectrum as chromatically homogenous: colour appearances would change continuously across the spectrum. What would be lacking is the familiar superordinate organisation, grouping, or categorical structure of these spectral hues which is so vividly apparent in our ordinary visual experience of colour.

To help get a better grip on total NCP, consider your visual experience of a sequence of shades within the category *red*, varying continuously from bright scarlet through to crimson. These shades do not all appear the same to you: you are able to discriminate differences between suitably separated pairs of shades all along the sequence. Moreover, shades that are closer together in the sequence appear more similar in respect of colour than shades that are farther apart. But I stipulate that there is no other respect – no *categorical* respect – in which shades at different positions in the sequence appear different, as there might be if the sequence ran from red through to orange.[143] A subject with total NCP would experience the entire colour spectrum in a way which is structurally analogous to your

[143] There is, of course, a respect in which all of the shades in the sequence appear the *same* to you; namely, in respect of their all belonging to the category *red*. This apparent respect of superordinate sameness would be lacking in the case of NCP.

ordinary experience of shades of red. The subject would be able to discriminate between neighbouring shades all the way along the spectrum – these shades would not all appear the same to her in respect of colour. The shades would appear to vary continuously from one end of the spectrum to the other, but there would be no other respect in which shades in one broad region appear different to shades in any other region. Two shades (*a* and *b*) close together in the sequence would appear more similar than two shades (*a* and *c*) which are further apart. But there would be no respect – no *categorical* respect – in which any two shades (*a* and *b*) appear the same, and in which any other shades (*a* and *c*) appear different.[144]

In grasping the notions of NCP and total-NCP, it may also be beneficial to consider an analogy with *speech* perception. CP for speech in some language *L* occurs when acoustic signals varying continuously along dimensions relevant to *L* speech are auditorily experienced as distinct linguistic units or *phonemes*. It is possible to produce acoustic signals which vary in this manner using a synthetic speech producer. One example is the dimension of Voice Onset Time (VOT), which is relevant to speech in many languages including English. The opposite ends of the VOT continuum correspond to voiced and voiceless stop consonants: voiced sounds are those with a short VOT, whereas voiceless sounds have longer VOTs.[145] Certain pairs of English phonemes lie at opposite ends of this continuum: /b/ (voiced) and /p/ (voiceless); /d/ and /t/; and /g/ and /k/. Using synthetic speech, we can set up VOT continua from /b/ to /p/ or from /d/ to /t/. English speakers report a categorical shift at some point along both continua: the continua of acoustic signals sound to them as if they are segmented into two discrete phonemic units. There are other acoustic continua, however, along which English speakers make no categorical phonemic distinctions. An example is the

[144] Perhaps subjects with total NCP represent all shades in the spectrum under the superordinate category *colour*. Even so, given the clarifications provided above, this would not class as CP for colour in the sense at issue.
[145] VOT relates to the 'time of onset of laryngeal vibration relative to the release gesture for the plosive', (i.e. relative to stopping airflow in the vocal tract), Rosen & Howell (1987: 113). Stop consonants are those produced by stopping airflow in the vocal tract.

retroflex-dental continuum between retroflex /d/ and dental /d/, a phonemic distinction made in Hindi but not in English. Tone languages such as Mandarin Chinese make phonemic distinctions along the continua of pitch or pitch contour, whereas such distinctions are not made in English. An example is the phonemic distinction in Mandarin between first or high-level tones (steady high pitch sounds) and second or rising tones (rising from mid-level to high pitch). The acoustic continuum ranging from a high-level tone to a rising tone would sound completely continuous to an English speaker, with no experienced segmentation of this acoustic interval into discrete phonemic units. English speakers, then, are non-categorical perceivers with respect to high-level and rising tone phonemes.[146]

As in the case of NCP for colour, we can extrapolate from the restricted notion of NCP with respect to some particular phonemic category to define a notion of *total* NCP for *L* speech. Most people have had auditory experiences of streams of speech which they can categorise into relatively few phonemic units. Extending on such experiences, we could conceive of a language for which we discern no phonemic structure whatsoever: all of the language's phonemic distinctions are made along acoustic continua that we perceive non-categorically, such as pitch continua. If we were to encounter such a language, our auditory perception of its speech would be analogous to the putative total NCP of the colour spectrum. We would experience all acoustic dimensions relevant to the expression of phonemes in that language as pure auditory continua with apparently no discrete phonemic structure.[147] Having introduced and explained the notions of CP and NCP, in the next section I argue that a subject could meet the constitutive standard for normal colour vision defended in the previous

[146] The variable here would be the difference in pitch between the end tone, and the tone at which the sound commences: a high-level tone would thus have value zero, as the sound has a constant high pitch, whereas a rising tone would have some positive value, as the pitch at which the sound commences is lower than that at which it ends. For illuminating discussion relevant to philosophy of auditory perception and sound, see O'Callaghan (2010).

[147] It should be noted that there is a phenomenological difference between experiencing a sound *as* speech versus experiencing it mere as a sound. The categorical/non-categorical distinction for speech thus assumes that the speech sounds are heard *as* speech, but not as phonemically distinct units of speech. If you like, subjects are able to represent auditory signals as phones but not phonemes. See O'Callaghan (2010).

chapters, and yet possess total NCP for colour. This argument provides the central support for the claim that colour vision does not enable us as part of its *nature* to see colours in the categorically committed sense.

3.3. The Dissociation of Colour Vision from Categorical Perception of Colour

The argument of this section is summarised as follows:

> 1. Necessarily, for any organisms *x*, *y*, *x* and *y* are alike in respect of their ability to perceive colour iff *x* and *y* are alike in respect of the visual discriminations they are able to make between the SSRs of objects across changes in ambient illumination.
>
> 2. It is possible that a subject with exactly the same surface reflectance discrimination capacities as a normal human could completely lack CP for colour.[148]
>
> 3. Therefore, it is possible that two organisms exactly alike in respect of their colour vision abilities could differ in respect of whether or not they possessed CP for colour.
>
> 4. Therefore, it is not part of the constitutive nature of colour vision that a subject is able to represent visually any superordinate colour categories; that is to say, colour vision does not, as a constitutive matter, provide subjects with the capacity for CP for colour.

The first premise is the constitutive thesis DEG which I introduced towards the end of chapter one. The second premise states that the key discriminatory capacity that is constitutive of colour vision is dissociable from CP for colour. This premise is supported by a thought experiment involving two subjects exactly alike in respect of their discriminatory

[148] As will become clear, I ultimately believe that the type of possibility in question is nomological, and not merely logical or metaphysical.

capacities, only one of whom possesses colour CP. It is further supported in the next section by some fascinating cases in cognitive neuropsychology. The conclusion of the argument is that CP is not part of the constitutive nature of colour vision.

Lest there be any immediate confusion, I offer the following clarifications. First, as should be clear from the discussion so far, I do not dispute that our ordinary conscious experience of colour involves CP. I take it to be well established empirically that in ordinary cases humans that possess colour vision also perceive colours categorically. The constitutive account thus implies neither that colour vision *does not*, nor that it *cannot*, involve (in some sense) seeing colour categories. The thesis is rather that CP is an inessential feature, or not part of the nature, of colour vision: an organism could class as a colour perceiver without seeing any colour categories. Moreover and perhaps more interestingly, two organisms could be exactly alike in respect of their colour vision abilities by the proposed constitutive standard, and yet differ in respect of whether they possess CP for colour. As far as understanding the nature of our own colour vision abilities is concerned, then, CP does not form part of the picture. Or so I shall argue.

Second, the claim that CP is not *part of the nature* of colour vision is entirely compatible with the possibility that CP is carried out in a neurophysiological region which forms a *proper part* of the colour vision pathway. It is likewise entirely possible that the mechanisms subserving CP are so anatomically and causally intertwined with the mechanisms subserving colour vision that it will be difficult or impossible to distinguish them neurophysiologically.[149] By analogy, consider that in most cases, human beings have *arms* as proper parts, but having arms clearly is not *constitutive* of being a human being. No plausible account of species individuation takes the feature, *having arms*, as a modally

[149] See Yoshioka, Dow, and Vautin (1996) on the neurophysiology of colour categorisation, and Davidoff (1997) on neuropsychological findings that suggest localisation of categorisation centres in the brain.

necessary or sufficient condition on membership of the species *homo sapien*. Equally, even if it were impossible to distinguish physiologically the mechanisms subserving my ability to ride a bike from the mechanisms subserving my ability to ride a skateboard, we should not therefore conclude that the ability to ride a bike is *constitutive* of the ability to ride a skateboard, or *vice versa*. It is clearly possible, indeed often actual, that an individual should possess one of these abilities and not the other.

Given the discriminatory characterisation of colour constancy and colour vision provided in chapter one, there might seem to be a relatively straightforward way to show that colour vision *per se* does not constitutively involve categorical perception of colour. Consider a colour vision system which is *completely basic* or *primitive* by this constitutive standard. For instance, consider an organism capable of making only *one* type of surface reflectance discrimination: presented with a suitably representative range of surface reflectances *R1*, *R2*, ... R_n, the organism would be able reliably to discriminate instances of *R1* from instances of *R2*, but would be completely unreliable in detecting *any* other differences in surface reflectance.[150] While *R1* and *R2* reliably appear to be different to the organism, say, the appearance of any reflectance *Ri*, such that $i > 2$, is so indeterminate or shifty that the organism cannot reliably detect whether R_i looks different to *R1* or *R2*. In virtue of its minimal reflectance discrimination abilities, the organism counts as possessing a minimal degree – a truly basic or primitive form – of colour vision. It seems intuitive, however, that this organism would not display the distinctive banding or similarity effects which are characteristic of CP. This case thus seems to show that colour vision *per se* does not as a constitutive matter provide subjects with the capacity for categorical perception of colour.[151]

[150] For the purposes of this argument, I stipulate that the sequence contains reflectances at discrete intervals and therefore is not dense.
[151] I thank Martin Davies for stressing the relevance of such primitive cases.

This point is well and good as far as it goes, but it goes not nearly far enough. The reader will recall that our assumed standard for CP requires the presence of differential similarity effects across at least three discriminable hues. In other words, CP involves the representation of determinable *similarity categories* among features that are discriminated by some other, typically more basic, perceptual process. It is predictably impossible, therefore, for an organism capable of making only one type of reflectance discrimination to meet this standard. This might seem something of a shallow victory, as it remains an open possibility that organisms possessing more *sophisticated*, or *full-blown*, forms of colour vision – such as humans and many other animals – might thereby possess CP. It is conceivable, for instance, that CP could be an *emergent* or *supervenient property* of constitutive colour vision systems, a capacity for representation that surfaces once organisms reach a certain degree of discriminatory sophistication. If this was the case it would follow that any organisms possessing a degree of colour vision comparable to that of normal humans would as a matter of metaphysical necessity possess CP for colour. Given that our primary philosophical concern undoubtedly lies with the colour vision abilities possessed by normal humans, I therefore intend to argue that even *our* colour vision abilities do not by virtue of their constitutive nature equip us with the capacity for CP for colour. That is to say, it is not part of the nature of *our* ability to perceive colour that we are able to see colours in the intuitive categorically committed sense.

I shall now argue for premise two, which holds that it is possible that a subject with exactly the same surface reflectance discrimination capacities as a normal human, and thus exactly the same in respect of colour vision by the foregoing constitutive standard, could completely lack CP for colour. The argument centres on a thought experiment involving two subjects: CAT, a normal adult human, and NONCAT, who has exactly the same visual discrimination abilities with respect to surface reflectance properties as CAT. That is to say,

schematically, given two objects *a* and *b*, with reflectances R_i and R_j, CAT can visually discriminate *a* from *b* in respect of reflectance iff NONCAT can visually discriminate *a* from *b* in respect of reflectance. From principle DEG, this is to say that by the lights of our constitutive standard for colour vision, CAT and NONCAT are exactly alike in respect of their ability visually to perceive colour. Now consider a context (a world *w*, at a time *t*) in which a colour shade *b*, which is agreed to belong to the category *red*, is viewed by CAT and NONCAT alongside another shade *a*, slightly different from *b*, but which is also in the category *red*, and a third shade *c*, also differing slightly from *b*, which is agreed to belong to the category *orange*. The shade pairs *a* and *b*, and *b* and *c*, (not *a* and *c*), are separated by the same *discriminable distance* for CAT and NONCAT: both subjects can *just discriminate a* from *b*, and *b* from *c*, say.[152] Neither the physical constitution of the three surfaces in question, the light which is incident upon them, nor any other contextually relevant condition, differs across the presentations of these shades to CAT and NONCAT. Now suppose, however, that CAT perceives *b* to be more similar to *a* than to *c*, whereas NONCAT perceives *b* to be just as similar to *a* as to *c*. This is to say that CAT, but not NONCAT, meets the assumed standard for CP with respect to *red*: CAT exhibits the requisite similarity effects at the boundary of *red*, whereas NONCAT does not.[153]

So far, I have argued that NONCAT has NCP for the category *red*. By extension, consider *any* colour category *C*, and any suitably placed triad of discriminable shades *a'*, *b'*, and *c'*, for which CAT exhibits the differential similarity effects required for CP with respect to *C*. By the stipulation regarding their discriminatory capacities, the shades *a'*, *b'*, and *c'*, will equally discriminable to NONCAT and CAT. But now suppose that NONCAT fails to exhibit the differential similarity effects required for CP with respect to *C*. NONCAT would

[152] This is to say that the pairs are separated by one *just noticeable difference* (JND). See Palmer (1999a: 671-72).
[153] I assume that NONCAT also fails to meet any of the other implicit or explicit standards for CP noted in section one.

thereby fail to exhibit CP for *any* of the categories for which CAT exhibits CP. Moreover, I stipulate that there are no *other* colour categories for which NONCAT exhibits CP, but for which CAT doesn't. NONCAT therefore would be exactly similar to CAT in respect of her reflectance discrimination abilities, and yet would have *total* NCP for colour. If genuinely possible, this scenario establishes that a subject exactly like a normal human in respect of the constitutive standard for colour vision could lack any form of CP for colour.

Does NONCAT present a genuine possibility? Is it possible that a subject could possess exactly the same reflectance discrimination abilities as a normal human, and yet completely lack CP for colour? It might be thought not, for CP is often if not typically characterised as a *discrimination effect*. That is to say, CP is typically characterised by the possession of finer 'discriminatory capacities' with respect to the colours of objects that lie across category boundaries than between the colours of objects that lie within the same category. In the introduction to a seminal collection of papers on CP, for example, Harnad (1987: 3) states that CP occurs when 'the subject can discriminate smaller physical differences between pairs of stimuli that straddle that boundary than between pairs that are within one category or the other'. In another seminal study, Bornstein and Korda (1984: 208) say that the central characteristic of CP with respect to some attribute is the 'nonmonotonic character of discriminability along their correlate physical dimensions'. On this view, two subjects differing in respect of CP for colour *ipso facto* would differ in their discriminatory capacities. This suggests that the scenario involving CAT and NONCAT does not present a genuine possibility after all, for CAT and NONCAT were stipulated to be exactly the same in respect of their reflectance discrimination abilities and yet utterly different in respect of CP for colour.

I offer three points in response. First, suppose that there is sufficient evidence of regularities in subjects' discrimination performance at colour category boundaries to support

a psychophysical law or generalisation governing the relationship between discrimination and CP. Even if this were the case, the proposed relationship would be at best nomologically necessary. It is still presumably *metaphysically* possible that a subject's CP with respect to some category could vary independently of her discrimination performance. This would establish NONCAT as a metaphysical possibility, if not a nomological possibility, which is all that is needed to establish the third premise.

Second, suppose for the sake of argument that the scenario involving NONCAT is not even metaphysically possible. That is, suppose it is metaphysically necessary that any organism with exactly CAT's discriminatory abilities possesses CP for the categories *red*, *yellow*, *green*, and so on. We can imagine a new case involving NONCAT2, who is exactly like CAT except that her discriminatory abilities vary just enough at category boundaries to guarantee that she has total NCP. It would also be metaphysically necessary that any organism with exactly NONCAT2's discriminatory abilities would completely lack CP for colour. Let's define the *cluster* be the set of all actual and possible creatures with the same discriminatory abilities as either CAT or NONCAT2. Organisms in the cluster are *extremely similar* in respect of their colour vision abilities and yet differ as a matter of metaphysical necessity, no less, in respect of whether they possess CP. At the very least, their discriminatory abilities are more similar to each other than either is, say, to the discriminatory abilities of some non-human species. CP for colour would not form part of an account of the abilities that the organisms in the cluster have in common. By focusing on the discriminatory abilities of the cluster rather than the *specific* abilities possessed by CAT, then, one could still argue that CP is not part of the nature of the colour vision abilities possessed by humans *as assessed across a range of closely related counterfactual cases*.

A third and more forceful response to this objection, towards which I am inclined, is to deny that the differential similarity effects characteristic of colour CP are purely a function

of discrimination performance. If this response is correct, then the scenario involving NONCAT is not just metaphysically but *nomologically* possible. In support of this view, Hardin (1997: 293) argues as follows,

> It is important to understand that the resemblance that connects two instances of the same colour category is not necessarily a function of the perceptual [i.e. discriminable] distance between them. It is not hard to find three colour samples A, B, and C, which are such that B is separated from A on the one side and C on the other by the same number of just-noticeable differences, and yet A and B are seen to belong to the same colour category whereas C is seen to belong to a different colour category.

The same view is expressed in the recent empirical literature. For example, Franklin and Davies (2004: 352, note 2) dissociate CP from discrimination for the following reason,

> [Inequality in discriminable distance] would not really be equivalent to adult categorical perception. Equivalent non-uniformities also occur within categories. Thus, quasi-categorical effects should also be found within categories. It might be thought that it is this variation in discriminability that drives category formation. Category boundaries occur at local discrimination maxima. This fits the blue-green boundary; there is a discrimination peak at about the right place. But, it is much less clear for green-yellow and yellow-orange-red.[154]

Franklin and Davies argue that the differential similarity effects of CP are not a function of discrimination because discrimination performance does not peak at the category boundaries between *green* and *yellow*, *yellow* and *orange*, and *orange* and *red*. Moreover they argue that if differences in discrimination were underlying CP, then we should expect visually to experience more colour categories than we in fact do, given the existence of local discrimination maxima within categories. Further evidence for this conclusion has been

[154] As it happens, Roberson et al. (2009) argue that discrimination performance also cannot explain the blue-green category boundary. See also Hanley & Roberson (2011).

presented by Roberson and colleagues (2009). Their study focuses on speakers of Korean, who display characteristic CP effects for three categories in the blue-green region, *chorok*, *cheongnok* and *parang*, whereas English speakers display these effects for only two: *blue* and *green*. Despite these robust differences in CP for colour, however, they found that discrimination thresholds across the spectrum were not significantly different for speakers of English and Korean. These studies present strong evidence against the claim that CP for colour is a pure discrimination effect and thus in favour of the view that the scenario involving CAT and NONCAT presents a genuine and indeed nomological possibility.

