

# INTRODUCTION TO THE PHILOSOPHY OF COLOUR

*Derek H. Brown and Fiona Macpherson*

## Why colour?

Why has the examination of many different aspects of colour been a prominent feature in philosophy, to such an extent that the topic is worthy of a handbook? Here are two related answers. First, colours are exceedingly familiar, seemingly simple features that become enigmatic under scrutiny, and they are difficult to capture in any familiar-sounding, unsophisticated theory. Second, through colour one can confront various problems that span the breadth of philosophy, including problems pertaining to perception, the mind-body relation, the nature of science, scepticism, vagueness, meta-ethics, and aesthetics. Let us elaborate.

Colour is, unsurprisingly, central to work in vision and, for better or worse, vision often dominates philosophy of perception. In this way, colour informs our understanding of the nature of perceptual states, the appearance–reality division, and of perceptual phenomenal character (that is, “what it is like” to have perceptual states; see Nagel, 1974). The study of colour has taught us about numerous surprising phenomena, including various kinds of illusions and hallucinations (roughly, cases in which experience fails to accurately represent the way the world is—or isn’t connected to it in the right way, in the case of veridical illusion or hallucination), the potential for perceptual constancies and variations to coexist (roughly, the potential for there to be cases in which we experience a colour to be the same under different illuminations, yet also to somehow look different), and particular associations between perception, cognition, and language. It exemplifies the rich relationship between empirical and philosophical research.

But colour’s reach is much broader than this. Through spectrum inversion scenarios,<sup>1</sup> “Purple Haze”,<sup>2</sup> and of course a certain Mary,<sup>3</sup> colour has anchored debates about physicalism, functionalism, and the explanatory gap. Colour is an ever-ready case study for perceptual variation. The same surfaces can look differently to a person in different conditions, and can look differently to different persons in the same conditions, thus helping fuel sceptical challenges and our accounts of the nature of knowledge. Colour is a stock example of vagueness, which continues to puzzle philosophical logicians. Models for balancing the perspectival nature of colour and the reality of colour have been analogized for use in meta-ethics. Studies suggest that in some respects colours are *out there* in the objective world yet in other respects inherently *perceiver-dependent*, in effect sitting at the intersection of perceiver and environment. This makes colour ontology—attempts to say what and where colours are—highly complicated. This list could continue.

Colour is thus a sparkling lens through which to philosophize. That is “Why colour?”—or at least a quick and dirty answer. A less quick but manageable answer is the rest of this volume. A long, rather unmanageable one, is the now thousands of works published on the topic in philosophy, and many thousands more found outside philosophy. Our focus in this volume is on the philosophical work, aimed at students and professionals inside and outside philosophy who want a picture of where we are at in that philosophical work. (Although we do include two papers that summarize the contemporary science of colour and empirical work on colour spaces.) However, before saying where philosophy of colour is at the moment, it is worth quickly remarking on how we got here.

Philosophers of colour frequently thank Larry Hardin for writing *Color for Philosophers* (1988). We wish to do so again. The work demonstrates, in beautiful detail, numerous ways in which empirical research can and should be brought to bear on basic questions about the nature and epistemology of colour. It stands today as a fine example of how to do empirically informed philosophy of mind, and it was central to what sparked the current interest in philosophy of colour. Although Hardin was unable to contribute to this volume directly, there is little doubt that his influence can be felt throughout.

There was also a broader group of philosophers who made a commitment to philosophy of colour in the last decades of the 20th century—and to great effect. Many of us became aware of this broader movement through Alex Byrne and David Hilbert’s two-volume set *Readings on Color* (1997, MIT). The science volume was a fantastic way for philosophers to introduce themselves to what scientists in the field were doing. The philosophy volume made it plain to all how much first-rate philosophy was being done on the topic. The numerous unresolved challenges and disparate approaches on offer showed many of us that a viable subfield in philosophy had been launched.

Since that time, several thousand philosophical works have been published on colour (as a quick search on PhilPapers reveals). And there is no sign of the field cooling off: several recent collections (for example, Cohen & Matthen 2010; Silva 2017) and monographs (for example, Chirimuuta 2015; Allen 2016; Gert 2017) are sure to stimulate research in the foreseeable future.

Unfortunately, the field of philosophy of colour is now extensive enough to be challenging to newcomers and to established philosophers interested in incorporating colour into their work. The field is even more difficult to penetrate for non-philosophers. We would benefit from a map of the terrain organized around issues of current interest. This Handbook aims to provide one such map. It is not a complete map. But we hope its scope locates the main countries of the field, and its details trace some of the interconnected networks within and between them.

The initial section—*The Importance of Colour to Philosophy*—surveys some broad areas in philosophy to which colour has direct application. This serves as a primer for why many philosophers are interested in colour, and why many more philosophers and non-philosophers should be too. The *Interlude* that follows contains summaries of the science and spaces of colour. These chapters swiftly bring readers to a practical baseline of knowledge on these topics, so that readers can delve into the many chapters that reference or discuss their topics. *Colour Phenomena* examines a suite of issues about colour that have been puzzling philosophers in recent years (and in many cases for much longer). The fundamental question ‘Where are colours and what is their nature?’ is the topic of *Colour Ontology*, introducing readers to a wide array of proposals defended in contemporary research. *Colour Experience and Epistemology* follows colour through the main, current models for perceptual experience and the challenges that notions like ‘looks’ generate in perception. Finally, *Language, Categories, and Thought* confronts readers with issues at the intersection of these three domains.

## **Overview of the sections**

### ***§1 The importance of colour to philosophy***

This section offers a broad sample of the many topics in philosophy that can be profitably studied through colour. Why does colour have such wide philosophical applicability? One way to address the question is to highlight aspects of colour that make it suitable for such applications. These include (but are not limited to) colour's phenomenal character, the persistence of colour variation, and the internal features of colours (that is, their apparent similarities and differences to each other). These aspects of colour have, perhaps unsurprisingly, given rise to a rather sophisticated space of options regarding the nature of colour and regarding the relation between colour and colour experience. We won't work through too many of the details here, but a few remarks on these issues will help readers conceptualize some of the many roles colour occupies through the chapters in this opening (and subsequent) sections.

One important aspect of colour is that it is both susceptible to empirical study, and remains a vivid example of the seeming gap between the world as it is experienced (with its blues and greens, reds and yellows) and the frequently held conception of a colourless physical world of spectral power distributions, light reflectance profiles, and neural pathways. We can think of this as part of the challenge of incorporating phenomenal character, or what we acquire via experience, into our understanding of the physical world. The challenge is non-trivial. We posit colours because we experience them, and we experience them to be in the world around us. Yet physics, chemistry, and neuroscience don't seem to need colours. The problem could be one of fitting *colours* into the physical world, or it could be one of fitting colour *experience* into the physical world, or both. To take but two examples, this issue has immediate impact on philosophy of mind and philosophy of science.

McLaughlin (Chapter 1) argues that both colour and colour experience create problems for physicalism, although only the latter directly impinges the mind-body problem. Part of the challenge colour poses is connected to its status as a secondary quality, which yields difficult questions about the relationship between colours and colour experience. Should we explain colour in terms of colour experience or colour experience in terms of colour? Various options are available; all bring distinct challenges.