If this last point is correct, why do so many people persist in characterising CP as a discrimination effect? This issue is especially puzzling given that even those who wish to dissociate discrimination and CP sometimes default to this characterisation. For example, in the very same paper quoted above, Franklin and Davies (2004: 350) rather confusingly adopt the standard line that 'Categorical Perception is shown by better between-category than within-category discrimination'. By way of diagnosis, I believe that this apparent inconsistency is due to a failure clearly to distinguish between notions of visual discriminability and similarity. In studies of CP, the notion of 'discrimination' is often used where 'similarity' (or relevant cognates) would be more appropriate.[155] To illustrate this point, consider the study of CP reported in Winawer and colleagues (2007). Their experiments involve an 'ABX test' paradigm in which subjects are presented with a target square together with two test squares (see figure seven above), one of which is identical to the target while the other 'distracter' square is drawn either from the same linguistic category *F* as the target (the 'within-category condition') or a different linguistic category *G* ('across-

[155] Note that *discrimination* and *similarity* are non-equivalent notions. Similarity is a relation of *resemblance*: *x* is more similar to *y* than to *z* (in respect *R*) iff *x* resembles *y* more than *z* (in respect *R*). In contrast, as noted in chapter one, discrimination is best viewed as an *epistemic* or *cognitive* notion: to say that *S* discriminates *x* from *y* (in respect *R*) is to say that *S* is able to acquire knowledge, or to token reliably accurate representations, with the content that *x* is distinct from *y* (in respect *R*).

category condition'). Subjects are asked to report which of the two test squares is the 'same' as the target. The assumed standard for possessing CP for *F* is that the subject will be faster at identifying the identical square in the *across*-category condition than the *within*-category condition. Such subjects are described as demonstrating 'faster discrimination' between the two test fields in the across-category condition than in the within-category condition. It simply does not follow from the set-up or the description of the experiment, however, that these subjects have '*better discrimination*' for colour in the sense at issue in this thesis. Nor does it follow that these subjects have '*better discrimination*' in the sense that the across-category fields are separated by more just noticeable differences (JNDs) than the within-category fields for subjects with CP. On the contrary, the result is entirely compatible with the test fields being separated by the same number of JNDs for all subjects. The difference in so-called 'discrimination' by subjects with CP might simply be due to the test fields *appearing more dissimilar* in the across-category condition than in the within-category condition. Given the above evidence dissociating discriminability from CP, presumably this is how the experiment *should* be interpreted in order to resolve the apparent inconsistency between these apparently conflicting characterisations of CP.

Moving on, perhaps the strongest objection to the third premise derives from the widespread attribution of apparently categorical processes in colour vision science. The objection can be sketched as follows. Colour vision science has progressed some considerable way towards elucidating the nature of our ability to perceive colour. Moreover, the physiological and computational processes and mechanisms described within colour vision science often are labelled with the categorical terms 'red', 'green', 'blue', and so on.

This practice strongly suggests that the ability to perceive colour categories is part of the nature of colour vision.[156]

In response, the history of colour vision science illustrates that researchers typically have been too quick to label physiological processes or mechanisms with categorical colour terms. A prominent early example is the *trichromatic theory* of colour vision presented by Young and Helmholtz, which proposed that all colour experiences were the product of the combined output of three retinal photoreceptor types, which were labelled as 'red', 'green', and 'blue', receptors. Now although the trichromatic theory correctly predicted the *number* of (chromatic) photoreceptor types in the human retina, it is now widely acknowledged that it was wrong to label these 'red', 'green', and 'blue' receptors, and that we are justified only in labelling them 'long', 'medium', and 'short' (L, M, S) wavelength cones. As Gordon and Abramov (2001: 103) note, 'an L-cone is not a red-cone; trivially, it responds to some degree to all wavelengths, and the wavelength at its spectral peak (560 nm) usually appears yellowish-green.' In a scolding passage, De Valois and De Valois (1975: 107) add that 'the spectral sensitivities of the L and the M cones are very nearly the same, peaking as they do at about 530-540 and 560-570 nm (both in the part of the spectrum we see as greenish yellow, compounding the absurdity of calling them "green" and "red" receptors as too many people do)'.

Similar over exuberance can be seen in attempts to provide physiological reductions of *opponent process theory* (OPT). OPT is founded on a set of observations about the phenomenal structure of ordinary human colour experience. There appear (phenomenally) to be four *unique hues*, that is, hues which do not appear to be a mixture of any other hue. These four hues are commonly labelled *unique red*, *unique green*, *unique yellow*, and *unique blue*.

[156] The objection to the constitutive account of colour vision posed by Thompson (1995b: 196-197) in section 1.4.6 is also relevant here.

All other shades of colour phenomenally appear to be *binary hues*, that is, hues which appear to be mixtures of other hues. Shades of *orange*, for example, appear to be mixtures of *red* and *yellow*, and shades of *purple* appear to be mixtures of *red* and *blue*. No shade ever appears to be a mixture of *red* and *green*, however, and likewise for *yellow* and *blue*. These pairs are therefore referred to as *phenomenally opponent colours*.[157] OPT provides strong reason to suspect that our ordinary conscious colour vision is subserved by opponent processes of some sort. Upon discovery of *physiologically opponent processes* in animal colour vision systems in the 1950s, researchers predictably were extremely quick to apply categorical colour terms to these processes.[158] This practice remains widespread across scholarly research articles and textbook introductions to colour vision. As Mollon (1997: 868-872) explains, however, these applications are once again premature,

> We still believe today that there are chromatically antagonistic channels in the early visual system—that is, channels that draw inputs of opposite sign from different classes of cone, but these channels simply do not correspond to the phenomenologically defined channels of Hering… In fact, no one has found a site in the visual system where colour appears to be represented according to Opponent Colour theory—that is, a site where the cells might be held to secrete redness and greenness or yellowness and blueness. Cells are found in the cortex that respond to restricted regions of chromaticity space, but they are by no means confined to the loci of pure hues. Thirty years ago we thought we understood the existence of four unique hues, hues that are phenomenally unmixed. Today this is perhaps the major unsolved problem of colour vision.[159]

[157] See Hardin (1988a: 26, 30, 37-39, 52). The *locus classicus* for opponent colour theory is Hering (1920/1964). Note that the theory describes how hues appear *phenomenally*: it is not necessary that binary hues, for instance, appear to have been produced via some *physical* mixture of lights or paints. Byrne & Hilbert (2008) and Byrne (2003: 642, fn. 7) provide further clarity on these notions.

[158] An example of a 'physiologically opponent process' is a cell that is excited by the activity of short wavelength cones, and inhibited by the activity of, say, medium wavelength cones. Such processes were first discovered in the retinal ganglion cells of fish by MacNichol & Svaetichin (1958), and in the lateral geniculate nuclei of macaques by De Valois et al. (1967). For a summary, see Gordon & Abramov (2001).

[159] See Gordon & Abramov (2001: 105-108) for more detailed discussion of these problems.

These apparently categorical scientific descriptions of the physiological mechanisms subserving colour vision therefore do nothing to undermine the view that colour vision does not constitutively involve the ability for categorical perception of colour.

A closely related objection is that, although no *physiological* reductions of OPT have been successful to date, the original *phenomenal* form of OPT itself indicates that human colour vision is essentially categorical in nature. Phenomenal OPT states that all human colour experience can be characterised by the ability to represent colours as shades of *red*, *green*, *yellow*, and *blue*, and as complexes of these categories, such as *reddish-blue*, *greenish-yellow*, and so on, subject to the constraints of the opponent hue pairings.[160] Many people take this account of the structure of colour experience to impose unassailable constraints on the nature of colour vision. Thompson (1995a: 9) thus proclaims that 'colour vision, whatever else it might be, is the ability to see visual qualities belonging to the phenomenal hue categories, red, green, yellow, blue'. If Thompson is correct, however, it surely follows that colour vision constitutively involves CP for colour, at least with respect to the categories *red*, *green*, *yellow*, and *blue*.

I offer two points in response. First, as noted in the introduction to this thesis and reiterated throughout, the constitutive approach pursued here is intended as an alternative to the standard experiential approach to understanding colour vision. It is simply missing the point, therefore, to insist that structural features apparent in our ordinary conscious visual experience of colour provide unassailable constraints on the nature of colour vision. One might insist that in rejecting such experiential constraints I have simply *changed the terms of the debate* or *changed the subject* with respect to others working in the philosophy colour such as Thompson. I completely reject these charges. All parties to the debate acknowledge,

[160] The 'hue scaling' techniques of Jameson & Hurvich (1959) and Sternheim & Boynton (1966) show that the terms 'red', 'green', 'yellow', and 'blue', and combinations of these terms, are necessary and sufficient to describe all of the spectral hues.

albeit some more explicitly than others, the foundational explanatory role of colour vision in the philosophy of colour. Far from changing the subject, I am offering an account of the nature of *that very psychological ability*. The fact that my account differs from the philosophical mainstream in both methodology and content does not imply that we are engaged with different subject matters. One is obviously free to object to any or all aspects of my account. What is not acceptable, however, is to impose an experiential or phenomenological conception of colour vision as a precondition on the debate.

Second, even if one remains unconvinced by the last paragraph, it is in any event mistaken to think that phenomenal OPT implies that human colour vision necessarily or constitutively involves CP. As we have seen, the four phenomenally unique hues postulated by OPT are typically labelled 'unique red', 'unique green', 'unique blue', and 'unique yellow'. This might suggest a straightforward relationship between perceiving the unique hues and seeing the colour categories *red*, *green*, *blue*, and *yellow*. This thought is bolstered by evidence suggesting that subjects possessing CP for these categories locate their *foci* or *paradigmatic examples* at or very close to the corresponding unique hues.[161] Despite the apparent correlations, however, there is no straightforward connection between the unique hues and CP for colour. Actual empirical cases suffice to make the point. The aforementioned study by Winawer and colleagues (2007) demonstrated that Russian subjects possess CP for the categories *light blue* and *dark blue*, which are expressed in Russian as 'siniy' and 'goluboy'. Now although Russian subjects presumably perceive a unique hue in the bluish region, this hue could not possibly form the focal point for the category *blue* for these subjects. Roberson and colleagues (2009) likewise found that Korean subjects possess CP for the three categories, *chorok*, *cheongnok* and *parang*, in the blue-green region. Korean subjects presumably perceive so-called 'unique blue' and 'unique green', but these hues

[161] McDaniel (1972) first linked Berlin and Kay focal colours to unique hues. See Wooten & Miller (1997: 83). MacLaurey (1997: 263) discusses some problems with this identification.

could not form focal points for the categories *blue* or *green* for these subjects. There is no limit on the range of possible variations of this sort. In particular, it is easy to imagine situations in which subjects accurately described by phenomenal OPT have CP for fewer, rather than more, colour categories than ordinary English speakers. An intermediate case would be a subject who possessed CP for a category that included both 'unique green' and 'unique blue' hues. This colour category might be 'multi-focal', in that two contrary hues would be classed as joint-best instances of the category. A more extreme case would be a subject like NONCAT who lacked CP for any colour categories whatsoever, and yet still experienced the four unique hues postulated by phenomenal OPT. Variations of this sort fatally undermine any putative constitutive connection between phenomenal OPT and CP for colour. By the same token, these cases demonstrate just how misleading it is to label the four unique hues postulated by phenomenal OPT with the category terms 'red', 'green', 'yellow', and 'blue'.

As will be clear to those with some knowledge of this area, the evidence adduced in the last paragraph undermines the extremely pervasive view that the colour categories apparent in our ordinary visual experience are *universal* features of colour vision. So-called *Universalism* has been the prevailing orthodoxy in psychological theories of colour categorisation at least since the seminal cross-cultural study by Berlin and Kay (1969). Universalism holds that the colour categories apparent in our visual experience of colour are the result of species-wide, largely innate, probably physiological, features of colour vision systems.[162] Key support for the Universalist position has been drawn from the following three claims,

[162] The debate has focused on so-called 'basic' colour categories, for which many behavioural and linguistic criteria have been proposed. See Berlin & Kay (1969) for details.

1. There appear to be a limited number of basic colour categories, including *red*, *yellow*, *green*, *blue*, and so on, which are expressed across many of the world's languages. This suggests the existence of certain underlying species-wide perceptual constraints on our colour categories.

2. Where cultures share terms for certain colour categories there is considerable agreement on the 'best' examples of these categories, suggesting that there exist certain shared phenomenally salient 'focal points' for these colour categories.

3. Across cultures, subjects find it easier to learn and remember names for the recognised focal colours of the categories *red*, *blue*, *green*, and so on.

Despite their obvious interest I do not intend to discuss these claims in detail. The key point for present purposes is that the Universalist view undoubtedly has permeated philosophy, perhaps partly explaining the widespread assumption that the ability to see colour categories such as *red* and *green* is somehow inherent to, or part of the nature of, colour vision.[163]

One of the most notable trends in recent cognitive psychology, however, has been a renewed interest in linguistic relativist or *Whorfian* views of perceptual categorisation. In its strongest form, the Whorfian view holds that all of our perceptual categories are determined (in some sense) without constraint by the language that we speak.[164] In the words of Whorf himself (1940: 213),

> The categories and types that we isolate from the world of phenomena we do not find there because they stare every observer in the face. On the contrary the world is

[163] Arguably the most widely read and influential philosophical introduction on the science of colour vision, Hardin's *Colour for Philosophers*, presents the standard Universalist line at (1988a: 156-169).
[164] To my knowledge, the mechanism by which this 'determination' is supposed to take place was never made clear. This remains an open issue in contemporary versions of the Whorfian thesis.

> presented in a kaleidoscopic flux of impressions which have to be organised in our minds. This means, largely, by the linguistic system in our minds.[165]

Nobody today would endorse such an extreme empiricist view. Indeed cognitive science and the associated philosophy of mind has flourished in the last fifty years largely as a result of highly rationalist and nativist programs of explanation with respect to our linguistic, perceptual, and more broadly cognitive capacities. Nonetheless, cross-cultural studies of the sort described above have provided compelling evidence for a weaker, *neo-Whorfian*, thesis,[166]

> (WH) Some structural differences between colour language systems will be paralleled by differences in the colour categories apparent or represented in the visual experience of the native speakers of these languages.

As an autobiographical note, early reflection on neo-Whorfian views in perceptual psychology seeded many of the ideas in this thesis. The view provided an intriguing alternative to the Universalist view of the psychological origin of colour CP and seemed highly conducive to the view that the abilities reflected in categorical perception of colour are not constitutively associated with colour vision. The preceding discussion of phenomenal OPT directly illustrates the relevance of neo-Whorfian views to the present argument. In the next section, I shall introduce some fascinating cognitive neuropsychological cases which have been cited as further support for the neo-Whorfian view. Intriguingly these cases also establish that the possibility of a complete dissociation between colour vision and CP, on which the foregoing argument rests, is not nearly as distant as one initially might think.

[165] Compare strong empiricist views such as Quine (1948: 11) who says that 'physical objects are postulated entities which round out and simplify our account of the flux of experience'.
[166] For a summary, see Kay & Regier (2006).

3.4. The View from Cognitive Neuropsychology

The neo-Whorfian view of colour categorisation is motivated not only by comparative, cross-cultural, studies but also by cases in cognitive neuropsychology. The relevant cases involve subjects with language-specific deficits who retain apparently normal visual discriminatory abilities but perform extremely poorly on *explicit* tests for CP for colour. As noted above, explicit tests for CP require the comprehension or production of a colour category name in response to a stimulus, selecting the focal instance of a given category, or the active sorting of colour chips into categories or groups. Poor performance on such tasks does not establish the absence of CP by the *implicit* standard of differential similarity effects at colour category boundaries that has been assumed here. Nonetheless their severe impairment on explicit tests for CP suggests that such subjects come extremely close to the complete dissociation between constitutive colour vision abilities and CP envisaged in the case of NONCAT.

Cognitive neuropsychology has identified a condition known as *visual colour anomia* or *optic aphasia for colours*. Aphasias are deficits in linguistic ability most typically resulting from a head injury or stroke. Anomia is a type of aphasia involving a specific impairment in the production of names or verbal labels. Optic aphasia for colours is characterised by a specific impairment in the naming of colours. Unlike *colour aphasia*, which is a purely verbal deficit, optic aphasia for colours is a *visuo-verbal* deficit resulting in fascinating (if extremely unfortunate) impairments in a subject's ability visually to categorise colours. The earliest reports of such cases appear in work by the German neurologists Goldstein and Gelb (1918), who described it as an 'impairment of the categorical attitude' towards colour, Sittig (1921), and Lange (1936: 88), who says that,

> The patients do no behave conceptually but in terms of concrete reality... The patients have no principle of classification in colour sorting, but depend on the experience of coherence between patterns and the other colours.[167]

Some decades later, important work by Beauvois and Saillant (1985) served to distinguish optic aphasia for colours from the more general linguistic deficit colour aphasia and also from forms of visual agnosia which arguably involve an impairment in the ability to perceive colour.[168]

Beauvois and Saillant (1985) describe the case of a patient MP who they diagnose as having optic aphasia for colours. MP's visual and linguistic abilities were assessed via purely verbal, purely visual, and combined visuo-verbal tasks. In the verbal tests, both the task stimulus and response and the intervening process or strategy used to complete the tasks were verbal. For example, a *colour name sorting test* involved questions such as, 'which category does *scarlet* belong to?', with the subject asked to respond verbally from a choice of 'brown', 'red', or 'yellow'. MP performed almost perfectly on such tasks. In the visual tests, the stimulus and response and the intervening strategy were visual. For example, in a *colour matching test* subjects were given two wool skeins and asked to decide if they were identical or not. MP was also tested on standard pseudo-isochromatic plates for colour discrimination. Once again MP's performance on such tasks was well within the normal range. In the visuo-verbal tests, the stimulus and response type differed in that one was verbal and the other visual or *vice versa*. For example, subjects were asked to produce a verbal colour label in response to a visually presented colour sample, or to point to a colour sample given a verbal request. These are clear examples of explicit tests for CP. MP's performance was significantly impaired on such tasks. MP's errors were also very difficult to understand: she

[167] As quoted in Roberson et al. (1999: 10).
[168] The relationship between optic aphasia for colour and colour agnosia is an important and complex issue that I aim to address in future work.

labelled a bright red sample as ‘light yellow’, pointed to navy blue when asked to point to orange, and pointed to pale blue when asked to point to the colour of parsley.