From the perspective of philosophy of science, Chirimuuta (Chapter 3) frames the problem of fitting colour and colour experience into the physical world as an important example of the tension between the manifest and scientific image (Sellars 1956). Chirimuuta first considers the possibility that the purported "gap" between the manifest and scientific image is mostly a 20th-century phenomenon. It has earlier roots, to be sure. But there is reason to believe that our routine attribution of this "gap" to earlier thought is inaccurate, something we have foisted on it in retrospect. Interestingly, work on the manifest and scientific image in contemporary philosophy of science continues to receive inspiration from the case of colour. Chirimuuta discusses two cases. In Wilson's (2006) deflationary approach, the many nuances of colour in our speech and actions are used to help deflate or significantly disintegrate the distinction between the objective and the subjective. Giere's (2006) perspectivism, a proposed middle ground in the debate between scientific realism and instrumentalism, is built on the idea that colours can be both real and perspectival (specifiable only relative to a perceiver). For him, scientific assertions are analogous, they are assertions about how the world is from this theory, or using this instrument. There is no unique, unified description of the world that is the "real" one.

Colour's usefulness in philosophy is also due to colour being an enduring example of perceptual variation: the experienced colour of a jacket can change depending on lighting

conditions, one's previous retinal state, and the absorption curves of different perceivers' cones, among other things. And the jacket can arguably remain intrinsically unchanged throughout. This ancient observation applies to many features outside colour, and has been reinforced through centuries of intellectual developments. In the case of colour, it is now underpinned by a wealth of solid empirical data, yielding numerous philosophical applications throughout this Handbook.

One philosophical application of perceptual variation occurs in Pritchard and Ranalli's chapter on scepticism in epistemology (Chapter 2). If, strictly speaking, we experience something variable in colour vision, and the objective world we purportedly inhabit isn't quite so variable, then arguably the world as it appears to us is different from the world as it is independently of us. Finding a way to accommodate variations in appearances with the acquisition of knowledge of an objective world is a central goal in epistemology, and Pritchard and Ranalli illustrate how colour is a profitable means of working through these struggles.

Another striking feature of colours is their robust categories (for example, the collection of all the blues, or of all the purples) and the undeniable respect in which they bleed into one another. This makes colour a staple case in studies of vagueness, the topic of Raffinan's contribution (Chapter 4). The blues merge into the purples, with no discernible point marking the boundary between them. Is the boundary there but inaccessible, leaving us in an epistemically impoverished state? Or is the boundary simply vague, leaving our statements about it susceptible to a host of different semantic analyses? Colour is particularly interesting here because it has been used to test whether or not the rough location of boundaries between its categories are stable when approached from different perspectives. For example, is the location of the boundary between blue and purple the same when approaching the blues from the purples, as opposed to when approaching the purples from the blues? Raffinan presents recent research suggesting that the boundary is direction sensitive.

In studying the logic of colour concepts, Gierlinger and Westphal (Chapter 5) examine the structure of colour propositions and inferences involving them in an effort to define colours (if possible) and, more modestly, to develop accounts of important colour phenomena. Such phenomena include, for example, colour opponencies (for example, red versus green and blue versus yellow), and colour incompatibilities more broadly, topics that re-emerge elsewhere in this Handbook (see below for guidance). Red and green are standardly taken to be incompatible in one or more senses. One can ask whether this incompatibility applies to the colours of things (for example, no thing can be both red and green all over at a time), to experience (for example, no perceiver can experience a reddish-green colour), or to colours themselves (for example, there is no reddish-green colour). When propositions containing these sorts of incompatibilities are regarded as true, what is the strength and nature of their truth and of the knowledge we might have of it? For example, are these propositions analytic, are they necessary, is our knowledge of them a priori? Further, if one isolates an incompatibility or set of incompatibilities that is unique to a colour, then one might use that to define the colour.

It is obvious that colour is significant to aesthetics—painters and dye producers, for example, have been at the leading edge of colour knowledge throughout our history. Yet it is daunting to try and summarize this significance. A good starting point is to focus on the importance of depiction to works of art, as Kulvicki does (Chapter 6). Depiction is roughly representation via 'picturing'. On one influential account, depiction essentially involves some form of resemblance between the depiction and what it depicts. For better or worse, philosophers of perception have often appealed to the painterly attitude in their discussion of art. In brief, the idea is that to produce a realistic painting—the staple form of depiction—the painter often looks at a three-dimensional scene as a two-dimensional artefact to be copied onto her canvas. Setting aside how

accurate this is of artistic practice, the idea is straightforward enough for geometric properties: the shapes of objects in the scene are “collapsed” into their outlines, and the sizes of those outlines are the sizes projected to an image at the canvas surface. What is the analogue for colour? A first thought is that one only need look at the colours of surfaces in the scene, and then find paints that match them to depict them. A moment’s reflection reveals, however, that the colours experienced in a scene are typically impacted not only by surfaces but also by illuminants, shadows, the medium through which one looks, and so on. An accurate two-dimensional depiction of this “collapses” these many elements into one colour at each point on the canvas. How can this be achieved? Is something lost in the process? And what might this teach us about colour vision? These are some of the issues brought forth by this inquiry.

At least some readers will see Gert’s Chapter 7 and wonder: what is the color/value analogy and why should I be interested in it? The aforementioned features of colour are part of the reason why there are numerous rich colour ontologies (accounts of the underlying nature of colour) and why understanding them and their relations is now rather complicated (see below and esp. §IV). As Gert illustrates, several of these colour ontologies have analogues in the space of value ontologies (accounts of the underlying nature of value). For example, there are theses asserting the reality or objectivity of colour/value, the relativity of colour/value, the fundamentally subjective or expressive character of colour/value, the error-ladenness of our claims about colour/value, and so on. On top of this, questions about how to fit colour into the physical world have interesting parallels with analogous questions about values. Work on the colour/value analogy stems from the intuitively plausible idea that the nature of colour can afford insight into the nature of value. The analogies aren’t perfect, and whether or not these differences dilute the usefulness of the analogy is part of the interest in the topic. In the context of this Handbook, this chapter is not merely an invitation for people interested in colour to consider whether views about colour may have valuable lessons for normative theory. It is also an invitation to consider the reverse.

## ***§2 Interlude: the science and spaces of colour***

There is now little question that informed philosophical contributions to colour theory often require basic technical knowledge concerning the science and spaces of colour, and in specific cases requires more than basic knowledge. Our goal has been to find a way to integrate technical colour knowledge into this Handbook that isn’t off-putting to newcomers, but doesn’t pretend that these details can generally be glossed over. Central to our strategy is this section, a basic course in the science and spaces of colour. They are suitable as introductions for new readers, and as reliable points of reference for all readers. Within this Handbook, these chapters allow authors of other chapters to presuppose this basic knowledge and, as is often needed, jump into particulars that are germane to their topic.