MP’s performance was also significantly impaired on the Gelb-Goldstein colour sorting test, which requires the subject to ‘select a skein of wool he likes and to pick out all the skeins which he believes can be grouped together with the chosen one’ (1985: 15). As Beauvois and Saillant (1985: 15) report,

> [A]lthough she could explain what the instructions meant, she behaved in an extremely strange way: she started with a pale grey and put it with the other grey, the white and the dark wools; then with them she put, in the following order: A pale pink, a pale yellow, a brown; she continued by adding all the pale pink and brown wools, and she finished by putting all the wools in the same category. When then asked to repeat the instructions, she correctly did so. On a second attempt she behaved in virtually the same way, the only difference being concerned with the order of the choice of the colours. On part 4 of the test, in which she was asked to select just the red wools, she picked up all the wools except for 10 (white, black, grey, one brown, four grey and two green); she seemed to select them without any apparent logic in order (e.g., she picked out a green wool just after a pink one, or a red just after a green).

The Gelb-Goldstein task is another paradigmatic explicit test for colour CP. MP was almost completely unable to complete this task, suggesting that although her visual discriminatory abilities were intact her ability consciously to perceive colour categories was damaged.

Roberson and colleagues (1999) describe the similar case of the patient LEW, who also meets the criteria for optic aphasia for colours. Like MP, LEW retained normal visual discriminatory abilities but performed extremely badly on explicit tests for CP. For example, when given an exemplary sample of a colour category such as *red* and asked to provide its name, LEW would invariably provide the wrong label. His choice was often radically wrong, as in calling a red shade ‘blue’. He was also impaired in his ability to point to an instance of a

colour given its name. LEW was also unable to pick out the 'best' or 'prototypical' example of a given colour category, his selections always falling outside the normal focal area for that category.

As with the patient MP, the most interesting aspect of LEW's behaviour came in his performance on a 'freesorting' categorisation task. The materials for the task included two groups of colour chips. The larger set included 58 chips spanning across the spectrum gradually from *red* to *purple*, through *pink*, *yellow*, *brown*, *green* and *blue*, with each chip abutting another in colour space. The second set included only 20 chips, with each chip considered by normal subjects to be a focal instance of the categories *red*, *yellow*, *pink*, *blue*, or *green*. The chips from one set were spread out on a table, and subjects asked to sort them into categories or groups in 'whichever way they saw fit'. As one would predict, normal subjects typically sort the chips into around eight groupings, corresponding to the basic colour categories *red*, *pink*, *yellow*, *brown*, *khaki*, *green*, *blue*, and *purple*. In general they perform the task with great ease and facility. In stark contrast, on one trial LEW split the large set into just four groupings, with one containing chips from *light purple*, *pink*, *red*, and *yellow*; the second containing some *red*, *brown*, *khaki* and *green*; the third containing *green* and *khaki*; and the fourth containing *blue*, *purple* and *red* chips. LEW's performance varied across trials, but he always generated highly unusual mixed groupings. LEW's strategy in achieving these groupings was of particular interest (1999: 9),

> [The] manner in which [LEW] proceeded was to pick up two stimuli and compare them to each other. LEW looked for two stimuli that were the most perceptually similar. If satisfied that they met his criteria for grouping he placed them together, later using one of them to carry out the same procedure with another stimulus.[169]

[169] In order to assess whether LEW's performance resulted from confusion, or perhaps overloading, given so many category boundaries in the large set of colour chips, LEW was also tested on smaller sets of colour chips ranging across only a single category boundary. However, even for chips spanning just the green/blue boundary,

LEW's unusual behaviour suggests that he was sorting or categorising the colour chips via the arduous method of pairwise similarity comparisons, rather than being guided by visually salient colour categories.[170] Despite his extreme difficulty with this *explicit* categorisation task, however, LEW's performance on *implicit* categorisation tasks, including the canonical test for differential similarity judgements across category boundaries, was within the normal range.[171] By the lights of the standard for CP assumed in this chapter, then, LEW classes as possessing CP for colour.

As Roberson and colleagues (1999) point out, the most plausible causal explanation of the highly unusual categorising behaviour in such cases is the language-specific deficit or aphasia. Roberson and colleagues (1999) thus present such cases as evidence for the neo-Whorfian view of colour categorisation. The success or otherwise of such empirical arguments is not my concern. The relevance to the present argument is that these cases serve as a substantial plausibility consideration in support of the crucial possibility involving CAT, NONCAT, and total NCP for colour. Although MP and LEW meet the official standard for CP assumed in this chapter, their radically impaired performance on explicit categorisation tasks suggests that they lack very closely related abilities. These subjects thus come extremely close to the sort of dissociation between our constitutive colour vision abilities and CP envisaged in the case of NONCAT. It is certainly no stretch to imagine a subject just like

---

for instance, LEW completed the task in the same way, assessing chips for pairwise similarity. The result was that LEW placed all of the chips into a single group.

[170] One way to describe this behaviour would be as a type of *sorites reasoning*. LEW seems to reason that for any particular pair of colour chips, if they fall below some threshold in perceptual similarity, then they should be assigned to the same colour category. This has the result that LEW's groups contain chips from clearly different and even opposing colour categories such as *red* and *green*. In much the same way the sorites paradox for 'bald', for instance, involves the repeated application of a *tolerance principle* such that two people differing in respect of only one hair must either both class as bald or both as not bald, with the result that everyone in the universe ends up being classed as bald (or as not bald). The relevance of this case for our understanding of the sorites paradox and vagueness will form the basis for future work.

[171] It is not unusual for patients with visual deficits as measured by performance on explicit tasks to perform relatively normally on implicit tasks. Nijboer (2007: 144-145) notes related findings in blindsight, object agnosia, and prosopagnosia.

MP or LEW also failing implicit tests for CP, which would suffice to establish the relevant possibility. Despite the apparent strangeness, then, the scenario involving NONCAT is not as distant as one initially might think.

3.5. Conclusions

In this chapter I have implemented the constitutive approach to understanding colour vision to argue that categorical perception of colour, a highly distinctive and vivid feature of our ordinary conscious visual experience of colour, is not part of the constitutive nature of colour vision. In the provocative phrase with which I started the chapter, colour vision does not by its nature confer the ability to *see colours*, in the natural categorically committed sense. The argument centred on the counterfactual case of NONCAT, a subject exactly like a normal human in respect of the proposed discriminatory standard for colour vision and yet completely lacking CP for colour. The plausibility of this case was further motivated by some fascinating cognitive neuropsychological evidence for actual dissociations between colour discrimination abilities and certain aspects of categorical perception for colour. As is often the case, the unexpected actual extends our knowledge as to what is possible. Such cases illustrate that our ordinary experience of colour possesses a latent complexity that is obscured by ordinary introspective approaches to understanding colour vision. Such complexity is one of the chief motivations for the constitutive approach, which seeks to isolate and refine our notion of the psychological kind *colour vision*, so that we might identify even the strangest instances of that perceptual ability.

One issue of outstanding importance in this chapter concerns the consequences of this argument for colour ontology. As noted in the introduction to this thesis, the mainstream experiential view of colour vision implies that our colour ontology ultimately must include properties such as *red*, *yellow*, and *green*. If the argument of this chapter is sound, then the

ability visually to represent such categories turns out not to be part of the constitutive nature of our colour vision abilities. The constitutive approach thus imposes no constraint on colour ontology to accommodate determinable colour categories such as *red* and *green*. Somewhat paradoxically, the constitutive approach thus implies that our ontology of colours will be free of the properties that we ordinarily refer to simply as *colours*. A closely related view is held by Smart (1975: 7) who claims that 'colours are the (perhaps highly disjunctive and idiosyncratic) properties of the surfaces of objects that explain the discriminations with respect to colour of normal human percipients, and also the experiences of these percipients'. Discussing Smart's view, Averill (1985: 15) notes that 'if colour is a matter of explaining the sorts of discriminations Smart has in mind, then it would seem that, strictly speaking, greenness is not a colour, but shades of green are colours'. Unfortunately limitations of space dictate that I cannot explore the consequences of this particular view of colour ontology, which therefore will have to be taken as an interesting corollary but not a central part of the present view.

4. Colour Similarities at Different Levels

As noted in chapter one, the constitutive account of colour vision as colour constancy *qua* reflectance discrimination naturally supports the view of colour ontology known as *reflectance physicalism*. So far this ontological thesis has remained very much in the background, although it has clearly surfaced in a number of places – most notably in discussing the colour appearance claim in chapter two and in concluding the discussion of categorical perception in the previous chapter. In this chapter I want to tackle the issue of colour ontology head-on in considering the *argument from colour similarity*, which many people take to pose the most serious threat to reflectance physicalism. I shall first argue that the putative dissociation between CP for colour and our constitutive colour vision abilities provides strong reason to think that phenomenally apparent similarities among the colours should not impose strong constraints on candidate theories of colour ontology. I shall then present two additional new problems for the argument from colour similarity, which are independent of the issues raised so far in this thesis. Both problems are inspired by Quine's observation that different systems of theoretical description and classification may have different associated standards for similarity. The first problem is that the argument ignores the potential for *contextual variation* in the similarity relation. It seems likely that the shift in the argument from an ordinary visual context to a reductive scientific context makes salient different respects of comparison between the colours and thus different notions of similarity. Although I doubt that this point is ultimately conclusive, it places the onus firmly on my opponent to clarify the operative notion of similarity in the argument, something that turns out to be much harder than one might think. This issue should encourage all parties to be more precise and critical in thinking about the notions of similarity operating in discussions

of colour and colour ontology. The second problem concerns the epistemic analogue of context sensitivity, namely the *presentation sensitivity of knowledge* of similarity. The argument from colour similarity depends crucially on the mistaken assumption that, if colours are identical with surface reflectance properties, then similarities between the colours must be *evident* from canonical physical descriptions of these properties such as their spectral reflectance curves. I present a range of analogous cases from chemistry and biochemistry in which this epistemic assumption fails. We shall see that the only way to save this assumption is to admit *interest relative* descriptions of properties as canonical physical presentations. This significantly reduces the threat posed by the argument, for the reflectance physicalist now can appeal to *biological* descriptions of reflectance properties, for example, which are better suited to the predictive and explanatory interests of vision science than the descriptions provided by optical science.

4.1. The Argument from Colour Similarity

Many people take the greatest threat to reflectance physicalism to be the *argument from colour similarity*.[172] The argument is part of the broader *argument from colour structure*, which was introduced to the contemporary literature by Hardin (1988a).[173] This broader argument rests on three commonly cited claims about our visual experience of colour, which for now I simply take at face value. The first is that some colours appear (phenomenally) to be *unique*, in the sense that they do not appear to be 'mixtures' of any other colours, whereas other colours appear to be *binary*, in that they appear to be 'mixtures' of other colours. The

[172] This threat is stressed by McLaughlin (2003: 111), Byrne & Hilbert (2003: 7), and Pautz (2006: 536). Most commentators agree that colour primitivism avoids the argument from colour similarity. See Pautz (2006: 536, 562) and Byrne (2003: 648) who claims that if the argument is successful then primitivism is 'the only realist theory of colour left standing'.

[173] The argument is also presented and endorsed by Maund (1995: 42, 141), Thompson (1995b: 128-130, 135-139), Clark (1996: 145-146), and Pautz (2006). Boghossian & Velleman (1991) and Johnston (1992) offer epistemological versions of the argument. Wittgenstein's (1977) notes on Goethe's *Theory of Colours* contain detailed discussion of various structural features of colour, although as one might expect there is no specific engagement with the philosophical tradition on primary and secondary qualities. For discussion, see McGinn (1991).

second claim is that some colours bear *opponent relations* to one another, in the sense that they appear (phenomenally) to 'cancel each other out', and cannot be 'mixed' with one another to form a binary hue. These claims were discussed briefly towards the end of the previous chapter. The third claim is that colours appear to bear various *similarity relations* to one another. The basic thrust of the argument is that none of these structural properties is shared by reflectances, which therefore cannot be identical with the colours.

By way of clarification, similarity relations are typically predicated of determinable colour categories, as in the claim that *red* appears more similar to *orange* than it does to *green*. At the end of the previous chapter I made it clear that a consequence of my view regarding colour vision and CP for colour is that strictly speaking determinable colour categories such as *red* and *green* are not colours. In what follows, then, I shall use subscripted colour terms such as '$blue_1$' and '$green_3$' to refer to determinate shades of colour and will focus on similarity claims such as that $blue_1$ appears more similar to $purple_2$ than $green_3$. It will be important to bear in mind that on my view perceiving the colour shade $blue_1$, for example, does not entail that one possesses categorical perception for the determinable colour category *blue*. The use of subscripted colour category terms is thus merely a *façon de parler* in order to secure reference to determinate shades of colour.

Cohen (2003: 81) summarises the argument from colour similarity as follows,

> "The argument from structure", is a simple instance of Leibniz's law: colours have these structural properties, and therefore any set of properties that is identical with the colours must have them as well. However, it is suggested, the only plausible candidates proposed by [reflectance physicalists] for being identical with the colours lack the structural properties in question, and therefore the purported identities cannot be sustained.

Here I shall focus exclusively on the similarity component of this argument, which can be unpacked as follows. Let's assume that reflectance physicalism holds that $blue_1$ is (type) identical with surface spectral reflectance type R_b, $purple_2$ is (type) identical with surface spectral reflectance type R_p, and $green_3$ is (type) identical with surface spectral reflectance type R_g,[174]

(PHYS) $\text{Blue}_1 = R_b$, $\text{purple}_2 = R_p$, $\text{green}_3 = R_g$.

From our visual experience of colour, it phenomenally appears as though the colours $blue_1$, $purple_2$, and $green_3$, stand in the following similarity relation,

(SIM) Blue_1 is more similar to purple_2 than green_3.

By Leibniz's Law, if PHYS is true, then SIM is true if and only if the corresponding claim about reflectances is true,

(LL) If $\text{blue}_1 = R_b$, $\text{purple}_2 = R_p$, $\text{green}_3 = R_g$, then blue_1 is more similar to purple_2 than green_3 iff R_b is more similar to R_p than R_g.[175]

It seems, however, that the reflectances R_b, R_p, and R_g do not stand in the requisite similarity relation,

?? (SURF) R_b is more similar to R_p than R_g.

What basis do we have for rejecting SURF? The primary basis is that the relevant similarity is not evident from the *canonical physical descriptions* or representations of these reflectances provided by optical science. The thought is that, if SURF is true, we should be

[174] Throughout this chapter it should be taken as read that colours ultimately might be identified with *productances* rather than reflectances. See the discussion in section 1.4.1 in chapter one. As Byrne & Hilbert (2003: 10-11) point out, the phenomenon of *metamerism* implies that determinate shades of colour such as $blue_1$ will be identical with reflectance/productance types rather than super determinate reflectances/productances. I would actually be inclined to identify colours with super determinate reflectances/productances but for the sake of argument I shall adopt the assumption that colours are reflectance/productance types.

[175] Strictly speaking the instance of Leibniz's Law will contain a single identity claim in the antecedent, but the argument could easily be repeated with other candidate identity claims, thus ensuring the requisite generality.

able to acquire knowledge of, or simply 'read off', the relevant similarity by casually reflecting on canonical scientific representations of the reflectances R_b, R_p, and R_g, such as their *reflectance spectra* or *spectral reflectance curves*, which are graphs of reflectance plotted against wavelength. Pautz (2006: 539), for example, says that 'the case for [rejecting SURF] is that it is obvious when we look at the reflectance curves for reflectances' R_b, R_p, and R_g that they do not stand in the requisite similarity relation, which we can see for ourselves in figure eight below. The argument proceeds by *modus tollens* on LL to derive the negation of PHYS, the reflectance physicalist thesis.

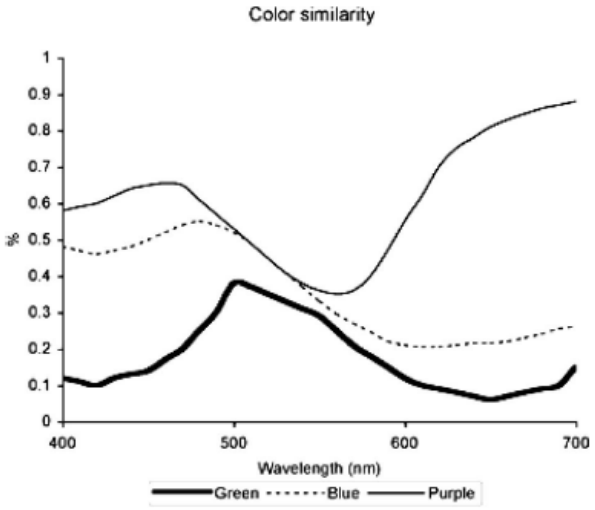


Fig. 8: Reproduced (with permission) from Byrne & Hilbert (2003: 13)

4.2. Standard Reflectance Physicalist Responses

What responses are available to the reflectance physicalist? The instance of Leibniz's Law is surely unassailable. One option would be to defend SURF, but reflectance physicalists have been reluctant to do this. Byrne and Hilbert (2003: 13), for example, agree with Pautz that 'there does not seem to be an obvious respect in which the ['blue'] reflectance-type is more similar to the ['purple' type] than it is to the ['green' type]'. Byrne (2003: 648) goes further in saying that 'we can be completely confident that any plausible physicalist candidates for the colours do *not* stand in the required genuine similarity relations'.[176] That leaves SIM, the

[176] Jackson (1998: 111) concurs. I discuss the notion of *genuine* similarity in detail below.

colour similarity claim. Given its apparent intuitiveness, however, no party to the debate seems to want to reject this claim. The argument from colour similarity thus leaves the reflectance physicalist little room for manoeuvre.

In broad terms, most reflectance physicalists have responded to the argument by reinterpreting the data about colour similarity as pertaining to features of *us* rather than features of the colours. The popular *experiential response* claims that SIM is elliptical for a claim about similarities between our *visual experiences* of $blue_1$, $purple_2$, and $green_3$ – more precisely, the *phenomenal characters* of these experiences.[177] Taking the value of the function E(*x*) be the type of phenomenal character of experiences typically caused by *x*, SIM is thus interpreted as follows:

(SIM-E) E($blue_1$) is more similar to E($purple_2$) than E($green_3$).

McLaughlin (2003: 115) thus notes that on this view the 'colours themselves participate in the similarity and difference relationships derivatively – in virtue of the participation of the visual experiences that they dispose their bearers to produce'.[178]

Other physicalists interpret the colour similarity data as pertaining to our distinctive mode of *visually representing* colours. An example is Byrne and Hilbert's (2003) 'hue magnitudes' approach.[179] Very roughly, the picture is that the colour vision system represents objects as having proportions of the opponent hues, which Byrne and Hilbert label 'red' (**R**)

[177] Shoemaker (1991), Lewis (1997: 339-340), McLaughlin (2003), Cohen (2003, 2009).
[178] Also see Cohen (2003: 88). As Pautz (2006: 554) usefully notes, this is in effect a dispositionalist or response-dependent account of colour structure, coupled with a physicalist or response-independent account of first-order colour properties.
[179] The commentaries and responses in Byrne & Hilbert (2003) and Pautz (unpublished a, b) present various criticisms and clarifications of the account. A precursor to the view appears in Hilbert (1987: 117-118), who suggests that triples of integrated reflectances might preserve similarity relations among the colours. Hilbert has since renounced this view: see Byrne & Hilbert (1997c: 285) and discussion in Cohen (2003: 81).

and ‘green’ (**G**), and ‘yellow’ (**Y**) and ‘blue’ (**B**), which allows colour similarities such as SIM to be explained as follows (2003: 15),[180]

> Objects that appear blue are represented as having a high proportion of **B** (and a lower proportion of either **G** or **R**); objects that appear purple are represented as having a roughly equal proportion of **B** and **R**, and objects that appear green are represented as having a high proportion of **G** (and a lower proportion of either **Y** or **B**). There is therefore a perceptually obvious respect in which blue is more similar to purple than to green. Namely, there is a hue-magnitude (**B**) that all blue-appearing objects and purple-appearing objects, but not all green-appearing objects, are represented as having.

Overall I am more sympathetic to the hue magnitudes response than the experiential response, and indeed some of what I say below is compatible with it. I shall not discuss these responses in further detail here, however, for my intention is to go on the offensive by introducing three new problems for the argument from colour similarity. The first problem builds on the argument in the previous chapter regarding categorical perception of colour. That argument established that many distinctive phenomenally apparent similarities among the colours are not related to the core abilities which are constitutive of colour vision. As such I believe that it is highly dubitable as to whether we should take such apparent colour similarities to impose strong constraints on colour ontology. The second and third problems are independent of these points and hence independent of the foregoing arguments in this thesis. I shall argue that the argument from colour similarity founders as it fails to acknowledge the *context sensitivity* of similarity and the *presentation sensitivity of our knowledge* of similarity. These failures significantly undermine the force of the argument and thereby allay the most significant threat to the reflectance physicalist.

[180] For reasons stated in the previous chapter I do not agree with the practice of labelling the opponent hues with the standard colour category terms ‘red’, ‘green’, and so on.

4.3. Categorical Perception and Similarity Effects

The argument from colour similarity is premised on the observation that colours ordinarily phenomenally appear to stand in various distinctive relations of similarity and difference. The objection to reflectance physicalism is that none of these phenomenally apparent colour similarities obtains among the relevant surface reflectance properties. In appealing to colour appearances in this way, the argument marks a clear application of the mainstream *experiential view* of colour vision, on which candidate views of colour ontology are subject to strong constraints derived from the representational and structural features of our ordinary conscious visual experience of colour. In this case the constraint imposed on reflectance physicalism is that it is required to recover the relational similarity structure between the colours that is apparent in our ordinary visual experience.

In this thesis I have been developing the alternative *constitutive approach* to understanding colour vision. This approach attempts to reduce our dependence on visual phenomenology as a source of evidence as to the nature of colour vision, and thereby to limit the constraints imposed on colour ontology by visual phenomenology. As I have reiterated at various points, my view is that introspection on phenomenology forms a poor foundation for colour ontology as it obscures the latent complexity in our visual experience of colour and is also liable to mislead as to the purported representational objectivity of colour vision. An advertised advantage of the constitutive approach to colour vision is that it helps us to avoid making mistaken inferences about colour ontology due to the inessential, anthropocentric, representational, organisational, or structural features apparent in our ordinary visual experience of colour.

In the previous chapter I deployed the constitutive approach in arguing that a particularly distinctive feature of our ordinary visual experience of colour, namely

categorical perception, is not part of the constitutive basis of colour vision. Colour CP was characterised via 'differential similarity effects' such that a subject possesses CP for *green*, for instance, just in case two within-category shades *a* and *b* would appear more similar to her than two across-category shades *b* and *c*, despite the fact that the pairs *a* and *b*, and *b* and *c*, are separated by exactly the same discriminable distance. One way to read the conclusion of this argument, therefore, is that at least some of the distinctive similarity relations that phenomenally appear to hold among the colours are strictly unrelated to the core representational, discriminatory, capacity that is constitutive of possessing colour vision.

This argument forms the thin end of a wedge with which we can permanently jam appeals to phenomenally apparent colour similarities in discussions of colour ontology. Suppose that reflectance physicalism is true and that the shades $\text{blue}_1 = R_b$, $\text{purple}_2 = R_p$, and $\text{green}_3 = R_g$. The argument from colour similarity contends that while ordinary visual experience suggests that blue_1 is more similar to purple_2 than green_3, it is not at all evident that R_b is more similar to R_p than R_g. This is supposed to defeat reflectance physicalism. Given the foregoing conclusion regarding CP for colour, however, we might wonder why *any* phenomenally apparent similarity relations should be taken as definitive in matters of colour ontology. By the argument of the previous chapter, two organisms exactly alike in respect of their ability visually to perceive colour could differ in respect of the phenomenally apparent similarity effects characteristic of CP for colour. The relevant phenomenally apparent similarity relations are thus deemed inessential to understanding the abilities which are constitutive of possessing colour vision. Extending on this case, for any colour similarity that is ordinarily phenomenally apparent to a subject *S*, we could raise the possibility that an organism possessing a comparable degree of colour vision to *S* could fail to experience the relevant shades of colour as standing in that exact same similarity relation. In such cases, the organism's constitutive discriminatory abilities with respect to colour would be comparable

to those possessed by *S*, but its categorisation, organisation, or visual perceptual grouping of these colours would be different. Given that these organisms *ex hypothesi* possess a comparable ability to perceive colour, it is unclear why the differences in these phenomenally apparent similarities should have any bearing whatsoever on the ontological question as to which properties this shared ability enables them to perceive. In the following two sections I shall introduce two further problems for the argument from colour similarity which do not depend on this admittedly fairly contentious stance. Fortunately these further problems alone suffice to defuse the argument from colour similarity and with it the greatest threat to reflectance physicalism.

4.4. The Context Sensitivity of Similarity

Suppose that I encounter a larger than average mouse in my kitchen. I utter the following:

(MOUSE) That mouse is big.

I then pause to ponder the following identity claim:

(PHYS-1) That mouse=that animal.

Now by Leibniz's Law,

(LL-1) If that mouse=that animal, then that mouse is big iff that animal is big.

Knowing what I know about animals, however, I confidently assert:

(ANIMAL) That animal is not big.

Adapting the reasoning employed in the argument from colour similarity, I infer by modus tollens on LL-1 that PHYS-1 is false: that mouse≠that animal. Now, clearly this is a poor argument. The problem here is not PHYS-1, but the tacit assumption that the predicate 'big' is *semantically univocal* across MOUSE and ANIMAL. In the remainder of this section, I

sketch an analogous context response to the argument from colour similarity. The proposal is that the change in context from SIM to SURF shifts the interpretation of the similarity relation, thus rendering the argument invalid.[181] Although this point does not conclusively undermine the argument, the onus is firmly on my opponent to eliminate the possibility of contextual variation by clarifying the operative notion of similarity. Unless such clarity can be achieved, the threat of the context response remains. As we shall see, it proves to be surprisingly difficult to pin down such similarity claims. The context response ultimately thus forces us to be far more critical in thinking about the notions of similarity operating in discussions of colour and colour ontology.

One reason to suspect that the argument from colour similarity is analogous to the mouse/animal example above is that similarity relations are notoriously context sensitive. As Lewis observes (1986: 254-255),

> Many different relations, some more stringent and some less, some stressing some respects of comparison and others stressing others, have a claim to be called "similarity". The exact meaning of... "similar" is neither constant nor determinate... I can say that Ted and Fred are very much alike, yet very different. Uniform resolutions [of these similarity relations] would make that a contradiction – so much the worse for them!

Empirical studies confirm that judgements of similarity are extremely variable. Medin, Goldstone, and Gentner (1993: 271) report experiments which 'show similarity to be dynamic and context dependent'. Murphy and Medin (1985: 296) say that 'the relative weighting of a feature... varies with the stimulus context and task, so that there is no unique

[181] Pautz (2006: 540, fn. 4) acknowledges the context sensitivity of the similarity relation but argues that this does not undermine the argument from colour similarity. His reasoning is that SURF 'does not come out true under any resolution of the vagueness compatible with the ordinary meaning of "x resembles y more than z".' I disagree, but in any event this does not affect the present proposal, which is that SIM and SURF involve different similarity relations, thus invalidating the argument.

answer to the question of how similar is one object to another'.[182] Given the clear change from a commonsense visual context in SIM to a reductive scientific context in SURF, a shift in the operative standards for similarity relation is not only conceivable but highly likely in this case.

The context response is bolstered by Quine's observation that different systems or levels of theoretical description typically have different associated standards for similarity. Quine held that the modification or revision of intuitive notions of similarity is integral to the process of finessing our ordinary categories into scientifically viable kinds. Quine (1969: 234) thought that grouping into kinds and notions of similarity or resemblance are 'variants or adaptations of a single notion'. As he explains (1969: 240),

> A sense of comparative similarity... is one of man's animal endowments. Insofar as it fits in with regular properties of nature, so as to afford us reasonable success in our primitive inductions and expectations, it is presumably an evolutionary product of natural selection. Secondly, ..., one's sense of similarity or one's system of kinds develops and changes and even turns multiple as one matures, making perhaps for increasingly dependable prediction. And at length standards of similarity set which are geared to theoretical science. This development is a development away from the immediate, subjective, animal sense of similarity to the remoter objectivity of a similarity determined by scientific hypothesis and posits and constructs. Things are similar in the later or theoretical sense to the degree that they are interchangeable parts of the cosmic machine revealed by science.

Quine (1969: 238) offers the following example by way of illustration,

> A crude example is the modification of the notion of fish by excluding whales and porpoises. Another taxonomic example is the grouping of kangaroos, opossums, and

[182] Goodman (1970: 444) argues that similarity, like motion, requires a frame of reference. See also Barsalou (1982), Tversky (1977).

> marsupial mice in a single kind, marsupials, while excluding ordinary mice. By primitive standards the marsupial mouse is more similar to the ordinary mouse than to the kangaroo; by theoretical standards the reverse is true.

Extending this to the case of colour, the thought is that while $blue_1$ is more similar to $purple_2$ than $green_3$ by the standards operating in casual observational contexts, this claim fails to hold by the standards for similarity operating in optics. These differences needn't undermine the proposed identities between colours and reflectances, any more than we take Quine's example to undermine the identities between ordinary and biological notions of *kangaroo*, *whale*, and so on. For colours as for other properties or kinds, different classificatory contexts simply invoke different standards for similarity.[183]

The foregoing points in support of the context response admittedly are more circumstantial than conclusive: we have established only that it is highly *likely* that the shift in context between SIM and SURF would affect the similarity relation. Nevertheless, the extreme context sensitivity of similarity together with Quine's observation about similarities at different theoretical levels means that the onus is on my opponent to rule out such effects in the argument from colour similarity. Fortunately for my opponent, there seems to be an obvious way to achieve this: namely, *make the argument more precise*! If SIM and SURF can be reformulated in a way that guarantees the univocity of the similarity relation, then the context response evaporates. I shall now consider a number of attempts to impose univocity

[183] For related examples, see Goodman (1970). Although Quine's primary concern is with natural kinds, his claims apply to property identifications more generally. Quine himself tended towards eliminativism about ordinary kinds and indeed the notion of similarity itself. Quine (1969: 237) says that 'the brute irrationality of our sense of similarity, its irrelevance to anything in logic or mathematics, offers little reason to expect that this sense is somehow in tune with the world – a world which, unlike language, we never made.' Needless to say, I do not share these views. I adopt a more permissive reading of Quine as establishing contextual variance in theoretical standards for similarity.

on the argument and argue that all of them are problematic. Despite their apparent simplicity and intuitiveness, colour similarity claims are extremely difficult to pin down.

A first attempt at imposing univocity across SIM and SURF would be to fix features of the context as follows:

(SIM-C) 'Blue_1 is more similar to purple_2 than green_3' is true as said by *S* in *w* at *t*.

?? (SURF-C) 'R_b is more similar to R_p than R_g' is true as said by *S* in *w* at *t*.

Fixing the subject, world, and time of utterance is often what people mean by 'fixing the context'. We should bear in mind, however, that *referring terms themselves* can affect the interpretation of predicated similarity relations, even given a fixed subject, world, and time triple. For example, consider the following sentences, all said by a speaker *S* in *w* at *t*:

4. That mouse is big and that elephant is not big.

5. Mickey is big and Jumbo is not big.

6. Every animal in the zoo is big.

In 4 the interpretation of the predicate 'big' shifts within a sentence at a context due to the different subjects to which the predicate is applied. One plausible explanation is that the different subject terms shift the *comparison class* for the predicate from *mice* to *elephants*: whereas that mouse is big {for a mouse}, that elephant is not big {for an elephant}.[184] The point applies to proper names (5) and even extends to *quantificational* subjects (6) as noted by Kennedy (2007: 8-9).[185] Assuming that predicates of similarity likewise can be influenced

[184] See Kennedy (2007: 21, fn. 19). One option is that the subject term *automatically* and *without exception* fixes the comparison class for the predicate, such that it is impossible to interpret 'big' in the context of 'that mouse is big' as meaning 'big {for an elephant}'. A more plausible option is that the subject term fixes a default interpretation for the comparison class but that this leaves room for further contextual variation, such that in unusual cases we might force a reading of 'that mouse is big' as 'that mouse is big {for an elephant}'.

[185] We might try to impose univocity by making the similarity relation in SURF *anaphoric* on SIM, as in 'blue_1 is more similar to purple_2 than green_3, and *R1*, *R2*, and *R3* stand in that exact same relation'. This won't work

by the subject terms to which they are applied, SIM-C and SURF-C leave open the possibility of contextual variation.

This point is closely related to a feature of Lewis's (1986) *counterpart theory*.[186] On Lewis's account, different names for the same object can have different associated counterpart relations. As a consequence, modal sentences of the form ⌜◊*Fa*⌝ and ⌜◊*Fb*⌝ can differ in truth value even when *a*=*b*. To illustrate, let the name 'Goliath' refer to the statue and 'Lumpl' refer to the lump of clay from which the statue is made. Now consider some counterfactual *w* in which the statue has been deconstructed and the clay put to some other use.

1. Lumpl would have survived in *w*.

2. Goliath would not have survived in *w*.

3. Therefore, Lumpl ≠ Goliath.

This argument is not valid on counterpart theory because the names 'Lumpl' and 'Goliath' have different associated counterpart relations. For Lewis (1986: 256, n.40) this means that the names have different associated transworld similarity relations, for counterpart relations are defined in terms of similarity. Now, although I do not endorse counterpart theory, I do endorse the claim that different names for the same object or property can have different associated similarity relations. The reflectance physicalist could thus insist that the terms 'blue_1' and 'R_b' represent the very same property (*de re*, no less) but that SIM and SURF differ in truth value on account of the different similarity relations associated with each term.

---

either, as shown by examples such as 'that mouse is big and that sky scraper is too'. See (Kennedy 2007: 8-9; 20-21).

[186] Here I am extremely indebted to personal correspondence with Adam Pautz, who brought to my attention the comparison with counterpart theory.

A second attempt at ensuring univocity across SIM and SURF would be to stipulate a specific interpretation for the similarity relation.[187] As the above quote from Lewis illustrates, different contexts typically influence the similarity relation by making salient different *respects* of comparison. Ted and Fred might be extremely similar in respect of height, say, but extremely dissimilar in respect of girth. We might thus eliminate the threat of contextual variation by specifying a respect of similarity in SIM and SURF as follows:

(SIM-R) Blue_1 is more similar to purple_2 than green_3 *in respect r*.

?? (SURF-R) R_b is more similar to R_p than R_g *in respect r*.

What might this respect of similarity be? Clearly not any old respect will do, for some respects of similarity will apply to the reflectances as much as to the colours. For example, both the colours and the reflectances satisfy the comparative relation of similarity in respect of *being favoured by Nora*, for Nora happens to like both blue_1 and purple_2 and R_b and R_p very much, but likes green_3 and R_g very little. What my opponent needs, then, is a more *discerning* respect of similarity that clearly holds between the colours and yet fails to hold between the reflectances. Byrne (2003: 641) expands on this thought as follows,

> Anything is similar to anything, provided the respects of similarity are allowed to be gerrymandered or gruesome, as Goodman observed. But similarity in non-gruesome or – as I shall say – *genuine* respects is much less ecumenical. Colours, it seems, provide a compelling illustration of the distinction as applied to similarities among *properties*. For instance, in innumerable gruesome respects, blue is more similar to yellow than to purple. But in a genuine respect, blue is more similar to purple than to yellow (genuinely more similar, as I shall sometimes put it).[188]

[187] If different colour terms do indeed have different associated similarity relations, then there inevitably will be a certain amount of linguistic artificiality in this approach. This causes more problems for advocates of the argument, who tend to rely on our commonsense understanding of similarity claims in order to motivate their premises. See for example Pautz (2006: *passim*).
[188] Also see Shoemaker (1975: 11).

If Byrne is correct about this, we could read SIM-R as positing some *genuine* respect *r* in which $blue_1$ is more similar to $purple_2$ than $green_3$. The question now, however, is what this respect might be. People on both sides of the debate have been remarkably vague on this point. The thought seems to be that it is just *intuitively obvious* that the colours are genuinely similar and *intuitively obvious* that the reflectances are not. Until someone clarifies which genuine respect of similarity is supposed to apply in SIM-R, however, there is simply no guarantee that SURF-R is being evaluated via the exact same standard of similarity and hence no guarantee against context playing some illicit role in shifting the operative respects of comparison.[189]

In any event, there is something dialectically quite odd about assuming that ordinary colour similarity claims such as SIM involve a notion of genuine similarity. Following Lewis, genuine similarities are typically understood as those which 'carve nature at the joints' in some sense.[190] Armstrong (1978: 7-9), for example, holds that genuine similarities are discovered only *aposteriori*, while Sydney Shoemaker (2003: 259) claims that 'a relation is a genuine similarity relation if it makes properties similar to the extent that their instantiation bestows similar causal powers.' On most accounts, then, genuine similarities are the province of the natural sciences, relating in some sense to the individuation of scientific kinds.[191] Given this fact, it seems highly unusual for my opponent to interpret SIM as predicating a genuine similarity among the colours. SIM is supposed to be obvious from ordinary visual

[189] My opponent might propose the following quantified reading: while there is *some* genuine respect in which $blue_1$ is more similar to $purple_2$ than $green_3$, there is *no* genuine respect in which R_b is more similar to R_p than R_g. My response would be that context could also play a role in restricting the domains of the quantifiers: different contexts of theoretical classification may be associated with different and potentially non-overlapping sets of respects – even *genuine* respects. For example, in the context of discussing some way in which Ted and Fred are extremely different, it might be true to say that there is simply no respect in which Ted is similar to Fred, whereas in other conversational contexts this would be false.