Byrne and Hilbert’s chapter on colour science (Chapter 8) provides an overview of the relevant topics in optics, physiology, psychophysics, the visual cortex, visual defects, and animal vision. From the perspective of optics, understanding colour vision requires consideration of numerous factors, including properties of light (especially wavelength), light reflectance properties of surfaces, and light transmission properties of media through which light travels. The light that enters the eye—the colour signal—is a function of these kinds of properties and generates an image on the back of the retina. At this point, physiology becomes crucial. In humans, colour vision utilizes the three different wavelength sensitive cone receptors (see below for exceptions), which send signals to retinal ganglion cells, which in turn send signals to the visual cortex, where a significant amount of colour processing occurs.

Scientists also approach the topic of colour from the perspective of psychophysics, the study of how subjects respond to coloured stimuli. This has led to various insights into types and variations in colour appearances and colour spaces, and notably to the opponent-processing theory of colour vision (though the status of the opponent-process theory is less secure today than it was in preceding decades).

In addition, colour scientists have deepened our understanding of colour through studying defects in colour vision. Many defects arise from abnormalities in cone receptors (yielding different kinds of what is commonly called “colour blindness”, a matter discussed below and in some of the chapters in this volume, especially Chapters 8 and 11). There are also defects due to abnormalities in the visual cortex, notably cerebral achromatopsia and colour agnosia. Very roughly, cerebral achromatopsia occurs when subjects exhibit few or no abilities to visually distinguish chromatic colours, while retaining a comparably sophisticated ability to distinguish visual form (that is, shapes and locations) by detecting differences in luminance. Very roughly, colour agnosia occurs when subjects exhibit sophisticated or even normal abilities to visually distinguish colours and to talk about colours in the abstract (for example, to recall that bananas are yellow), yet limited ability to talk about the colours they see. At present cerebral achromatopsia and colour agnosia are only partially understood phenomena. They are of interest for example for their capacity to provide insight into previously undocumented dissociations between colour and form perception (in the case of cerebral achromatopsia), and colour perception and colour language (in the case of colour agnosia). Finally, colour scientists have made remarkable progress in understanding colour vision by studying non-human animals (more on this below). These are, in broad outline, the topics discussed in Byrne and Hilbert’s colour science chapter.

Briggs’ chapter on colour spaces (Chapter 9) provides an overview of what colour spaces are, the main contemporary models of colour space, and the history of the topic. Colour spaces are representations of colours in some geometric space, where the dimensions of that space are defined in terms of colour-relevant attributes. There are numerous ways to do this (for example, Kuehni and Schwarz (2008) is a recent historical survey that identifies over 170 colour spaces). An important division within colour spaces is that between what one might call “perceptual” and “psychophysical” models. Whereas perceptual models are squarely focused on appearance or perceived colour, psychophysical models map classes of spectral distributions or reflectances that match in appearance under assumptions including a “standard” human observer. We will give a brief example of each.

Perceived colour can be described in terms of the attributes of hue, brightness (or lightness), and colourfulness (or saturation or chroma). Following Hering, it is now common to posit four unique or fundamental hues (blue, green, yellow, and red) for humans. Lightness is perceived position on a scale from black to white (for example, robin egg blue is a light colour, and navy blue a much darker one). Within these constraints it is easy to conceptualize many colour similarity relations (for example, yellows are more similar to oranges than to blues), and identify the fundamental colour opponencies (that is, blue–yellow, red–green, and black–white). In addition, colours possess robust categories (for example, the collection of all the blues, the collection of all the greens) that bleed into one another. Collectively, these features build up a rich and important example of a perceptual model of colour space.

By contrast, a psychophysical model maps colour specifications in terms of tristimulus values (such as CIE XYZ) or psychophysical measures such as dominant wavelength, purity, and luminance. Such a model has wavelength (or light reflectance) as a basic dimension. Since we know that the perceived or appearance colour induced by a stimulus can vary depending on a host of contextual factors, psychophysical models generally make operational assumptions to mitigate

these effects. For example, the perceived colour of an object can vary across illumination conditions. Thus, a psychophysical model that maps colours to surface reflectances might presume a “standard illuminant” so as to capture a mapping between perceived colour and stimulus that holds *if* those stimuli are perceived under a standard illuminant, even if the mapping is distorted when different illuminants are in play. Another important operational assumption might be that of a “standard observer”.

Colour spaces are of interest to philosophers for several reasons. Perceptual models aim to systematize the characteristics of colours themselves, and thus contribute to our understanding of the *explanandum* in theories of colour (for example, Chapter 10 delves more deeply into the hue facet of perceptual models). Psychophysical models illustrate various ways of mapping colour to stimuli, but require important assumptions to mitigate the variations between stimuli and perceived colour that occur in “real” scenarios. The significance of this variation between stimuli and perceived colour is a matter of great debate in philosophy of colour. Very roughly, one might view the presumption of a “standard illuminant” as helping us identify good epistemic conditions for perceiving colour, or one might view it as a means of ignoring crucial facts about the nature of colour. Psychophysical models are thus able to clarify what kinds of assumptions are needed to enhance the mappings between perceived colour and stimuli, providing an important empirical touchstone for these philosophical debates. The history of colour spaces naturally illuminates the development of our thinking about colour, including ways that this thinking is influenced by, and influences, factors of significance to the individual or time period. Thus, one might view the centrality of colour-wavelength mappings to Newton’s colour space as not only illuminating on its own, but as also reflecting his well-known preoccupation with the nature of light. The insights contained in Newton’s colour space are remarkable, but also important is its arguably limited use when considering the colours associated with surfaces as opposed to lights. Finally, one might argue that the mere proliferation of colour spaces contains an important lesson about colour. There are comparatively few perceptual or psychophysical spaces for other domains such as shape, sound, odours, and so on. Our understanding of colour is thus systematizable, but there are glaring constraints on how “unified” the resulting understanding is, making colour an intriguing topic of study. There is thus much of philosophical interest in the study of colour spaces, and Briggs’ chapter provides the requisite background for thinking about these issues.

### **§3 Colour phenomena**

Colour phenomena are what theories of colour perception and colour ontology must account for—they are the explananda for colour theory. There are far more colour phenomena than are treated in this section, but each of the eight phenomena that appear here have received focused philosophical attention in recent years. Although all are relevant to questions in colour perception or colour ontology, the focus of each chapter is not on a particular view about these matters but instead on the phenomenon at issue. Collectively, these chapters illustrate the diversity of phenomena to be explained in a philosophy of colour, and implicitly how challenging it is to build a theory that does justice to them all. It is worth saying a few words about each.

As mentioned above, it is common to recognize four basic hues in human colour vision: blue, green, red, and yellow. Matthen’s topic is *unique* or *pure* hues (Chapter 10), instances of basic hues that contain no trace of other hues. For example, a blue that has no reddishness and no greenishness but only bluishness. Sticking with the example, unique blue is a collection of blues, as such blues can vary in lightness and saturation without inheriting any other hues. Unique colours are first and foremost identified through perceptual experiences; the nature of

stimuli that give rise to these experiences can be subsequently studied. The phenomenon is of interest to philosophers of perception for at least two reasons. Unique hues are examples of unitary or pure perceptual qualities, identifiable through perceptual experience. They thus provide insight into the basic machinery of the (human) colour domain, a foundation from which other colours can be built. Second, we now have good empirical evidence for there being considerable variation in what stimuli induce experiences of a given unique hue across normal human trichromats. (Trichromats have three types of cones cells in the eye that are responsive to different wavelengths—in humans, short, medium, and long wavelengths.) For example, a stimulus that induces an experience of unique green in one perceiver might induce an experience of a strongly bluish-green in another perceiver. Variation in unique hue perception is thus an example of perceptual variation across normal perceivers that sits at the foundations of colour theory.