[190] Lewis held that two things are genuinely similar to the extent that they share *natural* properties, which are those that carve nature at the joints. Given that Lewis took naturalness to come in degrees, this implies that there are degrees of genuineness of similarity. Ultimately Lewis thought the only *truly* natural properties were those posited by fundamental physics. Hence the notion of genuine similarity as it applies to other properties is somewhat tendentious.

[191] See also Goodman (1965: 121).

experience, a bit of pre-theoretic data. Surely no party to this debate, however, thinks that it obvious from visual experience that colour similarities carve nature at the joints. Indeed some advocates of the argument from colour similarity surely would want to deny this![192] Perhaps the thought rather is that *if* reflectance physicalism is true, then visually apparent colour similarities such as SIM *must* carve nature at the joints in some sense. The reasoning might be that if colours are surface reflectances, then phenomenally apparent similarities among the colours are *ipso facto* similarities among scientific kinds and hence genuine similarities. The argument from colour similarity would then take a slightly different form, with the addition of the following premise:

> (GEN) Necessarily, if reflectance physicalism is true, then all phenomenally apparent colour similarities are similarities in some genuine respect.

Reflectance physicalists such as Byrne and Hilbert (1997c: 276) and Byrne (2003: 655) would deny this premise, however, because they identify colours with types or classes of reflectance that are not of any particular interest to physicists or optical scientists, which they take to rule out the possibility of genuine similarity relations between these types. As we shall see in the next section, I believe that this is somewhat premature for it is arguable that even such *interest relative* similarities can class as genuine similarities. If this is correct, then the reflectance physicalist can respond to the argument even if the premise GEN is granted. Either way, the appeal to genuine similarity fails to save the argument from colour similarity.

An alternative reading of SIM-R and SURF-R would be that *r* is some *intrinsic* respect of similarity, as suggested by Pautz (2006: 554),

[192] For this reason, I am highly sceptical of the *epistemic* form of the argument from colour similarity as presented by Johnston (1992), which simply takes it for granted that visually apparent colour similarities are genuine similarities. For related discussion see Byrne (2003).

> It seems perceptually obvious that the fact that blue resembles purple more than green obtains by virtue of the intrinsic character of the colours. To see that the colours fall into this resemblance order, you only need to focus on the colours themselves.[193]

Again, however, it is far from clear whether the notion of intrinsic similarity should be applied to the colours.[194] The claim that *a* and *b* are similar in some intrinsic respect is equivalent to the claim that there is some intrinsic *property* of *a* and *b* in respect of which *a* and *b* are similar. If *a* and *b* are *objects*, this equivalence is unproblematic; for presumably we can all agree on some candidate intrinsic properties in respect of which objects can be more or less similar. Assuming that *a* and *b* have their *shapes* intrinsically, for example, similarity in respect of shape would qualify as an intrinsic similarity between *a* and *b*. In contrast, I can think of no straightforward examples of intrinsic properties of *properties*. Perhaps *being a property* is one, but this is not a respect in which properties could be more or less similar. One might be tempted to claim that $blue_1$ is more similar to $purple_2$ than $green_3$ *in respect of hue*, where hues are properties that colours have intrinsically.[195] As Armstrong (1978: 105-108) noted years ago, however, this proposal fails because $blue_1$, $purple_2$, and $green_3$ do not *instantiate* hue properties; they *are* hue properties. By the same token, it makes no sense to say that *6'2"* is more similar to *6'3"* than *6'0" in respect of height*, for *6'0"*, *6'2"*, and *6'3"* *are* heights.[196]

A possible objection is that the demand to *specify* the operative intrinsic respect of similarity in SIM-R and SURF-R is too strong. There are many cases in which we know that two things are similar without being able to specify the precise respect in which they are

[193] Also see Pautz (2006: 538, 550).
[194] Byrne & Hilbert (1997c: 277) also find it unclear 'what people are up to' in making such claims.
[195] A related proposal was made by Adam Stewart-Wallace in a response to this work.
[196] Armstrong's (1978: 109, 111, 116) resolution of this difficulty was to analyse property resemblances via a relation of *partial identity* between universals. Roughly, the idea is that universals resemble each other to the extent that they have common universals as parts or constituents. Pautz (1997) rejects Armstrong's account, however. For further criticism, see Eddon (2007).

similar. By the same token, the objection runs, we can know that two things are *intrinsically* similar while remaining ignorant of the precise respect in which they are intrinsically similar. In particular, we know that the *colours* stand in relations of intrinsic similarity, as Pautz remarks, because we can acquire knowledge of these similarities just by visually 'focusing on the colours themselves'. It is thus evident that colour similarities obtain in virtue of the nature of the colours themselves, and not due to any extraneous factors. By contrast, from what we know about the nature of the reflectances R_b, R_p, and R_g, as revealed by their reflectance curves, it is evident that these properties fail to stand in corresponding relations of intrinsic similarity. We could therefore reformulate the argument from colour similarity as follows: while there is *some* intrinsic respect (whatever that might be) in which blue_1 is more similar to purple_2 than green_3, there is simply *no* intrinsic respect in which R_b is more similar to R_p than R_g.

My response is that this objection fails to address the problem at issue. As noted above, any intrinsic similarity between *a* and *b* requires there to be some intrinsic property of *a* and *b* in respect of which they are similar. The problem is that it is entirely unclear what the relevant second-order intrinsic properties of colour properties could be, or indeed whether there are any such properties. This is a deep and difficult issue in the metaphysics of properties, and not something that can be settled by intuitions derived from 'focusing on the colours themselves' in visual experience. Such intuitions at best establish that colour similarities satisfy an *epistemic* notion of intrinsicness, on which a thing's intrinsic properties are those that we can know about simply by inspecting the thing itself.[197] The question at

[197] This distinction is noted by Dunn (1990: 178): 'epistemologically, an intrinsic property would be a property that one could determine by inspection of the object itself - in particular, for a physical object, one would not have to look outside its region of space-time'. It is unclear to me, however, how to understand epistemic intrinsicness as applied to properties: in what sense do we 'inspect', 'attend to', or 'see' colour properties themselves in visual experience? Are colour properties the 'objects' or colour perception in a sense analogous to that in which tables and chairs are the 'objects' of object perception? Our epistemic and perceptual position with respect to colour properties *as such* is a topic in need of much further attention. See Byrne (2003: 655) for related comments.

issue, however, is whether colour similarities are *metaphysically* intrinsic, that is, whether these relations hold by virtue of the nature of colour properties themselves. Our visual experience of colour is simply ill-equipped to answer this question.[198] As such, my opponent cannot simply help herself to the notion of similarity in some intrinsic respect in interpreting SIM, for SIM is supposed to report our commonsense visual intuitions about the colours.

Where does this leave us with regards to the context response to the argument from colour similarity? As noted above, the extreme context sensitivity of similarity puts the onus on advocates of the argument to eliminate the possibility that SIM and SURF involve different notions of similarity. Fixing standard features of the context such as the subject, time, and world of utterance will not suffice, for different referring terms for the colours may have different associated similarity relations. A more promising route is to impose a univocal interpretation on the similarity relation by specifying a particular *respect* of similarity. This respect of similarity has to be *discerning* in order to rule out its application to the reflectances. Having canvassed some candidates, however, it remains extremely difficult to see what this respect of similarity might be. This point is not exactly conclusive against the argument. Perhaps there are candidate respects that I haven't considered. Perhaps there are other ways to eliminate the possibility of contextual variation. Nonetheless, the context response shifts the burden firmly onto my opponent to propose some such solution or to withdraw the argument. More generally, the context response forces us to be more cautious and critical in thinking about colour similarity and its relationship to colour ontology. Until we can achieve clarity about the nature of similarity as it applies to the colours, how can we hold reflectance physicalism to account for failing to find corresponding similarities among the reflectances? In the next section, I shift attention to the epistemic phenomenon of the

[198] Advocates of *Revelation*, the principle that every fact about the nature of a colour property is somehow made entirely manifest in visual experience, would think otherwise. For example, see Johnston (1992). This principle clearly stands in diametric opposition to the constitutive approach to colour vision and colour ontology.

presentation sensitivity of knowledge of similarity. The argument from colour similarity makes some highly questionable assumptions about the extent of our ability to acquire knowledge of similarities. Once these assumptions are rejected, the argument falls once and for all.

4.5. Presentation Sensitivity and Knowledge of Colour Similarities

The crucial move in the argument from colour similarity was to reject SURF, the reflectance similarity claim. As noted above, the basis for this move was that the truth of SURF was not *evident* from the reflectance spectra for R_b, R_p, and R_g: there was no *obvious* sense in which the reflectance spectrum for R_b seemed more similar to the reflectance spectrum for R_p than the reflectance spectrum for R_g. At the crucial point, then, the argument relies on the following epistemic principle:

> (EVIDENT) If R_b is more similar to R_p than R_g, then it is evident from canonical physical descriptions of the reflectance properties R_b, R_p, and R_g, that R_b is more similar to R_p than R_g.

Note that EVIDENT entails the following:

> (EVIDENT*) If reflectance physicalism is true, then if blue_1 is more similar to purple_2 than green_3, then it is evident from canonical physical descriptions of the reflectance properties R_b, R_p, and R_g, that R_b is more similar to R_p than R_g.[199]

Byrne (2003: 642, fn. 5) highlights the central role played by EVIDENT* in the argument,

[199] Given the arguments of the previous section it is pertinent to note that EVIDENT* only holds on the assumption that there is some univocal interpretation of the similarity relation as it applies to both the colours and the reflectances. Once again it is therefore crucial to the argument of colour similarity that my opponent can make good on this assumption in interpreting EVIDENT*. For the purposes of argument, in this section I shall ignore this issue. By the by, we should also note that deriving EVIDENT* from EVIDENT does not require the illicit substitution of co-referring terms into an opaque context.

> The argument from similarity assumes that if physicalism about colour is true, then any genuine respects of similarity between the colours will be evident at the level of the canonical physical description of those properties. Absent this assumption, the failure of colour science to find the appropriate genuine respects of similarity would not show that such respects did not exist, thus blocking the argument.[200]

To make the reasoning here explicit, Byrne claims that if EVIDENT* is false, then EVIDENT is false, and if EVIDENT is false then this undermines the crucial inference from the lack of evident similarity between the reflectance spectra for R_b, R_p, and R_g, to the falsity of SURF. In what follows, I shall argue that on most readings EVIDENT* is indeed false.[201] On these readings, the reflectance physicalist should reject the assumption that all similarities among the colours are evident from canonical physical descriptions of reflectance properties.[202] This would critically undermine the argument from colour similarity, for the fact that SURF is not evident from the canonical physical descriptions of R_b, R_p, and R_g would be no barrier to its truth. There is one reading on which EVIDENT* might be true, but on this reading the argument from colour similarity poses no great threat to the reflectance physicalist.

As everyone knows, all knowledge is presentation sensitive. I can know that Dame Edna is glamorous while disbelieving that Barry Humphreys is glamorous, even though Barry Humphreys *is* Dame Edna.[203] Knowledge of similarities is no exception, of course. Suppose that blue_1 is Cameron's favourite colour, purple_2 is Brown's favourite colour, and green_3 is Blair's favourite colour. Absenting auxiliary knowledge of these identities, I am not in a position to know that Cameron's favourite colour is more similar to Brown's favourite colour

[200] Byrne is discussing the distinct *epistemological* form of the argument from colour structure as presented by Johnston (1992). Nevertheless as we saw above the assumption that he outlines also underlies the metaphysical form of the argument discussed here.
[201] One way to summarise this position is that the colour physicalist should advocate the metaphysical, but not epistemic, reduction of colour properties to reflectances.
[202] Byrne (2003) accepts this assumption. This marks a key difference between our approaches.
[203] Some hardcore Millians or direct reference theorists about propositional content would deny this. I assume the falsity of such views.

than Blair's favourite colour. To pick another example, suppose that we come across three seated individuals with heights *6'0"*, *6'2"*, and *6'3"*. Each individual has a slightly different seated posture and so under these presentations of their heights, I am unable to acquire the knowledge that the height of the second individual is more similar to the height of the third individual than the height of the first individual.[204]

No party to this debate would assume, then, that two similar things can be known to be similar under all presentations. Nevertheless, the assumption EVIDENT* presupposes the following restricted thesis of presentation insensitivity:

> (PRESENT) For all *canonical physical* presentations *x*, *y*, *z*, if *ox=a*, *oy=b*, and *oz=c*, then if *a* is more similar to *b* than *c*, it is knowable that *ox* is more similar to *oy* than *oz*.

Here and throughout the variables *a*, *b*, *c*... range over objects, *x*, *y*, *z*... over presentations of objects, and *o* is a function from presentations to objects presented, so ⌜*ox*⌝ is to be read as 'the object presented by *x*'. Let $a=blue_1$, $b=purple_2$, and $c=green_3$. If reflectance physicalism is true, then the spectral reflectance curves for R_b, R_p, and R_g are canonical physical presentations of *a*, *b*, and *c* respectively.[205] PRESENT then entails that if $blue_1$ is more similar to $purple_2$ than $green_3$, it is knowable that the property presented by the reflectance spectrum for R_b is more similar to the property presented by the reflectance spectrum for R_p

[204] I have learned much from related discussion of the presentation sensitivity of discrimination, understood as the acquisition of knowledge of distinctness, in Williamson (1990). This notion of discrimination played a key role in the constitutive account of colour vision provided in chapter one. It is important to note that no heavy duty metaphysics of presentations is intended or needed; in short, we don't have to be Fregeans to admit presentations. Presentation also need not involve *re*presentation. While properties can of course be presented by token descriptions ('the colour of my armchair') or by particular perceptual states (my state of visually perceiving the colour of my armchair), they can also be presented simply by their instances at particular places and times (my armchair).

[205] Notice that PRESENT is trivially satisfied if *a*, *b*, and *c* have no canonical physical presentations. If colour physicalism is true, however, then there will of course be canonical physical presentations of the colours, namely their spectral reflectance spectra. I have intentionally neglected to provide criteria as to what constitutes a canonical physical presentation of some object or property. I shall address this point more directly later on. For now I assume some intuitive understanding of this notion.

than the property presented by the reflectance spectrum for R_g. I shall now argue that PRESENT is false.

The argument is that, for some theoretical reductions of *Fs* to *Gs*, some similarities among the *Fs* are not evident from all canonical physical presentations of the *Gs*. This argument does not rely on the assumption of some retrograde or gerrymandered notion of similarity, but applies also to *genuine* similarities, such as they may be. This allows me to present my opponent in the strongest possible light, granting that the notion of genuine similarity can be applied to the colours. Relating this back to the previous section, I believe that the reflectance physicalist can grant the assumption GEN and still respond to the argument from colour similarity. As we shall now see, whether or not genuine similarities at the *reduced* level are evident at the *reducing* level depends on one's choice of physical description or representation of the reducing properties. Whereas above it was suggested that similarity claims might be true in some linguistic contexts and false in others, the point here is that similarity claims might be knowable in some epistemic contexts and not knowable in others.[206] Here is an example. Many people would judge intuitively that there is some genuine sense in which diamonds are more similar to sapphires than coal. Now as reductive claims go, it is uncontroversial that diamond and coal are allotropes of carbon – tetrahedrally bonded and amorphous carbon, respectively – whereas sapphire is a form of aluminium oxide. Hence chemical descriptions or presentations of these compounds will class as canonical physical presentations of diamond, sapphire, and coal. One common and

[206] There are some interesting parallels here with the knowledge argument for phenomenal qualities, (although I stress that I do not take colour similarities to be phenomenal facts, that is, facts about what it is like to experience colours). The proposal is that some properties of colour properties are not knowable from canonical physical descriptions of these properties. If you like, Mary could know all the colour physics facts and yet fail to know that blue is more similar to purple than green. Colour physicalists might worry that this entails the existence of non-physical colour facts. (This is, in effect, the objection that Byrne raises to my approach at 2003: 642, fn. 5.) Following the lead of physicalist responses to the knowledge argument, however, one could argue either that not every physical colour fact is knowable under canonical physical presentations, ('Mary didn't know all the physical facts'); or perhaps that canonical physical presentations of colour facts simply do not enable us to *recognise* the relevant similarities, ('Mary learns an old fact under a new guise'). These connections will be explored further in future work.

presumably canonical way of describing these compounds is by their *chemical* or *molecular formulae*, which represent facts about the type and number of atoms constituting each compound. Under these presentations, however, it is not at all evident that diamond is genuinely more similar to sapphire than coal. On the contrary, relative to the respects of comparison made evident by their chemical formulae, diamond is more similar to coal than sapphire. Of course, another canonical way of presenting these compounds is by their *structural formulae*, which additionally represent the type of bonding between atoms and thus make evident the intuitive similarity between diamond and sapphire. PRESENT still fails, however, for this genuine similarity is evident only from *some*, but not all, canonical physical descriptions of these chemical compounds.

It might be objected that molecular formulae are not canonical physical presentations of chemical compounds, which would explain why the genuine similarity between diamonds and sapphires is not evident under these presentations. One might think that molecular formulae simply do not provide enough information about the *nature* of chemical compounds to make evident all genuine similarities among these kinds. In response, although molecular formulae admittedly present comparatively little information about chemical compounds, they are still among the most prominent and widely used descriptors within chemistry. As such I see no principled reason to deny that they are canonically physical descriptions. Even if this point is not granted, however, it is easy enough to rerun the argument using chemical descriptors that all parties would agree are canonically physical. For example, presumably everyone would grant that structural formulae are canonical physical presentations of chemical compounds. Many genuine chemical (dis)similarities are not evident under such presentations. To illustrate, *stereoisomers* are compounds with exactly the same atomic composition and bond structure but different geometrical profiles. Geometrical differences in the relative rotation of bonded atoms can affect the *potential energy* of a compound.

Similarities or dissimilarities in potential energy would not be evident or predictable from the structural formulae for stereoisomers, as these do not represent the *spatial position* of atoms within molecules. As in the diamond/sapphire case, different chemical descriptions might be better suited to predicting these similarities. *Newman projections*, for example, represent the angles between atoms, a crucial determinant of potential energy. The key point here is that while all of these descriptions class as canonically physical, some of them fail to make evident certain genuine similarities between chemical compounds. That is to say, PRESENT is false.

My opponent might counter with the following revision to PRESENT:

(PRESENT+) For *some* canonical physical presentations *x*, *y*, *z*, *ox=a*, *oy=b*, and *oz=c*, and if *a* is (genuinely) more similar to *b* than *c*, it is knowable that *ox* is (genuinely) more similar to *oy* than *oz*.