Recall that using basic hues (blue, green, red, yellow), saturation, lightness, and colour categories we can build a perceptual space for human colours. Are there colours or experiences of colours that are outside this space—are there *novel colours*? The question can be asked of human colour perception and of non-human perception. In the human case a central issue is whether or not a human can perceive a colour that is not localizable within the human colour space or somehow at odds with that space. Macpherson (Chapter 11) discusses reasons to believe in novel colours, including a fascinating case study, namely experiments that purport to induce experiences of reddish-green and bluish-yellow colours in humans. Since in our standard model red and green are colour opposites, experiences of reddish-green colours are by default impossible. Thus, these experiments create a dilemma: either the experiences are somehow misdescribed or the red-green opponency in the standard model is not an in-principle opponency.

Within humans, questions about novel colours also arise from the existence of ‘colour blind’ humans, which Macpherson also discusses. “Colour blind” is a term generally designating any human dichromat, monochromat, or anomalous trichromat. Red-green colour blind people don’t fail to experience colours when looking at red or green things. Instead, roughly speaking, they experience the same colour when looking at red or green things—whatever that colour is. On the traditional picture, when looking at red or green things (things in the “confusion zones”), colour blind people experience one familiar colour (for example, brown or green). But, according to another more recent theory, they experience a colour that is novel to the rest of us. Either way, many people have concluded that since trichromacy is the statistical norm, on this picture the reds and greens that trichromats experience are the actual or real colours of the things at issue, and the colours experienced by the colour blind are abnormal and hence erroneous colours.

Reflection on this picture might prompt one to wonder why the colour a dichromat sees when looking at an apple is the wrong colour—what’s wrong with it? Suppose it is true that such people don’t see starkly different colours when looking at Granny Smith versus Red Delicious apples. In this way their discriminatory capacities are less powerful than those of trichromats. But still, why does that make the colour they see a wrong one, as opposed to simply a different one from those with normal colour vision?

Inspired by the thought that the colour blind’s experience of colour is not incorrect, some people have fundamentally rethought the nature of veridicality in colour perception and, in some instances, the nature of colour blindness itself. For example, there is evidence that dichromats can and do see different colours when looking at various objects in their confusion zones. It is just that they need some time to inspect or explore these objects’ colours, to look at them from different angles, against different backgrounds, and so on. Their deficiency thus might not be the principled one that they can’t see reds and greens (to stick with our example), but instead



the more limited one that it is harder for them to see these colours. If so, there are still difficult questions about what colours such persons experience when a red object and a green object look to have the same colour. But if they can distinguish between reds and greens given the opportunity, then they may be in a better position than we previously thought to tell us about any experience of novel colours that they have.

Synaesthesia, Brogaard's topic (Chapter 12), occurs when people have an extra experience, or element of experience, in response to a certain stimulus. Sometimes synaesthesia is loosely described as consisting of cases of hearing colours or tasting shapes, etc. These cases involve, for example, people experiencing sounds of a certain pitch when they experience different colours, or feeling shapes when they experience different tastes. A person who has synaesthesia can communicate that they have this extra experience in their mental life—often called a *concurrent*—although disagreement remains about the exact nature of that experience, for example, whether the concurrent is somehow attributed to the inducing stimulus or not. An interesting philosophical aspect of synaesthesia is that it involves aspects of perceptual experience being produced in unexpected ways, and it has been intensely studied in recent years.

Interestingly, it turns out that the most common type of synaesthesia isn't cross-modal. It is instead intramodal: experiencing visually perceived graphemes (for example, letters, numbers) as having certain colours—irrespective of the colour of the ink that they are printed in. There are interesting questions about different forms of synaesthesia. In some cases, the concurrent is experienced as being in the world in front of subjects. In other cases it is experienced imaginatively, as being in the "mind's eye". Some concurrents seem to be experienced in response only to perceptual stimuli and their basic perceptual features, while some concurrents are experienced in response to the concept of the relevant inducer, and might be brought on by thought of that inducer as well as perception of it. In the last few years, synaesthesia has grown from an elusive, barely understood phenomenon, to a multi-faceted one that has been discovered to be present in a significant portion of the population.

Philosophers are interested in synaesthesia for several reasons. The triggering of colour experience by non-traditional stimuli is often classified as a form of colour illusion, and the extent and nature of colour illusions is always a crucial topic to philosophy of perception. Beyond this, synaesthesia provides insight into various stable, unexpected pathways for triggering perceptual experience, and, as such, there has been discussion of whether synaesthesia provides a challenge for functionalist or representationalist theories of perceptual experience.

Synaesthesia brings our attention to unexpected associations between experiences of colour and other perceptual qualities. Colour blindness highlights the potential for colour experiences that cannot be communicated between humans with disparate colour visual systems. Spectrum inversion is an extreme case involving both unexpected experiential associations and potentially incommunicable experiences. It is Ross's topic (Chapter 13). Locke wondered whether the colour idea or sensation produced in me by a marigold might be the same as the one produced in you by a violet, noting that we could arguably never detect this difference. In theory, one could take any quality, such as pitches of sounds or degrees of sourness, and imagine some kind of inversion scenario about it, but it is standard to use colour as the focal case. If Locke is right, spectrum inversion would involve an otherwise unexpected swapping, across inverted perceivers, of associations between colour experiences and objects. This would arguably generate a seemingly impenetrable communication barrier between them, guaranteeing the privacy of an individual's colour experience.

Independently of Wittgenstein's invective against private languages, spectrum inversion has variously impacted contemporary philosophy. It poses a challenge for functionalist theories of mind, if two similarly functioning perceivers might nonetheless have inverted colour

experiences. It highlights the tenuous relationship between colour experience and the meanings of colour terms, and makes us question the representational character of said experience. In addition, there are now important arguments in favour of the empirical possibility and indeed probability of spectrum inversion within a small portion of humanity (Nida-Rümelin, 1996)—although such cases have raised vital questions about whether such an inversion could indeed go undetected. For some, the theoretical possibility of spectrum inversion is enough to pose the above philosophical problems, but for those who want a non-zero, or better non-trivial, empirical probability that it obtains, these latter arguments are of great interest.

Although *interspecies* studies in colour vision don't provide evidence for an extreme circumstance like spectrum inversion, Allen demonstrates (Chapter 14) that they do provide evidence for robust variations in the nature of colour visual systems and of uses of colour vision. Among other things, this has helped us to formulate a less human-centric conception of colour, forcing us to: more broadly conceive of what the *functions* of colour vision are for different organisms; recognize substantive differences in colour *perceptions* across species; wonder whether or not various non-human animals have categorically different colour *experiences* from our own; and revisit how we think of colour illusion. Further, by virtue of these items making colour appear highly species specific, they challenge non-relativized, objectivist views about colour.

The above phenomena belong to a long list of phenomena that inform questions about the nature and scope of error or “bad cases” in colour perception. Perceptual error has always been central to philosophy of perception, especially to our understanding of the nature of perceptual experience and to perceptual epistemology. We typically distinguish illusion from hallucination, taking the former to involve misperceiving something before you and the latter to involve experiencing something that (in some sense) is not there at all. Visit *The Illusions Index* ([www.illusionsindex.org](http://www.illusionsindex.org)) to experience numerous relevant examples. Two important philosophical issues concern the *scope* of perceptual error (that is, which perceptions count as erroneous and which do not), and the *nature* of perceptual error (that is, how should we understand the nature of the experiences one undergoes during colour illusions and hallucinations). The nature of error in colour perception is discussed throughout this volume, though most directly in §V. Watkins' contribution (Chapter 15) is largely focused on the scope of colour illusion.

The scope of colour illusion is of particular interest because it has proven so difficult to identify an agreed conception of what colours are, and such an agreed conception would otherwise be used to identify when illusion occurs. By comparison, we have an agreed conception of what shapes are and what the shapes of things in our world are. With this in hand we can examine shape perceptions, and when these deviate from perceived shapes we have a solid basis from which to infer that perceptual illusion occurs. But in the case of colour there is no agreement on what colours are (see §IV), and thus we must find another way to motivate judgements about when colour illusion occurs.

One traditional means of identifying colour illusions is by appeal to colour variation. Due to perceptual variation, the colour some object *O* (or some suitably uniform part of *O*) appears to have varies as a host of perceptual factors change. On the traditional view *O* can't have all these colours in itself. On that view, due to the aforementioned colour incompatibilities, *O* can have at most one colour in itself (that is, if *O* is green all over at some time then it isn't also any other colour at that time). It follows that either *O* doesn't have colour in itself, or all but one of *O*'s apparent colours are illusory.

Traditionally, advocates of the subjectivity or mind-dependence of colour opt for the first disjunct, and advocates of the objectivity or mind-independence of colour opt for the second. One difficulty for objectivists is that colour variation is now recognized as being very pervasive. It is found for example within any reasonable construal of normal or ideal perceptual conditions.

This makes it difficult to identify *which* apparent colour is *O*'s colour, and more deeply undermines the credibility of the commitment to *O* having a unique colour in itself. In recent years, objectivists have developed various lines of response. One is to question the extent of *colour incompatibility*: why did we think that *O* (or the relevant uniform part of *O*) can only have one colour in itself? Perhaps *O* can have one human colour and at the same time have a different pigeon colour. Another is to question the idea that a given colour can only correctly manifest or display one colour appearance. Perhaps blue things can sometimes look green, without error being present—blue can simply appear green in some contexts. These and other strategies are developed in various chapters in this Handbook. The aim of Watkins' chapter is not so much to defend colour objectivism as it is to demonstrate the various means objectivists have available to them to minimize the ascription of colour illusion in the face of colour variation. Watkins' more general goal is to dissuade us from thinking that we can or should aim to settle which colour perceptions are illusory first, and proceed to reach conclusions about colour ontology from there.

In the final chapter in this section, Brown's (Chapter 16) focus is not on colour variation but colour constancy. Very roughly, colour constancy is the tendency of a colour visual system to register a stable colour in a surface across variations in illumination conditions that surface is under. This is important in part because it can serve as a counter to colour variation: certainly illumination conditions change and, to some extent, colour perception changes with it, but there is an underlying sameness in colour perception that remains throughout. Although the details matter (and they are complicated), in this rough way, colour constancy can counter-balance some of the pressures variation places on objectivist theories of colour, and interplay with the above-mentioned connections between variation and illusion.

There are related reasons for interest in colour constancy. The phenomenon is found in many non-human species, making it a central facet of colour vision. Its existence implies that colour vision isn't only concerned with analysing incoming light, but is also concerned with isolating aspects of one's environment that remain constant across light changes. This helps one shift from thinking of colour vision as focused on the proximal signal to a system that also has interest in distal things—in the surface qualities of objects and therefore in those objects themselves. Because of this one might think that colour constancy helps explain *why* we see colours on rocks and trees (and not as aspects of the light arrays striking our eyes), and why we experience colours as illumination-independent properties (if in fact we do).

Finally, difficult questions about the nature of colour perception, experience, and judgement arise when trying to "fit" colour constancy with colour variation. Consider a simplified example. When, by hypothesis, we experience a wall as differently illuminated and stably coloured, does the experienced illumination variation constitute a difference in experienced colour? In at least some respects it seems that it does: the wall looks darker in colour in the shadowed part and lighter in colour in the unshadowed part. But then in what way is the wall experienced as stably coloured—how can something look to be both the same colour and differently coloured? Perhaps some aspects of the case are reflections of colour *experience*, and others are reflections of *judgements* made about that experience. Alternatively, perhaps we can construct a more sophisticated account of colour experience that resolves the tension between the overlapping apparent colour difference and sameness.

Colour constancy has thus emerged as a fruitful phenomenon: for defending the objectivity of colour against pressures from colour variation; for helping us conceive of colour perception as targeting distal things in our environment, as opposed to light arrays striking our retinas; and for providing us with a difficult and basic case study situated at the intersection of colour experience and judgement.

## §4 Colour ontology

What are colours exactly, and where precisely are they? These are arguably the guiding questions for colour ontologists. They are daunting ones, as should be apparent given our discussion thus far about the numerous ways colour is used in philosophy, the rich knowledge we are acquiring about the science and spaces of colour, and the variety of colour phenomena that need explanation. Our aim in organizing this section was to identify a cluster of colour ontologies that accurately represent those ontologies that people hold and discuss in contemporary philosophy. We will present them by locating each within a broader landscape of colour ontologies. Unfortunately, some locations in this landscape do not have dedicated chapters in this volume—we had to be selective—but we provide references to sources for these which the interested reader can pursue. To keep this introduction manageable, our remarks aim to differentiate various colour ontologies but *not* to motivate them. Readers can learn about those motivations by consulting the relevant chapters.

As just stated, the overarching tasks of colour ontology are to determine what kinds of properties colours are and where we can find them in our world (if anywhere). These tasks should be clearly differentiated from the related ones of determining what causes colour experiences and determining what causes some object to have a given colour, although determining these things may be helpful in carrying out the overarching task. The issue is also not to determine what makes some experience a colour experience or what makes a perception one of colour. In colour ontology, the focus is on colours themselves. That said, adopting an account of colour experience, colour perception, the causal antecedents of colour, and of course of colour illusion and hallucination, can impact one's colour ontology—and vice versa.

There are numerous ways to parse the space of colour ontology. Our parsing centres on two broad and orthogonal dimensions, the subjectivist-objectivist dimension and the reductionist-non-reductionist dimension. The former concerns the extent to which colours are dependent on perceivers (and thus subjective) or independent of perceivers (and thus objective). The latter roughly concerns the extent to which colours are physically reducible or have non-physical aspects (and are therefore non-reductive).

Regarding the *subjectivist-objectivist* dimension, the broad options are that colours are entirely independent of perceivers (objectivism), that colours are only dependent on perceivers (mentalism), and that they are dependent on both perceivers and the world outside of them (relationalism).