PRESENT+ states that there will be at least one canonical physical presentation of *a*, *b*, and *c*, such that if *a* is (genuinely) more similar to *b* than *c*, then this presentation will make evident that similarity. If sound, this principle revives the threat to SURF, for our opponent will insist that there is simply *no* canonical physical description of the reflectances R_b, R_p, and R_g under which it is evident that R_b is (genuinely) more similar to R_p than R_g.

As an opening gambit, the reflectance physicalist might insist that, for all we know, *some future* canonical physical description of surface reflectance properties will make evident the fact that R_b is more similar to R_p than R_g. This would be enough to satisfy PRESENT+ and hence to rebut the argument from colour similarity. Given that reflectance physicalists differ markedly on their optimism on this point,[207] I propose to take a more direct approach in arguing against PRESENT+. So far I have intentionally declined to take an explicit stand on

[207] Compare McLaughlin (2003: 116) and Byrne (2003: 648).

what counts as a 'canonical physical description' of a property. I shall now argue that, if we take a *conservative* view on this matter then PRESENT+ is false, whereas if we adopt a more *liberal* view of canonically physical descriptions then PRESENT+ may be true, but the principle no longer poses a significant threat to reflectance physicalism. Either way, PRESENT+ fails to save the argument from colour similarity.

My argument proceeds via further scientific examples involving identities between *chemical* and *biochemical* molecules. There are some similarities between biochemical molecules that a) plausibly are genuine similarities, b) are not evident under any standard chemical molecular descriptions or representations of these molecules, and c) can be made evident only via *interest relative* descriptions or representations, by which I mean descriptions that encode features of chemical structure relevant to a molecule's biochemical activity. This presents us with a dialectical fork. On the one hand, if we take the conservative view that such interest relative descriptions are *not* canonical physical descriptions of chemical molecules, then PRESENT+ is false, for there are genuine biochemical similarities between molecules that are not evident under any canonical physical presentations. On the other hand, if we take the more liberal view that such interest relative descriptions *are* canonically physical, then PRESENT+ might well be true; but the reflectance physicalist now can be confident of meeting the challenge posed by the argument from colour similarity. No longer restricted to *optics*-relative descriptions and representations of reflectance properties such as their reflectance curves, interest relative descriptions provide the reflectance physicalist with a host of new explanatory possibilities with respect to colour similarities. One potentially promising route, inspired by the analogy with biochemistry, would be to appeal to *biological* descriptions of reflectances that highlight features relating to the property's *biological* role. The prospects for such an account will be discussed briefly below.

My argument draws on discussions of drug discovery methods in pharmaceutical research.[208] A fundamental assumption in this field is the 'structure-activity relationship', described by prominent theoretician Kubinyi (1998: 225) as follows,

> The underlying concept is that chemical similarity is reflected by similar biological activities – i.e. chemically closely related analogues should be related in their mode of action, as well as in their relative potencies. This fundamental assumption has, indeed, been used in medicinal chemistry research, and has led to many valuable drugs.

What do we mean here by 'chemical' and 'biochemical' similarity? As the expression 'structure-activity relationship' suggests, chemical similarity relates to *chemical structure*,[209] whereas biochemical similarity relates to *biological activity*, understood specifically in terms of *ligand-receptor relations and interactions*, the potential that a particular molecule – the ligand – has to bind with a biological macromolecule.[210] The structure-activity relationship guides a predictive practice in which the relative chemical similarity between two molecules is used to predict how biochemically similar these molecules will be. This certainly seems intuitive, given that biochemical molecules are type identical to chemical molecules. The assumption is that biochemical similarities therefore will be evident under some presentation or description of the chemical structure of these molecules; in other words, that we will be

[208] The drug discovery process generally starts with a 'lead structure', a molecule that is biochemically efficacious although not optimally so, which is then refined via an evolutionary process of systematic alteration or substitution of 'functional groups', or structural modifications such as the rigidification of unstable molecules, where changes that improve specificity, efficacy, side effects, and so on, are retained, others discarded. There is a sizeable literature dedicated to developing metrics for similarity between chemical compounds. For a review, see Nikolova & Jaworska (2003).

[209] Particularly relevant to determining chemical structure are features of molecular *constitution*, *configuration*, and *conformation*. *Constitution* relates to the manner and sequence of bonding between atoms in a molecule. *Configuration* concerns the 3D arrangement of atoms in the molecule, particularly the valence angles of atoms as discussed above in connection with stereoisomers. *Conformation* relates to the thermodynamically stable spatial arrangements of atoms within a molecule.

[210] Biochemical similarities of this sort are surely *genuine* similarities, whatever sense we might attach to that term. Assuming the above understanding of genuine similarities as 'carving nature at the joints', for example, biochemical similarities class as genuine to the extent that we take biochemistry to describe empirically robust kinds and laws.

able to acquire knowledge of biochemical similarities by evaluating how similar two molecules are in respect of chemical structure. This is a form of *epistemic reductionism* with respect to biochemical similarities, analogous to the assumption EVIDENT* operating in the argument from colour similarity. Framing this assumption in line with PRESENT+, the thought is that biochemical molecules *a*, *b*, and *c*, will have some canonical chemical presentations *x*, *y*, and *z*, and if *a* is biochemically more similar to *b* than *c*, then it will be knowable that *ox* is more similar to *oy* than *oz*.

As Kubinyi goes on to note, however, there are many cases in which apparently chemically similar molecules are nonetheless extremely biochemically dissimilar and vice versa.[211] Such cases are well known in the extant literature. Bajorath (2002: 890) refers to them as 'similarity paradoxes', whereby 'minor chemical modifications of otherwise similar molecules can render them [biochemically] either active or inactive'. Nikolova and Jaworska (2003: 1016) likewise report that 'several surprising structure-activity relationships demonstrate that chemically similar compounds may have significantly different biological actions and activities and different molecules can be very similar in their biological activities...' These cases prima facie establish that for some biochemical molecules *a*, *b*, *c* such that *a* is biochemically more similar to *b* than *c*, it is not evident under any standard chemical structural description or presentation *x*, *y*, *z* that *ox* is more similar to *oy* than *oz*. In other words, some genuine biochemical similarities are simply not evident under any of the structural descriptions commonly employed to assess chemical similarity between molecules.

This is not the end of the story, however. While the similarity paradoxes indicate that standard chemical structural descriptions are sometimes unsuitable for predicting biochemical similarity, there are other ways to describe chemical molecules that are better suited to this

[211] For example, Kubinyi (1998: 228) and (2002: 244): 'The [biochemical] selectivity of these analogues against the two [adrenergic] receptors differs by nearly eight orders of magnitude, despite their close chemical similarity!'

task. The intuitive distinction here is between descriptions suited to understanding chemical structure in its own right, which people seem to have in mind when evaluating chemical similarity, and descriptions that capture patterns or regularities in chemical structure that are relevant to biological activity. As Kubinyi (1998: 249) notes, while the former category yield 'target-independent similarity indices', the prediction of biochemical similarity requires target-*dependent* descriptions which take account of the specific geometrical and topological features, functional groupings, and wide range of bonding behaviours that determine the relations between a ligand and biological macromolecule. Examples include *pharmacophoric descriptors*, as discussed by Patterson and colleagues (1996: 3049),

> [A] synthetic chemist will regard two molecules as similar when their topological descriptions – their networks of atoms and connecting bonds – contain a sufficiently large number of features in common. But, with increasing knowledge of receptor-ligand complex structures, it seems plausible that similarity of pharmacophoric descriptors, critical geometric arrangements of structural features, may be a more important determinant than topology of whether two molecules will bind to the same receptors. And there are very many other properties, such as overall molecular size and lipophilicity, which affect biological potency and may, therefore, be useful diversity descriptors.

Target-dependent descriptions such as pharmacophoric descriptors represent complex aggregations of chemical molecular features.[212] In this respect, they might be considered purely chemical structural descriptions. A key difference from standard chemical molecular descriptions, however, is that the component features are selected specifically on the basis of their ability to predict biological activity. Methods include the complex analysis of high-dimensional datasets in order to find patterns or regularities among the features of biological

[212] Such features include molecules' volume and surface potentials, hydrophobic and polar regions, hydrogen bond donor potentials, acceptor potentials, and molecular electrostatic potentials. See Kubinyi (2002).

active molecules.[213] I therefore class these as *interest relative descriptions*, in the sense that the features represented are more closely related to the classificatory and predictive needs and interests of the biochemist than the chemist. The conclusion that I wish to draw, then, is that some biochemical similarities between chemical molecules are evident only under interest relative descriptions of these molecules, and not under any standard chemical structural descriptions.

How does this conclusion impact on the principle PRESENT+? This depends on whether or not we class interest relative descriptions such as pharmacophoric descriptors as canonical physical presentations. On the one hand, suppose that interest relative descriptions are deemed *not* canonically physical, leaving only standard chemical structural descriptions in this category. PRESENT+ then arguably is false, for as we have seen some genuine biochemical similarities are only evident under interest relative descriptions of chemical molecules. This critically undermines the argument from colour similarity, for without PRESENT+ we are not entitled to infer the falsity of SURF from the mere fact that this similarity is not evident from canonical physical descriptions of the reflectances R_b, R_p, and R_g. Pursuing this line, the reflectance physicalist should hold that the colours can be *metaphysically* but not *epistemically* reduced to surface reflectances: despite the identity between colours and reflectances, some similarities between the colours simply are not evident, predictable, or explicable from any physical description or characterisation of the reflectances. The situation here would be no worse than that facing biochemistry, which exhibits analogous failures in the epistemic reduction of biochemical similarities to chemical similarities.

On the other hand, suppose that interest relative descriptions *are* deemed canonically physical. PRESENT+ might then be true, but the threat posed by the argument from colour

[213] This is referred to as *pharmacophoric analysis*.

similarity is greatly diminished. On this interpretation, the reflectance physicalist can satisfy PRESENT+ by providing some interest relative descriptions of the surface reflectance properties R_b, R_p, and R_g, under which it is evident that R_b is more similar to R_p than R_g. Importantly the reflectance physicalist is no longer restricted solely to *optics*-relative presentations of surface reflectance properties, such as the reflectance spectra for R_b, R_p, and R_g, illustrated in figure eight above. One promising approach, inspired and motivated by the foregoing discussion of biochemical similarities, would be to develop presentations of reflectances that are more suited to the classificatory and predictive interests of *biology*. Philipona and O'Regan (2006: 332) have recently outlined a highly germane approach to classifying reflectance properties which involves

> ...constructing a biological restriction of the physicist's notion of reflectance. From the physicist's viewpoint, reflectance is the relationship that holds between the spectrum of the light illuminating a surface and the spectrum of the light reflected by that surface. To take into account the loss of information induced by photopigments, it is therefore natural for the biologist to study the relationship between the accessible information (instead of the spectrum) about the illuminant and the accessible information (instead of the spectrum) about the reflected light... [W]ith respect to standard sets of natural illuminants, physical surfaces' reflectance induce a linear constraint between the accessible information about the light reflected by that surface. This suggests a biological conception of surface reflection properties as *linear operators* in the three-dimensional space of the accessible information.[214]

It is not my intention here to discuss or evaluate the details of this account. It has been introduced solely to illustrate the potential for a programme of interest relative explanation in which surface reflectance properties are described in ways specifically designed to predict

[214] The 'accessible information' about a light or surface is that information which is accessible through an organism's photopigment set. This need not be limited to the photoreceptors of any particular organism or species. One could instead specify a function from any possible set of photoreceptors to the information about the illumination and surface reflectance that would be made accessible to the visual system of an organism with these photoreceptors. This approach would be more suited to the constitutive approach to understanding colour vision as it would not be tied to the specific features of our actual visual physiology.

their biological role. In essence the programme seeks to describe reflectances as dispositions to make certain information accessible to biological photoreceptors, in contrast to the standard optical characterisation of these properties as dispositions to reflect incident light at all wavelengths in the visible spectrum. The approach is thus similar to Harman's (1996: 255) proposal that colours be specified 'in terms of *biological* mechanisms of colour perception'. A key difference with Harman's account, however, is that on the present approach colours are type identical with surface reflectances (or perhaps productances). The approach is thus biological only insofar as it seeks to provide *interest relative descriptions* of these optical properties that will enable us to predict how they will interact with the contingencies of specific biological mechanisms, paradigmatically biological photoreceptors. As commentators have noted in connection with Harman's account, this target-dependent characterisation of reflectances does not collapse into a dispositionalist, secondary quality, view of colour because no reference is made to the *visual experiences* typically produced by reflectances.[215] Assuming that this programme can be carried out, the reflectance physicalist therefore can be confident that she will be able to satisfy PRESENT+ on the more liberal, interest-relative, reading of 'canonically physical presentations'.

4.6. Conclusions

In this chapter I have presented three new problems for the argument from colour similarity against reflectance physicalism. The first problem concerned the assumption, typical among proponents of the mainstream experiential view of colour vision, that phenomenally apparent colour similarities impose strong constraints on colour ontology. Extending on the argument of the previous chapter in which I dissociated the differential similarity effects characteristic of categorical perception from the constitutive basis for colour vision, I argued that we have good inductive reason to doubt whether *any* phenomenally apparent similarities should be

[215] This is noted by Byrne & Hilbert (1997a: xxiii).

definitive in matters of colour ontology. The second problem concerned the potential for contextual variation in the similarity relation across the premises of the argument. Similarity is a notoriously slippery notion and the onus accordingly is on advocates of the argument to ensure a univocal reading of the similarity relation. The various proposed solutions to this problem seem to raise more questions than they answer, however, about the nature of similarity between the colours. The third problem concerns the presentation sensitivity of our knowledge of similarities. The argument from colour similarity assumes that if reflectance physicalism is true, then any similarities between the colours will be evident from canonical physical descriptions of surface reflectance properties. I argued that this principle is false on most readings, even when genuine similarities are involved. The reflectance physicalist therefore can insist that on her view there is no reason to expect colour similarities – even *genuine* colour similarities, such as they may be – to be evident under canonical physical descriptions of reflectances. There is perhaps one reading on which this epistemic principle is true, if we grant that *interest relative* descriptions class as canonically physical presentations. On this reading the argument loses much of its force, however, for the reflectance physicalist can be confident of finding some interest relative descriptions of surface reflectances which will make evident all visually apparent similarities among the colours. I briefly considered the prospects for a programme of *biological* characterisation of reflectances that could meet this challenge. Further development of this view will provide the basis for future work. In any event, the arguments in this chapter spell the end for the argument from colour similarity and with it the greatest threat to reflectance physicalism.

5. Concluding Remarks

This thesis has ranged over four interlocking issues: colour vision, categorical perception of colour, colour similarity, and colour ontology. Taking centre stage throughout the thesis has been the novel framework for thinking about issues in the philosophy of colour that I call the *constitutive approach* to colour vision. The aim of this approach is to provide an account of colour vision *qua* distinctive psychological kind. The approach thereby seeks to downplay the significance of visual phenomenology in understanding colour vision, which marks a significant departure from the mainstream experiential view of colour vision that predominates in the philosophy of colour. On the constitutive approach visual phenomenology does retain a significant and inalienable role in aiding our initial identification and characterisation of our visual perceptual abilities. In the ultimate account of the *nature* of these abilities, however, specific features of our visual phenomenology play second fiddle to broader considerations deriving from the empirical study of vision and actual or counterfactual dissociations between abilities that are normally indistinguishable via ordinary introspection on our conscious experience of colour.

I have argued that the most plausible account as to the nature of colour vision is that it involves the ability to achieve *colour constancy*. The standard appearance invariance view characterises colour constancy in phenomenological terms as involving the invariance of phenomenally apparent colour across changes in illumination. I developed the alternative *reflectance discrimination view*, on which colour constancy is understood as constitutively involving the ability to discriminate differences in reflectance (or perhaps productance) across changes in illumination. This account employed a broadly cognitive, epistemic, notion of visual discrimination, which requires a subject to possess capacities for the visual

representation of distal features of the environment and also some type of first personal or organism level access to, or registration of, the relevant representations of objective differences. This discriminatory account of colour constancy is surprisingly explanatorily powerful. As we saw in chapter two, for example, the account predicts that it could be advantageous in achieving colour constancy for an organism consciously to represent the illumination conditions in the scene. The account is thus able to accommodate the putatively illumination-dependent aspect of apparent colour that forms the basis for the argument from illumination. In this regard the reflectance discrimination view performs much better than appearance invariantism as an account of the nature of colour constancy.

I then deployed this constitutive account in arguing that the distinctive and vivid feature of our ordinary experience of colour known as *categorical perception* should be philosophically dissociated from colour vision. This argument serves as an illustrative case study as to how the constitutive account of colour vision can be applied to illuminate the distinctive nature of colour vision *qua* psychological kind. The argument provides a clear example of an aspect of the phenomenology of apparent colour that is not strictly related to the constitutive function of colour vision. The argument thus highlights the latent and typically underappreciated complexity of the phenomenology of apparent colour, and lends support to the idea that visual phenomenology is insufficiently discerning to provide a secure route to knowledge as to the distinctive nature of colour vision. This point crystallises the difference between the constitutive approach and the mainstream experiential approach to colour vision.

Throughout the thesis, the issue of colour ontology has been a constant presence, albeit often in the background rather than at the forefront of discussion. The proposed constitutive account of colour vision as the ability to discriminate reflectances provides a clear motivation for *reflectance physicalism*, the view that colours are type identical to

surface spectral reflectance properties (or perhaps productances). The argument dissociating colour vision from categorical perception of colour also supports the view that colour categories such as *red*, *green*, and *blue*, are not strictly speaking colours. Unfortunately limitations of space dictate that I have been unable fully to explore the consequences of this particular brand of reflectance physicalism. I shall take up this topic in future work. My strategy has rather been to take the putative motivation for reflectance physicalism at face value and then to consider how this ontological thesis stands up against the *argument from colour similarity*, which is widely regarded to be its greatest threat. The argument of chapter three provides initial reason to be sceptical as to whether phenomenally apparent similarities among the colours should impose strong constraints on colour ontology. I also presented two independent problems for the argument from colour similarity, based on the extreme context sensitivity of similarity and the presentation sensitivity of our knowledge of similarity. Considered together these issues significantly undermine the force of the argument and thus relieve much of the pressure on reflectance physicalism.

Throughout the thesis I have highlighted various issues arising out of this work that I aim to explore and develop in future research. One such issue concerns the proposal at the end of chapter four regarding a program of *interest relative*, perhaps *biological*, descriptions of surface reflectance properties. To my mind this general approach offers by far the most promising route for explaining the apparently conflicting similarities frequently encountered across all disciplines at different levels of theoretical description. I also hope to extend the discussion of perceptual similarity to other visually perceptible qualities such as shape and size, and to other sense modalities such as audition and olfaction. It seems to me that the complex relationship between constitutive, representational, discriminatory perceptual functions, on the one hand, and perceptual categorisation, structure, organisation, or grouping, on the other, has been insufficiently appreciated and explored. In improving our

understanding of these distinct but closely interrelated psychological abilities, we will undoubtedly gain a more secure grip on the role played by perception in providing us with knowledge about distal features of the world.