What would it mean for colours to be *objective* (Chapters 17 and 18)? With respect to the debate in colour ontology, what is meant is that the nature of colour does not require appeal to perceivers at all. Thus, bananas are yellow and the nature of that yellowness does not depend on any facts about any perceivers. One version of this theory, defended by Byrne and Hilbert (Chapter 17), draws on contemporary colour science and identifies a particular shade of yellowness with a class or disjunction of surface spectral reflectances (SSRs). Call that class  $C_I$ . The SSR of a surface is the percentage of light that it reflects of each wavelength of light. We know that objects that look to normal human perceivers in normal conditions to have the same shade of colour can have various different SSRs. The disjunction of those SSRs,  $C_I$ , causes experiences of that particular shade of yellow in normal perceivers in our world (as opposed to other possible worlds where it may not), and that shade may bear other interesting relations to the perception of yellow. Furthermore, it may be that it is only via experiences of yellow (in this world) that we are able to identify  $C_I$  as the property yellow. Nonetheless, what objectivists insist on is that  $C_I$ 's nature, that is the nature of that shade of yellow, does not depend on any of these relations (for objectivists, contingent relations) to perceptual states— $C_I$  is a wholly

objective property. It can be specified in wholly objective terms—in terms of SSRs—without mention of the nature of its effects on perceivers.

In contrast to the notion of objectivity just outlined, there is an alternative notion of what it is for colours to be objective that one might have taken to be plausible: colours are objective if and only if they are properties of things that persist independently of perceivers. What is crucial for objectivity on this view is that bananas could have been yellow before perceivers were around and could maintain their yellowness were perceivers to disappear. This is a weaker account of objectivity, for it is consistent with the nature of yellowness requiring appeal to perceptual states. For example, yellowness might be defined as the power to induce perceptions of yellow in perceivers, regardless of whether or not colour perceivers are around to help bananas exercise that power. Even though this kind of Lockean ontology satisfies one intuitive notion of ‘objective’, in colour ontology discussions it is typically regarded as a perceiver-dependent or subjectivist one. In fact it is considered a form of *colour relationalism*. This is because what it is for something to be a colour is defined in terms of its relation to perceivers of a certain sort (in terms of its power to affect them)—albeit a sort of perceiver that may not exist at a given time or place.

There is another feature of colour relationalism that is worth considerable attention. Colour relationalists maintain that colours minimally depend on both perceivers and the perceived. A classic example is the Lockean dispositionalist just mentioned, for whom yellow is a power to produce experiences of yellow in perceivers. However, in most contemporary literature colour relationalists typically maintain that the colour of a thing, say our banana, is a complex relational property with relata that include not only particular perceivers and the banana, but also the lighting conditions, neighbouring colours, and perhaps other features of the environment. Such a view is the topic of Cohen’s contribution (Chapter 19). According to it, colour depends not only on perceiver and perceived, but also on perceptual conditions. When speaking precisely, these contemporary relationalists won’t say that a banana has a colour *simpliciter*, but instead that a given banana has some colour for some perceiver in some viewing condition, where various facets of the perceiver and viewing condition need specification to fully explicate the relevant colour. Despite this difference between the views, both Lockean and contemporary relationalist views assert that colours are relational as opposed to intrinsic properties. And, crucially, the kind of relational property colours are is the kind that depend on the nature of perceptual states. The nature of colour for relationalists therefore depends both on environmental and perceptual factors.

We previously saw that there are two notions of “objective”, and that the stronger, rather than weaker, one was what was that employed by those holding colour objectivism. The situation is similar with respect to “relational”. Relationalists about colour hold that colour is a relational property that makes reference to both perceivers and the colour-relevant facets of their environment. However, there is a much looser sense of “relational” which might be applied to colours, according to which colours are relational as opposed to intrinsic properties *simpliciter*. That is, their nature involves specification of some form of relation—but it needn’t be a relation between perceivers and their environment. This is a different idea and isn’t typically what is meant by “colour relationalism”. One example of such a view would be the identification of colours with SSRs, for SSRs are relational properties. They are properties specifying the ways in which a given surface reflects different wavelengths of light. The relation in question is one that relates parts of the environment to each other: lightwaves to surfaces. One who identifies colours with classes of SSRs is a colour relationalist in this looser sense, but is definitely *not* a colour relationalist in the first sense. Similarly, one might identify colours with colour sensations or colour qualia. Qualia are private mental items posited to explain the seemingly

subjective character of consciousness. Qualia are typically conceived of as mind-dependent features. Thus, someone who believes that colours are a kind of qualia would arguably hold that colours are mind-dependent. One would thereby be a colour relationalist in the looser sense, but again wouldn't be in the first more relevant sense. The relation employed here does not involve the environment of the perceiver. We stick to current convention and by "colour relationalism" mean the view that colours are properties whose nature depends on both environmental and perceptual factors.

In contrast to both colour objectivism and relationalism, the colour *mental*ist takes colours to be properties whose nature depends only on perceptual states, narrowly construed to include only states whose nature can be specified by saying what is going on within the head or within mental states. Colour mentalism is the focus of Robinson's contribution (Chapter 21). Mentalists might take colours to be features of neural states or to be a subset of mental qualia, and it is reasonable to regard those who take colours to be properties of sense-data to be mentalists in this sense (sense-data are mind-dependent perceptual objects discussed below and in detail in Chapter 24). Mentalists agree with their rivals that environmental factors like SSRs and relations between surfaces, illumination conditions and the like, are critical to understanding when a perceiver will likely enter into a given perceptual colour state. These environmental factors are thus also critical to understanding when a given colour will likely be instantiated in a world. However, these environmental factors are not regarded as part of the nature of colour, but only as part of what typically causes colours to be instantiated (in the minds or brains of perceivers). With this rough understanding of the subjectivist-objectivist dimension of colour ontologies, we proceed to the reductionist-non-reductionist one.

As the subjective-objective and reductionist-non-reductionist dimensions are orthogonal, each of objectivism, relationalism, and mentalism can be developed in a *reductionist* or *non-reductionist* manner. The issue, as it manifests itself in colour studies, concerns the extent to which colour properties are physical. (In different domains within philosophy, what counts as a successful reduction differs. For example, in moral philosophy reduction of goodness or badness to mental states, counts as a success. In aesthetics, reduction of beauty to colour and form would be a success. In philosophy of mind, reduction to physical states is what counts.)

Let us start with reductivism. One influential form of colour objectivism, which we have encountered already, is a reductionist view. It is the reductionist view that identifies surface colours with classes of SSRs and colours more broadly with ways that things manipulate or produce light (Chapter 17). Contrasting with this, though not a popular view, mentalism in a reductive form identifies colours with properties of neural states. Different again is a relationalist reductionist view. In order for a relationalist to be a reductionist they must hold that all of the relata that make up a colour relation are physical. To take a simple example, if blue is the power to produce experiences of blue in perceivers, then both the power and the experience must be physical.

While physicalism is a dominant force in contemporary analytic philosophy, we mentioned at the outset that colour has been a persistent source for resistance by the non-physicalist. Sometimes the resistance stems more directly from considerations about the ability to naturalize colour experience qua experience, rather than from considerations about colour ontology, but not always. There are a variety of non-reductive ontologies. *Colour primitivism* is the thesis, defended by Gert (Chapter 18), that colours are wholly objective properties that are *sui generis*. According to primitivists, colours bear important relations to physical properties like SSRs, perhaps in some sense supervening on them. But colours themselves are not identifiable with anything like SSRs. Instead they are non-physical, objective properties of things in our environment. Here are two straightforward examples of non-reductivist views that are forms of

mentalism (Robinson's topic in Chapter 21). Those who take colours to be properties of sense-data are non-reductionist colour mentalists, at least to the extent that sense-data are taken to be non-physical and mind-dependent (as they usually are). Similarly, people who identify colours with colour qualia will be non-reductionists, to the extent that they take qualia to be non-physical properties (as they often are). Finally, a straightforward example of a relationalist position that is non-reductionist is the view that that colours are powers to produce colour experiences in perceivers, and that colour experiences are non-physical states.

In addition to chapters on the preceding views, the other two ontology chapters in this volume are titled "Eliminativism" and "Monism and Pluralism". Both headings are potentially ambiguous so it is worth stating how they should be understood in this volume. On one reading—that which is the focus of Wright's contribution (Chapter 22)—colour eliminativism asserts that nothing in one's *surrounding environment* has colour. Bananas aren't yellow or any other colour, nor are trees green, postboxes red, or blueberries blue. Such a view doesn't entail colour mentalism, as it is strictly the negative thesis that nothing in one's surrounding environment is coloured. However, it is consistent with mentalism, since one's mental states (again narrowly construed) do not extend into one's surrounding environment.

A stronger reading of 'eliminativism' asserts that *nothing in our world* has colour, not bananas and blueberries, but also not sense-data, neural states, or mental states. On this view, colours are not the kind of property one finds instances of in our world at all. Advocates of this stronger form of eliminativism are obliged to give some kind of positive story to explain our various colour commitments. One proposal might be that colours are useful fictions (Gatzia, 2010), and another that colours are properties found in some other more perfect world such as Eden (Chalmers, 2006)—though in theory these proposals could be consistent with one another. Either way, nothing in *our* world has colour. There is unfortunately no chapter in this volume dedicated to this stronger eliminativism.

The colour pluralist believes that there are multiple colours, in some more interesting way than that there are blues and greens. However, there are different ways in which this might be the case. One could for example believe that there is a sense in which colours are mental properties and a sense in which they are objective properties, yet see no credible means of collapsing colour ontology into one or the other. So why not accept both? This idea has been defended by some authors (Maund 1981; Brown 2006), and while it could be described as a kind of colour pluralism, it is more usefully referred to as the *dual-referent view* or the *colour ambiguity view*. There is unfortunately also no chapter dedicated to this idea. The phrase "colour pluralism" is typically reserved for a different and, in a way, more radical idea.

The more radical idea is that the same part of the same object simultaneously has multiple colours. A ripe banana may be yellow all over but it is also other colours all over at the same time. This is the colour pluralism Kalderon discusses (Chapter 20). It is *prima facie* in tension with colour incompatibilities, like the purported truism that nothing can be black and white and red all over. Under scrutiny, which incompatibilities a given colour pluralism excludes depends on which colours that pluralism allows to be co-instantiated. Unrestricted pluralism recognizes no genuine colour incompatibilities (for example, everything can have every colour). More restricted pluralisms might recognize incompatibilities between say the colours that humans experience (for example, something can't be red and purple all over), but no such incompatibilities between the colours that humans experience and those that pigeons experience, presuming that those are different (for example, something can be "human orange" all over and have some other "pigeon colour" all over as well).

One form of colour pluralism stems from colour relationalism. A given object can, in theory, enter into multiple colour relations with perceivers in different perceptual conditions. Since the

relationalist holds that these relations *are* colours, if we assume that there is no principled means of singling out one of these scenarios as uniquely correct, a given object will have multiple colours. Via this reasoning, contemporary relationalists tend towards being unrestricted colour pluralists. By contrast, traditional Lockean dispositionalists typically do not have this commitment. For them, the perceptual relation that obtains in “normal” conditions is often presumed unique and uniquely correct—although this is an optional feature of their view.

There can also be pluralist versions of objectivism about colour. Recall that according to objectivism colours are not perceiver-dependent, but instead classes of SSRs or *sui generis* objective properties, or *somesuch*. Either way, different colours can be possessed by a surface at a given time. How is this explained? If colours are classes of SSRs, then the explanation proceeds by pointing out that a given SSR can be a member of various colour classes. Thus, a surface with a given SSR can have many different colours. If colours are *sui generis* properties, then the explanation stems from the account what these *sui generis* properties are like. One can certainly develop an account according to which a surface can have multiple such properties.

### **§5 Colour experience and epistemology**

Recall that part of the philosophical challenge of colour stems from its status as a secondary quality, which in part means that it is difficult to formulate accounts of colour independently of accounts of colour experience, and vice versa. Because of this, experience has an elevated role in colour theory, compared for example, to its role in shape theory. The aims of this section are to introduce readers to how colour relates to three dominant approaches to perceptual experience, and to provide readers with a very brief introduction to “looks” statements involving colour terms.

We begin with a brief outline of the three accounts of perceptual experience and subsequently apply them to colour. The accounts are representationalism, naïve or direct realism, and indirect realism (for example, sense-datum theory). There are numerous ways of developing these accounts, and some controversies over how to characterize them. We nonetheless hope that the following sketch, which leaves out many subtleties, is informative.

Pautz (Chapter 23) discusses the currently dominant representational approach to perceptual experience. Its central tenet is that perceptual experiences are inherently representational states, which is to say that they express “contents” which may, to varying extents, be satisfied by (or be accurate of) the relevant part of world before the subject. For example, a common experience might be one that represents that there is a blue square in front of me. The content of that experience is that there is a blue square in front of me. This content would be accurate if there was indeed a blue square in front of me and would be inaccurate if there was not. Because perceptual contents can diverge from the world before the subject, the two are distinct kinds of things, and in many cases it is helpful to think of contents as abstract entities (for example, propositions). Within this framework, the relationship between the phenomenal character of perceptual experience and perceptual content is a contentious issue: phenomenal character might be determined by perceptual content, or vice versa, or both; phenomenal character might outstrip perceptual content, or vice versa, or both; phenomenal character might be identical to perceptual content; and so on. We won’t discuss the nuances of these debates, but it is fair to say that the representational approach is committed to *explaining* phenomenal character in terms of perceptual contents, in some substantive sense of ‘explain’. In Pautz’s stronger formulation, “experiencing is *nothing but* representing” (p. 366).

Representational approaches can be contrasted with acquaintance approaches. The latter assert that perceptual experience constitutively involves those things that a subject experiences,



for example objects, properties (or property instances), relations, or, most commonly, some combination thereof. Crucially, since the perceptual experience constitutively involves the relevant entity, there is no opportunity for the experience to be variously accurate of, or satisfied by, that entity. That is, experiential states conceived in this way make no appeal to truth-conditions or accuracy-conditions in the way that contents do, and thus they are not representational.