One issue of foundational importance that is raised by this thesis, but that is not directly addressed within it, concerns the nature of property perception *per se*. This thesis has provided an account of a particular type of property perception, namely visual *colour* perception. At various points I have also appealed to other types of property perception such as visual shape, size, distance, and motion perception. There remains a lack of clarity, however, surrounding the philosophically fundamental notion of property perception itself. This notion has some inherently puzzling features. To motivate this point, let's assume two widely held views. First, let's assume that properties are *universals*, that is to say, abstract, non-spatiotemporal, acausal, multiply instantiable entities. Second, let's assume that visual perception necessarily involves a causal relation. That is to say, necessarily, a subject *S* visually perceives *x* only if *S* is causally related to *x* (in a non-deviant way). The puzzle is now immediately evident. Given that properties are abstracta and bear no causal relations to anything, how is it possible for anybody to perceive properties? One immediately obvious suggestion is that property perception involves representing or being related in some other way to property instances or *tropes*. Indeed this is the background view operating in this thesis. There remain many important questions, however, regarding the interplay between the metaphysics of property instances or tropes, on the one hand, and our distinctive ability to perceive properties, on the other. What is the best metaphysical account as to how tropes should be individuated? Are determinate property instances identical to determinable property instances? Are both determinate and determinable property instances causally efficacious? And crucially, how do our answers to these questions play into accounts of property perception? For example, would the claim that only determinate property instances

are causally efficacious support the view that colour vision does not constitutively involve the ability to see colour categories such as *red* and *green*? If property perception relates us to particular tropes such as *that shade of red* or *the shape of that apple*, what consequences would this have for the representational contents of perception? Might the content of property perception include a kind of quasi-objectual, quasi-singular, referential device that picks out the particular property instance to which we are related in the world? Is it conceivable that property perception could constitute a more basic form of objective representation than object perception itself? How would this connect with the discussion of perceptual constancy provided in this thesis? These questions highlight a few of the intriguing puzzles raised by the topic of property perception. The evident pregnancy of these issues signifies just how much work remains to be done in advancing understanding of the philosophical foundations of perception.

Bibliography

Abrams, A., Hillis, J., & Brainard, D. (2007) 'The Relation Between Colour Discrimination and Colour Constancy: When Is Optimal Adaptation Task Dependent?', *Neural Computation*, Vol. 19, pp2610-2637.

Adelson, E. H. (2001) 'On Seeing Stuff: The Perception of Materials by Humans and Machines', *Proceedings of the SPIE*, Rogowitz & Pappas (Eds.), Vol. 4299, pp. 1-12, Human Vision and Electronic Imaging VI.

Amano, K., Foster, D. H., Nascimento, S. (2003) 'Red-green colour deficiency and colour constancy under orthogonal-daylight changes', in Mollon, Pokorny, & Knoblauch (Eds.) *Normal and defective colour vision*, Oxford: Oxford University Press, pp. 225–230.

Arend, L. (1994) 'Surface colours, illumination and surface geometry: intrinsic image models of human vision', in Gilchrist (Ed.) *Lightness, brightness, and transparency*, Hillsdale, NJ, England: Lawrence Erlbaum Associates, pp159-213.

Arend, L. & Reeves, A. (1986) 'Simultaneous colour constancy', *Journal of the Optical Society of America A*, 3(10), pp1743–1751.

Armstrong, D. M. (1968) *A Materialist Theory of Mind*, London: Routledge & Kegan Paul.

Armstrong, D. M. (1978) *Universals and Scientific Realism: A Theory of Universals Vol. II*, Cambridge University Press.

Averill, E.W. (1985). 'Colour and the Anthropocentric Problem', in Byrne & Hilbert (1997a), pp11-32.

Bajorath, J. (2002) 'Integration of virtual and high-throughput screening', *Nature Reviews Drug Discovery*, Vol. 1, pp882-894.

Balkenius, A. & Kelber, A. (2004) ‘Colour constancy in diurnal and nocturnal hawkmoths’, *The Journal of Experimental Biology*, Vol. 207, pp3307-3316.

Barsalou, L. W. (1982) ‘Context-independent and context-dependent information in concepts’, *Memory and Cognition*, Vol. 10, No. 1, pp82-93.

Bäuml, K. H. (1999) ‘Simultaneous Colour Constancy: How Surface Colour Perception Varies with the Illuminant’, *Vision Research*, 39: 1531 – 50.

Bayne, T. & Spener, M. (2010) ‘Introspective Humility’, *Philosophical Issues*, Vol. 20, pp1-22

Beauvois, M. F. & Saillant, B. (1985) ‘Optic aphasia for colours and colour agnosia: A distinction between visual and visuo-verbal impairments in the processing of colours’, *Cognitive Neuropsychology*, Vol. 2, No. 1, pp1-48.

Berlin, B. & Kay, P. (1969) *Basic Colour Terms: Their Universality and Evolution*, University of California Press, Berkeley, CA.

Boghossian, P., A., & Velleman, D., J. (1989) ‘Colour as a secondary quality’, in Byrne & Hilbert (1997a), pp81–103.

Boghossian, P., A., & Velleman, D., J. (1991) ‘Physicalist Theories of Colour’, in Byrne & Hilbert (1997a), pp. 105-136.

Bornstein, M. H. (1981) ‘"Human Infant Colour Vision and Colour Perception" Reviewed and Reassessed: A Critique of Werner and Wooten (1979a)’, *Infant Behaviour and Development*, Vol. 4, pp119-150.

Bornstein, M. H. & Korda, N. O. (1984) ‘Discrimination and matching within and between hues measured by reaction times: some implications for categorical perception and levels of information processing’, *Psychological Research*, 46: 207-222.

Boynton, R. M. (1996) ‘History and current status of a physiologically based system of photometry and colourimetry’, *Journal of the Optical Society of America: A*, Vol. 13(8), pp1609-21.

Braddon-Mitchell, D. & Jackson, F. (1996) *The Philosophy of Mind and Cognition*, Oxford: Blackwell Publishers.

Brainard, D. H. (1998) 'Colour constancy in the nearly natural image: 2. Achromatic loci', *Journal of the Optical Society of America A*, Vol. 15, pp 307–325.

Brainard, D. H., Kraft, J. M., and Longere, P. (2003) 'Colour constancy: developing empirical tests of computational models', in Mausfeld & Heyer (Eds.) *Colour Perception: Mind and the Physical World*, Oxford University Press, pp307-334.

Brainard, D. H. & Maloney, L. T. (2011) 'Surface colour perception and equivalent illumination models', *Journal of Vision* 11(5), pp1–18.

Broackes, J. (1992). 'The Autonomy of Colour', in Byrne & Hilbert (1997a), pp191-225.

Broackes, J. (2010). 'What do the colour-blind see?', in Cohen & Matthen (Eds.), *Colour Ontology and Colour Science*, Cambridge, Mass.: MIT Press, pp291-405.

Burge, T. (2005) 'Disjunctivism and Perceptual Psychology', *Philosophical Topics*, Vol. 33 (1), pp1-78

Burge, T. (2009) 'Perceptual Objectivity', *Philosophical Review*, 118 (3), pp285-324.

Burge, T. (2010) *Origins of Objectivity*, Oxford University Press

Byrne, A. (2003) 'Colour and Similarity', *Philosophy and Phenomenological Research*, Vol. 66, No. 3, pp641-665.

Byrne, A., & Hilbert, D. eds. (1997a), *Readings on Colour Vol. 1: The Philosophy of Colour*, Cambridge, Mass.: MIT Press.

Byrne, A., & Hilbert, D. eds. (1997b) *Readings on Colour. Vol. 2: The Science of Colour*, Cambridge, Mass.: MIT Press.

Byrne, A., & Hilbert, D. (1997c) 'Colours and Reflectances', in Byrne & Hilbert (1997a), pp263-288.

Byrne, A. and Hilbert, D. (2003) ‘Colour realism and colour science’, *Behavioural and Brain Sciences*, 26: 3–64.

Byrne, A., & Hilbert, D. (2008) ‘Basic Sensible Qualities and the Structure of Appearance’, *Philosophical Issues*, Vol. 18, pp385-405.

Byrne, A., & Hilbert, D. (2010) ‘How do things look to the colour blind?’, in Cohen & Matthen (Eds.) *Colour Ontology and Colour Science*, Cambridge, Mass.: MIT Press.

Campbell, J. (1993) ‘A Simple View of Colour’, in Byrne & Hilbert (1997a), pp. 177-190.

Campenhausen, V. C. (1986) ‘Photoreceptors, lightness constancy and colour vision’, *Naturwissenschaften* 73: 674 -675.

Carlson, V. R. (1960) ‘Overestimation in Size-Constancy Judgments’, *The American Journal of Psychology*, Vol. 73, No. 2, pp199-213.

Chomsky, N. (1980) *Rules and Representations*, Oxford: Basil Blackwell.

Chomsky, N. (1986) *Knowledge of Language*, New York: Praeger.

Churchland, P. S. (1982) ‘Is the visual system as smart as it looks?’, in Schwartz (2004), pp217-228.

Clark, A. (1996) ‘True theories, false colours’, *Philosophy of Science*, Vol. 63 Supplement Issue, pp143-150.

Cohen, J. (2003) ‘On the Structural Properties of the Colours’, *Australasian Journal of Philosophy*, Vol. 81, pp78 – 95.

Cohen, J. (2008) ‘Colour constancy as counterfactual’, *Australasian Journal of Philosophy*, Vol. 86, No. 1, pp61-92.

Cohen, J. (2009) *The Red and the Real: An Essay on Colour Ontology*, Oxford University Press.

Craven, J. & Foster, D. H. (1992) ‘An operational approach to colour constancy’, *Vision Research*, Vol. 32, No. 7, pp1359-1366.

Davidoff, J. (1997) ‘The neuropsychology of colour’, in Hardin & Maffi (1997), pp118-134.

Decock, L.& van Brakel, J. (2003) ‘Orange laser beams are not illusory : The need for a plurality of “real” colour ontologies’, *Behavioural and Brain Sciences*, 26(1), pp27-28.

Dedrick, D. (2003) ‘Productance physicalism and *a posteriori* necessity’, *Behavioural and Brain Sciences*, 26(1), pp28-.29.

Delahunt, P. B. & Brainard, D. H. (2004) ‘Does Human Colour Constancy Incorporate the Statistical Regularity of Natural Daylight?’ *Journal of Vision*, 4: 57–81.

Denton, M. J. (1999) ‘The Inverted Retina: Maladaptation or re-adaptation?’, *Origins and Design*, Issue 37, Vol. 19.

DeValois, R., L., & DeValois K., K. (1975) ‘Neural Coding of Colour’, in Byrne and Hilbert (1997b), pp93-140.

DeValois, R. L, Abramov, I., & Mead, W. R. (1967) ‘Single cell analysis of wavelength discrimination at the lateral geniculate nucleus in the macaque’, *Journal of Neurophysiology*, Vol. 30, pp415-433.

DeYoe, E. A. & Van Essen, D. C. (1988) ‘Concurrent processing streams in monkey visual cortex’, *Trends in Neurosciences*, 11(5), pp219-226.

Dörr, S. & Neumeyer, C. (2000) ‘Colour constancy in goldfish: the limits’, *Journal of Comparative Physiology*, Vol. 186, pp885-896.

Dretske, F. (2004) ‘Change blindness’. *Philosophical Studies*, 120: 1-18.

Dunn, J. M. (1990) ‘Relevant Predication 2: Intrinsic Properties and Internal Relations’, *Philosophical Studies*, 60: 177-206.

Eddon, M. (2007) ‘Armstrong on Quantities and Resemblance’, *Philosophical Studies*, Vol. 136 (3), pp385–404.

Epstein, W. (1989) ‘”Why do things look as they do?” What Koffka might have said to Gibson, Marr and Rock’, in Schwartz (2004), pp247-260.

Fodor, J. A. (1983) *The Modularity of Mind*, MIT Press.

Foster, D. H., Amano, K., Nascimento, S. (2003) ‘Tritanopic colour constancy under daylight changes?’, in Mollon, Pokorny, & Knoblauch (Eds.) *Normal and defective colour vision*, Oxford: Oxford University Press, pp218–224.

Foster, D. H., Linnell, K. J. (1995) ‘Evidence for relational colour constancy in red-green colour-deficient human observers’, *Journal of Physiology*, 485: 23–24.

Franklin, A. & Davies, I. R. L. (2004) ‘New Evidence for Infant Colour Categories’, *British Journal of Developmental Psychology*, 22: 349-377.

Funkhouser, E. (2006) ‘The Determinable-Determinate Relation’, *Nous*, Vol. 40, No. 3, pp548-569

Gibson, J. J. (1986) *The Ecological Approach to Visual Perception*, Lawrence Erlbaum Associates.

Gilchrist, A. (2006) *Seeing black and white*, Oxford: Oxford University Press.

Goldstein, K. & Gelb, A. (1918) ‘Psychologische Analysen hirnpathologischer Falle auf Grund von Untersuchungen Hirnverletzer’, *Zeitschrift fur die gesamte Neurologie und Psychiatrie*, Vol. 41, pp1–142.

Golinkoff, R. M., Mervis, C., & Hirsh-Pasek, K. (1994) ‘Early object labels: The case for a developmental lexical principles framework’, *Journal of Child Language*, Vol. 21, pp125-55.

Goodale, M. A., & Milner, A. D. (1992) ‘Separate visual pathways for perception and action’, *Trends in Neurosciences*, 15, 20–25.

Goodman, N. (1965) *Fact, fiction, and forecast*. New York: Bobbs-Merrill.

Goodman, N. (1970) ‘Seven strictures on similarity’, in Foster & Swanson (Eds.) *Experience* and *Theory*, Boston: University of Massachusetts Press.

Goodman, N. (1977) *The Structure of Appearance*, (3rd Ed.), D. Reidel Publishing Co.

Gordon, J. and I. Abramov. (2001) ‘Colour Vision’, in Goldstein (Ed.) *The Blackwell Handbook of Perception*, Oxford: United Kingdom, Blackwell, pp92-127.

Hanley, J. R. & Roberson, D. (2011) ‘Categorical perception effects reflect differences in typicality on within-category trials’, *Psychonomic Bulletin & Review*, Vol. 18, No. 2, pp355-363.

Hardin, C., L. (1983) ‘Colours, Normal Observers, and Standard Conditions’, *Journal of Philosophy*, Vol. 80, pp806-13.

Hardin, C., L. (1988a) *Colour for Philosophers: Unweaving the Rainbow*. Indianapolis, MA: Hackett.

Hardin, C., L. (1988b) ‘Phenomenal colours and sorites’, *Nous*, Vol. 22.

Hardin, C., L. (1992) ‘The Virtues of Illusion’, *Philosophical Studies*, Vol. 68, No. 3, pp371-382.

Hardin, C., L. (1997) ‘Colour Quality and Colour Structure’, in Byrne & Hilbert (1997a), pp289-302.

Hardin, C. L., & Maffi, L. eds. (1997) *Colour Categories in Thought and Language*, Cambridge University Press.

Harman, G. (1996) ‘Explaining Objective Colour in Terms of Subjective Reactions’, in Byrne and Hilbert (1997a), pp247-262.

Harnad, S. (1987) *Categorical perception: The Groundwork of Cognition*, New York: Cambridge University Press.

Hatfield, G. (1992) ‘Colour perception and neural encoding: does metameric matching entail a loss of information?’, *PSA 1192*, 1: 492-504.

Hellie, H. (2005) ‘Noise and Perceptual Indiscriminability’, *Mind*, Vol. 114, pp481-508.

Helmholtz, H. V. (1924) *Helmholtz's Treatise of Physiological Optics*, Rochester, NY: Optical Society of America.

Hering, E. (1920/1964). *Grundzüge der Lehre vom Lichtsinn*, translated as *Outlines of a theory of the light sense*, (L.M. Hurvich & D. Jameson, Trans.), Cambridge, MA: Harvard University Press.

Hilbert, D., R. (1987) *Colour and Colour Perception: A Study in Anthropocentric Realism*, Stanford: Centre for the Study of Language and Information.

Hilbert, D., R. (1992) 'What is colour vision?', *Philosophical Studies*, 68: 351-70.

Hilbert, D., R. (2005) 'Colour constancy and the complexity of colour', *Philosophical Topics*, Vol. 33, No. 1, pp141-158.

Hilbert, D., R. (Forthcoming) 'Constancy, content, and inference', available online at http://tigger.uic.edu/~hilbert/papers/constancy_penn-paper.pdf

Hillis, J. & Brainard, D. (2005) 'Do common mechanisms of adaptation mediate colour discrimination and appearance? Uniform backgrounds', *Journal of the Optical Society of America: A*, Vol. 22(10), pp2090-2106.

Hirsh-Pasek, K., Golinkoff, R. M., Hennon, E. A., & McGuire, M. J. (2004) 'Hybrid theories at the frontier of developmental psychology: The emergentist coalition model of word learning as a case in point', in Hall & Waxman (Eds.) *Weaving a lexicon*, Cambridge, MA: MIT Press, pp173-204.

Hollich, G. J., Hirsh-Pasek, K., Golinkoff, R. M., Brand, R. J., Brown, E., Chung, H. L., et al. (2000) 'Breaking the language barrier: An emergentist coalition model for the origins of word learning', *Monographs of the Society for Research in Child Development*, 65 (3), ppi-iv + 1-135.

Hsia, Y., & Graham, C. H. (1965) 'Colour Blindness', in C. H. Graham, ed., Vision and Visual Perception, in Byrne & Hilbert (1997b), pp201-229.

Hubel, D. H., Livingstone, M. S. (1987) 'Psychophysical evidence for separate channels for the perception of form, colour, movement, and depth', *Journal of Neuroscience*, Vol. 7, pp3416-3468.

Hume, D. (1975) *Enquiries Concerning the Human Understanding and Concerning the Principles of Morals*, Oxford University Press

Hurlbert, A.C. (1998) 'Computational Models of Colour Constancy', in Walsh & Kulikowski (1998), pp283-322.

Hurvich, L. M. (1982) 'Chromatic and Achromatic Response Functions', in Byrne & Hilbert (Eds. 1997b), pp67-92.

Jackson, F. (1998) *From Metaphysics to Ethics: A Defence of Conceptual Analysis*, New York: Oxford University Press.

Jackson, F., & Pargetter, R. (1987) 'An Objectivist's Guide to Subjectivism About Colour', in Byrne & Hilbert (1997a) pp67-79.

Jakab, Zoltan & McLaughlin (2003) 'Why not colour physicalism without colour absolutism?', *Behavioural and Brain Sciences*, 26(1), pp34–35.

Jameson, D., & Hurvich, L. M. (1959) 'Perceived colour and its dependence on focal, surrounding, and preceding stimulus variables', *Journal of the Optical Society of America*, Vol. 49, p890.

Jameson, D. and Hurvich, L. M. (1989) 'Essay Concerning Colour Constancy', in Byrne and Hilbert (1997b), 177–98.

Johnston, M. (1992) 'How to Speak of the Colours', *Philosophical Studies,* 68: 221 – 64.

Kalderon, M. (Forthcoming) 'Colour and the problem of perceptual presence', to appear in *Dialectica*.

Kay, P. & Regier, T. (2006) 'Language, thought, and colour: recent developments', *TRENDS in Cognitive Science*, Vol. 10, No. 2.

Kelber, A. & Osorio, D. (2010) 'From spectral information to animal colour vision: experiments and concepts', *Proceedings of the Royal Society B*, Vol. 277, pp1617-1625

Kennedy, C. (2007) ‘Vagueness and grammar: the semantics of relative and absolute gradable adjectives’, *Linguistics and Philosophy*, 30: 1–45.

Kinoshita, M. and Arikawa, K. (2000) ‘Colour constancy of the swallowtail butterfly *Papilio xuthus*’, *Journal of Experimental Biology*, Vol. 203, pp3521-3530.

Kubinyi, H. (1998) ‘Similarity and Dissimilarity: A Medicinal Chemist’s View’, P*erspectives in Drug Discovery and Design*, Vol. 9-11, pp225–252.

Kubinyi, H. (2002) ‘Chemical similarity and biological activities’, *Journal of Brazilian Chemical Society*, vol.13, n.6, pp717-726.

Kuehni, R. G. ‘Olive green or chestnut brown?’, *Behavioural and Brain Sciences*, 26(1), pp35-36.

Kuehni, R. G. (2005) *Colour: An Introduction to Practice and Principles (2nd Edition)*, John Wiley & Sons: New Jersey.

Land, E. H. (1986) ‘Recent advances in retinex theory’, in Byrne & Hilbert (1997b), pp143-160.

Lange, J., (1936) ‘Agnosia and apraxia’, in Brown (Ed.) *Agnosia and Apraxia: Selected Papers of Liepmann, Lange and Poetzl*, Institute for Research in Behavioural Neuroscience, New York.

Lewis, D. (1986) *On the Plurality of Worlds*, Oxford: Blackwell.

Lewis, D. (1997) ‘Naming the Colours’, *Australasian Journal of Philosophy*, 75: 325 – 42.

Livingston, K., Andrews, J. and Harnad, S. (1998) ‘Categorical Perception Effects Induced by Category Learning’, *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 24(3): 732–753.

Livingstone, M. S. & Hubel, D. H. (1987) ‘Psychophysical evidence for separate channels for the perception of form, colour, movement, and depth’, *Journal of Neuroscience*, 7(11), pp3416-3468.

Livingstone, M. S. & Hubel, D. H. (1988) ‘Segregation of form, colour, movement, and depth: Anatomy, physiology, and perception’, *Science*, 240(4853), pp740-749.

Lotto, R. B., & Purves, D. (2000) ‘An empirical explanation of colour contrast’, *Proceedings of the National Academy of Sciences USA*, Vol. 97, pp12834–12839.

MacAdam, D. L. (1985) ‘The physical basis of colour specification’, in Byrne & Hilbert (1997b), pp33-63

McDaniel, C. K. (1972) *Hue perception and hue naming.* Unpublished BA thesis, Harvard University, Cambridge, MA.

McDowell, J. (1985) ‘Values and Secondary Qualities’, in Honderich (Ed.) *Morality and Objectivity: A Tribute to J. L. Mackie*, London: Routledge & Kegan Paul.

McGinn, C. (1983) *The Subjective View*, Oxford: Oxford University Press.

McGinn, M. (1991) ‘Wittgenstein's "Remarks on Colour"’, *Philosophy*, Vol. 66, No. 258, pp. 435-453

McLaughlin, B. (2000) ‘Colours and colour spaces’, in Cubb-Stevens (Ed.) *Proceedings of the Twentieth World Congress Of Philosophy*, Volume 5, Philosophy Documentation Centre.

McLaughlin, B. (2003) ‘Colour, Consciousness, and Colour Consciousness’, in Smith & Jokic (Eds.) *Consciousness: New Philosophical Perspectives*, New York: Oxford University Press, pp97–154.

MacLaurey, R. (1997) ‘Skewing and darkening: dynamics of the *cool* category’, in Hardin & Maffi (1997), pp261-282.

MacLeod, D. (2003) ‘New Dimensions in Colour Perception’, *Trends in Cognitive Sciences*, Vol 7, No 3, pp97-99.

MacNichol, E. F., & Svaetichin, G. (1958) ‘Electrical responses from the isolated retinas of fishes’, *American Journal of Opthamology*, Vol. 46, No. 3, p26.

MacPherson, F. ed. (2010) *The Senses: Classic and Contemporary Philosophical Perspectives*, Oxford University Press.

Maloney, L. T. (1984) ‘Computational approaches to colour constancy’, Unpublished Technical Report, Stanford University.

Maloney, L. T. (2003), ‘Commentaries on Brainard, Kraft, and Longère: Surface Colour perception and its environments’, in Mausfeld & Heyer (Eds) *Colour: Connecting the Mind to the Physical World*, Oxford: Oxford University Press, pp329-330.

Maloney, L. T. & Wandell, B. A. (1986) ‘Colour Constancy: A method for recovering surface reflectance’, *Journal of the Optical Society of America A*, 3: 29-33.

Markman, A. B. & Ross, B. H. (2003) ‘Category use and category learning’, *Psychological Bulletin*, vol. 129, pp. 592–613.

Marr, D. (1982) *Vision: A computational investigation into the human representation and processing of visual information*, New York: Freeman.

Martin, M. G. F. (2004) ‘The Limits of Self-Awareness’, *Philosophical Studies*, Vol. 120, No. 1/3, pp37-89.

Matthen, M. (1988) Biological functions and perceptual content, *Journal of Philosophy*, 85: 5—27

Matthen, M. (1999) ‘The Disunity of Colour’, *Philosophical Review*, 108 (1), pp47-84.

Maund, B. (1995) *Colours: Their Nature and Representation*, New York: Cambridge University Press.

Mausfeld, R. (1998) ‘Colour perception: from Grassman codes to a dual code for object and illumination colours’, in W. Backhaus, R. Kliegl & J. Werner (Eds.), *Colour Vision: Perspectives from Different Disciplines*, De Gruyter, pp219-250.

Mausfeld, R. & Niederée, R. (2003) ‘Can a physicalist notion of colour provide any insight into the nature of colour perception?’, *Behavioural and Brain Sciences*, 26(1), pp41-42.

Maximov, V. (2000) ‘Environmental factors which may have led to the appearance of colour vision’, *Philosophical Transactions of the Royal Society of London B*, Vol. 355, pp1239-1242.

Medin, D. L., Goldstone, R. L., & Gentner, D. (1993) ‘Respects for similarity’, *Psychological Review*, 100(2), 254-278.

Mervis, C. B., & Bertrand, J. (1994) ‘Acquisition of the Novel Name-Nameless Category (N3C) Principle’, *Child Development*, 65 (6), pp1646-62.

Milner, D., & Goodale, M. A. (1995) *The visual brain in action*, New York: Oxford University Press.

Mollon, J. (1997) ‘”…On the Basis of Velocity Clues Alone”: Some Perceptual Themes 1946-1996’, *The Quarterly Journal of experimental Psychology*, Vol. 50A, No. 4, pp859-878.

Morland, A. B., MacDonald, J. H., & Middleton, K. F. (1997) ‘Colour constancy in acquired and congenital colour vision deficiencies’, in Dickenson, Murray, & Carden (Eds.) *John Dalton's colour vision legacy*, London: Taylor & Francis, pp 463-468.

Murphy, G. L. and Medin, D. L. (1985) ‘The Role of Theories in Conceptual Coherence’, *Psychological Review*, 92(3), 289-316.

Nassau, K. (1980) ‘The Causes of Colour’, in Byrne and Hilbert (1997b), pp3–29.

Neumeyer, C. (1981) ‚Chromatic adaptation in the honeybee: successive colour contrast and colour constancy‘, *Journal of Comparative Physiology*, Vol. 144, pp543-553

Neumeyer, C. (1998) ‘Comparative aspects of colour constancy‘, in Walsh & Kulikowski (1998), pp323-351.

Neumeyer, C., Dörr, S., Fritsch, J., Kardelky, C. (2002) ‘Colour constancy in goldfish and man: influence of surround size and lightness’, *Perception*, Vol. 31, pp171-187.

Nijboer, T. (2007) ‘Neuropsychology of Colour Vision: Studies in patients with acquired brain damage, healthy participants, and cases with developmental disorders’, PhD Thesis, University of Utrecht.

Nikolova & Jaworska (2003) ’Approaches to Measure Chemical Similarity - a Review’, *QSAR & Combinatorial Science* 22 (9-10): 1006–1026.

Noë, A. (2004) *Action in Perception*, Cambridge, MA: MIT Press.

Novella, S. (2008) ‘Suboptimal optics: Vision problems on as scars of evolution history’, *Evolutionary Education Outreach*, 1:493-497.

O’Callaghan, C. (2010) ‘Experiencing Speech’, *Philosophical Issues*, Vol. 20, pp305-332.

Palmer, S. E. (1999a) *Vision Science: Photons to Phenomenology*, Cambridge, MA: MIT.

Palmer, S. E. (1999b) ‘Colour, Consciousness, and the Isomorphism Constraint’, *Behavioural and Brain Sciences*, Vol. 22, pp923-989.

Palmer, S. E., Brooks, J., Nelson, R. (2003) ‘When does grouping happen?’, *Acta Psychologica*, Vol. 114, pp311-330.

Patterson, D. E., Cramer, R. D., Ferguson, A. M., Clark, R. D., & Weinberger, L. E. (1996) ‘Neighbourhood Behaviour: A Useful Concept for Validation of “Molecular Diversity” Descriptors’, *Journal of Medicinal Chemistry*, Vol. 39, pp3049-3059.

Pautz, A. (1997) ‘An argument against Armstrong’s analysis of the resemblance of universals’, *Australasian Journal of Philosophy*, Vol. 75, No. 1, pp109-111.

Pautz, A. (2006) ‘Can the Physicalist explain colour structure in terms of colour experience?’, *Australasian Journal of Philosophy*, Vol. 84, No. 4, pp535-564.

Peacocke, C. (1983) *Sense and Content*, Oxford University Press.

Peacocke, C. (1992) *A Study of Concepts*, Cambridge MA.: MIT Press.

Philipona, D.L., & O'Regan, J.K. (2006) 'Colour naming, unique hues, and hue cancellation predicted from singularities in reflection properties', *Visual Neuroscience*, 23(3-4), 331-339.

Pylyshyn, Z. 1999. 'Is vision continuous with cognition? The case for cognitive impenetrability of visual perception', *Behavioural and Brain Sciences*, Vol. 22, pp341-365.

Quine, W. V. O. (1948) 'On What There Is', in J. Kim & E. Sosa (Eds.) *Metaphysics: An Anthology*, Blackwell Publishers, 1999, pp4-12.

Quine, W. V. O. (1969) 'Natural Kinds', in J. Kim & E. Sosa (Eds.) *Metaphysics: An Anthology*, Blackwell Publishers, 1999, pp223-242.

Quine, W. V. O. (1976) 'Grades of Discriminability', *Journal of Philosophy*, Vol. 73, No. 5, pp113-116.

Roberson, D., Davidoff, J. and Braisby , N. (1999) 'Similarity and categorization: neuropsychological evidence for a dissociation in explicit categorization tasks', *Cognition*, 71, pp1–42.

Roberson, D., Hanley, J.R., & Pak, H.S. (2009) 'Thresholds for colour discrimination in English and Korean speakers', *Cognition*, 112, pp482-487.

Robilotto, R., & Zaidi, Q. (2004) 'Limits of lightness identification for real objects under natural viewing conditions', *Journal of Vision*, Vol. 4, pp779-797.

Rock, I. (1985) *The Logic of Perception*, MIT Press.

Rosen, S. & Howell, P. (1987) 'Auditory, Articulatory and Learning Explanations of Categorical Perception in Speech', in Harnad (1987).

Russell, B. (1967) *The Problems of Philosophy*, Oxford University Press.

Rüttiger, L., Mayser, H., Serey, L., Sharpe, L. T. (2001) 'The Colour Constancy of the Red-Green Colour Blind', *Colour Research & Application*, 26: S209–S213.

Schwartz, R. ed. (2004) *Perception*, Blackwell Publishing.

Schwitzgebel, E. (2008) ‘The unreliability of naïve introspection’, *The Philosophical Review*, 117/2: 245-73.

Shea, N. (2012) ‘Methodological Encounters with the Phenomenal Kind’, *Philosophy and Phenomenological Research*, Vol. 84, Issue 2, pp307-344.

Shepard, R. N. (1987) ‘Evolution of a mesh between principles of the mind and regularities of the world’, in John Dupré (Ed.) *Beyond the Best: Essays on Evolution and Optimality*, Cambridge, Mass.: MIT Press.

Shepard, R. N. (1990) ‘A possible evolutionary basis for trichromacy’, in Brill (ed.) *Perceiving, Measuring, and Using Colour. Proceedings of the SPIE/SPSE Symposium on Electronic Imaging: Science and Technology*, 1250: 301-9.

Shoemaker, S. (1975) ‘Phenomenal Similarity’, *Critica,* Vol. 7, No. 20, pp3-34.

Shoemaker, S. (1991) ‘Qualia and Consciousness’, *Mind,* 100: 507–24.

Shoemaker, S. (2003) ‘Content, Character, and Colour’, *Philosophical Issues*, 13: 253-278.

Siegel, S. (2004) ‘Indiscriminability and the phenomenal’, *Philosophical Studies*, Vol. 120, Nos. 1-3, pp91-112.

Singer, J. L. (1952) ‘Personal and environmental determinants of perception in a size constancy experiment’, *Journal of Experimental Psychology*, Vol. 43(6), pp420-427.

Siok, W. T., et al. (2009) ‘Language regions of brain are operative in colour perception’, *Proceedings of the National Academy of Science*, 106(20): 8140-8145.

Sittig, O., (1921) *Stoerungen im Verhalten gegenueber Farben bei Aphasischen*, Monatsschrift fuer Psychiatrie und Neurologie, 49: 63–68.

Smart, J. J. C. (1975) ‘On Some Criticisms of a Physicalist Theory of Colours’, in Byrne & Hilbert (1997a), pp1-10.

Smithson, H. (2005) ‘Sensory, computational and cognitive components of human colour constancy’, *Philosophical Transactions of the Royal Society: Biological Sciences*, 360 (1458): 1329-1346

Sternheim, C. E., and Boynton, R. M. (1966) ‘Uniqueness of perceived hues investigated with a continuous judgemental technique’, *Journal of Experimental Psychology*, Vol 72(5), pp770-776.

Stoerig, P. & Cowey, A. (1997) ‘Blindsight in man and monkey’, *Brain*, Vol. 120, pp535-559

Strawson, P. F. (1979) ‘Perception and its objects’, in MacDonald (Ed.) *Perception and Identity*, London: Macmillan, pp41-60.

Thompson, B. (2006) ‘Colour Constancy and Russellian Representationalism’, *Australasian Journal of Philosophy*, *84*, 75-94.

Thompson, E. (1995a) ‘Colour Vision, Evolution, and Perceptual Content’, *Synthese*, Vol. 104, pp1-32.

Thompson, E. (1995b) *Colour Vision: A Study in Cognitive Science and the Philosophy of Perception*, London and New York: Routledge.

Thompson, E., Palacios, A., Varela, F. J. (1998) ‘Ways of Colouring: Comparative Colour Vision as a Case Study for Cognitive Science’, in Noë & Thompson (Eds.) *Vision and Mind: Selected Readings in the Philosophy of Perception*, MIT Press, pp351-418.

Troost, J. M., & De Weert, C. M. M. (1991) ‘Naming versus matching in colour constancy’, *Perception and Psychophysics*, Vol. 50, No. 6, pp591-602.

Tversky, A. (1977) ‘Features of Similarity’, *Psychological Review*, Vol. 84, No. 4, pp327-352.

Tye, M. (2000) *Colour, Consciousness, and Content*, MIT Press.

Van Essen, D. C., & DeYoe, E. A. (1995) ‘Concurrent processing in the primate visual cortex’, in Gazzaniga (ed.) *The cognitive neurosciences*, Cambridge, Mass.: MIT Press, pp383-400.

Vorobyev, M. (2004) ‘Ecology and evolution of primate colour vision’, *Clinical and Experimental Optometry*, Vol. 87, pp230-238.

Walsh, V. (1999) ‘How does the cortex construct colour?’, *Proceedings of the National Academy of Science,* 96(24): 13594-13596.

Walsh, V. & Kulikowski, J. eds. (1998) *Perceptual constancy: why things look as they do*, Cambridge University Press, Cambridge.

Wandell, B. A. (1989) ‘Colour Constancy and the Natural Image’, in Byrne & Hilbert (1997b), pp161-176.

Werner, A., Menzel, R. and Wehrhahn, C. (1988) ‘Colour constancy in the honeybee’, *Journal of Neuroscience*, Vol. 8, pp156-159.

Whorf, B. L. (1940): 'Science and Linguistics', *Technology Review*, 42(6): 229-31, 247-8, reprinted in Whorf (1956) *Language, Thought and Reality* (ed. J. B. Carroll), Cambridge, MA: MIT Press.

Williamson, T. (1990) *Identity and Discrimination*, Basil Blackwell.

Winawer, J, Witthoft, N, Frank, M ,Wu, L, Wade, A, Boroditsky, L. (2007) ‘The Russian Blues reveal effects of language on colour discrimination’, *Proceedings of the National Academy of Science*, 104(19): 7780-7785.

Wittgenstein, L. (1977) *Remarks on Colour,* Oxford: Blackwell Publishing.

Wooten, B., & Miller, D. (1997) ‘The psychophysics of colour’, in Hardin & Maffi (1997), pp59–88.

Yoshioka, T, Dow, B, Vautin, R. (1996) ‘Neuronal mechanisms of colour categorisation in areas V1, V2 and V4 of macaque monkey visual cortex’, *Behavioural Brain Research,* Vol. 76, pp51-70.

Zaidi, Q. (1998) ‘Identification of illuminant and object colours: heuristic-based algorithms’, *Journal of the Optical Society of America*, Vol. 15, No. 7.

Zaidi, Q. (2001) ‘Colour constancy in a rough world’, *Colour research and application*, Supplement Vol. 26, pp192-200.

Zeki, S. M. (1978) ‘Functional specialisation in the visual cortex of the rhesus monkey’, *Nature*, 274(5670), pp423-428.