The two acquaintance-based chapters in this Handbook differ on what they take perceivers to be acquainted with. According to naïve or direct realism, the topic of Campbell's contribution (Chapter 25), in perception, subjects are acquainted with the objective reality before them. Thus, perceptual experience constitutively involves objective or mind-independent entities in our world. To have a perceptual experience is to be *open to* the world before you. By contrast, the indirect realist views discussed by Maund (Chapter 24) hold that in perception subjects are acquainted with mind-dependent entities like sense-data. Thus, perceptual experience constitutively involves sense-data. Experiences of these mind-dependent entities are typically taken to generate the possibility of indirect perceptions of the objective world. Before introducing more machinery, it is worth making three clarifications.

First, it is fairly straightforward that whereas direct realism is constructed to secure direct perception of the objective world, indirect realism is constructed to secure at most indirect perception of the objective world. One might ask: should the representational approach be understood as a model that permits direct perception of the objective world, or as a model that permits only indirect perception of the objective world? This is a contentious issue, in part, but not only because there are controversies surrounding what counts as direct or indirect perception. It is nonetheless fair to say that representationalists have sought to avoid asserting that we can only indirectly perceive the objective world. This may be because representationalism permits direct perception of the objective world, or because it is a theory in which the question of how directly we can perceive the world doesn't apply in any straightforward sense. Detractors roughly argue that, because contents are essential to the representationalist account of perceptual experience, and by their nature contents are distinct from the objective world, perceptions of the objective world are in some sense "mediated" by contents. We won't wade into these waters more deeply than this.

Second, is the indirect realist approach a representationalist view? While it is true that on this view perceivers are acquainted with mind-dependent entities like sense-data, this question can nonetheless arise in a number of ways. Might perceivers also perceptually represent sense-data? Might sense-data represent the objective world? Perhaps most crucially, aren't the indirect perceptions perceivers have of the objective world representational states (perceivers certainly aren't acquainted with the objective world on this account)? It is a well-known historical fact that indirect realist views were originally called "representative realist" views, a matter that caused confusion when what we're calling "representationalism" came to the fore. Our own view is that this is a complicated question (see for example Maund's chapter in this volume and Macpherson 2014), and that on many, though not all, formulations of indirect realism perceptual representation is utilized in some way or other. Let us move on.

Third, there is much value to hybrid views proposing that perceptual experience involves both acquaintance with *and* representation of the objective world. However, since there is no consensus on the landscape for such views (two sources are Brown 2012 and Schellenberg 2014) there is no chapter in this Handbook dedicated to the topic.

Given the complications inherent in representationalism and indirect realism, and the simplicity of the naïve realist thesis that perceptual experience is constituted by being acquainted with the objective reality before the subject, why not be a naïve realist? Matters are not this

straightforward because we have yet to focus on how these accounts deal with perceptual illusions and hallucinations, otherwise known as the “bad” cases in perception. Indeed representationalism and indirect realism were developed in part with an eye to giving a compelling account of (a) what is happening in bad cases and (b) why there seem to be similarities between the bad and good (veridical perceptual) cases. Regarding (a), there are numerous kinds of perceptual illusions and hallucinations that need explanation, something that can be safely inferred from our discussion to this point. Regarding (b), in theory a hallucination of a blue ball can be subjectively indistinguishable from a perception of a blue ball, and illusions in which one experiences a blue ball to be purple can be indistinguishable from experiences of a purple ball as purple. Here is an over-simplified guide of how to incorporate these issues into our three accounts of perceptual experience.

According to representationalism, perceptual experience is explained in terms of perceptual contents. Thus, perceptual illusions and hallucinations are states whose contents are to some degree or other not satisfied by the world before the subject—or it is simply luck that they are. Similarities between good and bad perceptual cases are explained in terms of similarities in perceptual contents. Suppose one is having an experience as of a blue ball. This means that one is perceptually expressing a content such as <there is blue ball>. If the state is hallucinatory then that content is either not satisfied by the world before the subject—or it is merely luck that it is (to allow for the possibility of veridical hallucination). If the state is illusory then at least part of the content is not satisfied by the world before the subject (for example, a purple ball is before the subject)—or it is merely luck that it is so satisfied (to allow for the possibility of veridical illusions). If the state is a case of veridical perception (a good case) then the whole content is satisfied by the world before the subject, and that content is connected up to the world in the right way to rule out luck playing a role in the content being satisfied. In our example, there is a blue ball before the subject and the blue ball is connected to the content in the right way.

The indirect realist model is similar, except perceptual experiences are explained in terms of sense-data instead of perceptual contents. Thus, perceptual illusions and hallucinations involve being acquainted with sense-data whose features to some degree or other do not “match” the objective world before the subject—or are not connected up to it in the right way. Subjective similarities between good and bad perceptual cases are explained in terms of similarities between the sense-data experienced in these cases. Suppose again that one is having an experience as of a blue ball. It follows that one is acquainted with a blue, round sense-datum. If the state is hallucinatory, then the world before the subject contains no blue ball—or it is merely luck that it is (to allow for the possibility of veridical hallucination). If the state is illusory, then there is a partial match between the sense-data and the world (for example, a purple ball is before the subject)—or it is merely luck that it is so satisfied (to allow for the possibility of veridical illusions). If the state is purely veridical (a good case), then there is a “complete match” between the sense-data and the world before the subject, and the sense-data are connected up to the world in the right way to rule out luck playing a role in the content being satisfied. In our example, there is a blue ball before the subject, and one’s round, blue sense-datum is appropriately related to the world.

The naïve realist picture is very different. Perceptual experiences are constituted by acquaintance with the objective reality before the subject. Since bad cases (perceptual illusions and hallucinations) involve failures to pick up on the objective reality before the subject, the naïve realist regards these as fundamentally different kinds of cases from good cases. This is why naïve realism is also called “disjunctivism”: one is either in a good case or a bad case, and these are very different kinds of states, despite there being subjective similarities between them. Disjunctivists

disagree about how to characterize bad cases, and about how to characterize the subjective similarities between good and bad cases. However, they are often engaged in a collective effort to minimize the number of bad cases, for example arguing that some purported cases of hallucinations are cognitive instead of perceptual, and that some purported cases of illusion involve not perceptual error but awareness of things like objective appearance properties. These important details, however, can be set aside.

What does all of this have to do with colour? As should be plain to this point, colour is a fantastic case study for theories of perceptual experience. Colour is a vivid, familiar phenomenon. It has been studied intensely, yielding a rich stock of explananda. And colour experience plays a substantive role in debates about colour ontology, in part owing to the difficulties many have had with defending colour objectivism. In theory, there are numerous colour ontologies that one could attach to each theory of experience. In practice, there are natural or familiar pairings that are emphasized by each author. With that in mind, here is a brief remark on what readers can find in each chapter in this section.

After outlining representationalism, Pautz (Chapter 23) considers and assesses various colour ontologies that could be expressed within representationalism. These notably include relationalist (what he calls “response-dependent”), objectivist (what he calls “response-independent”), and eliminativist (what he calls “irrealist”) ontologies. Maund (Chapter 24) discusses various motives for indirect realism, numerous ways of developing the view, and focuses on applying indirect realism to two colour ontologies: mentalism and relationalism. Campbell (Chapter 25) matches naïve realism with primitive objectivism about colour—this arguably being the most natural ontology for naïve realists. From this perspective he discusses various perceptual phenomena, including the transparency of experience, perceptual colour space, imagination *de re*, Mary the colour scientist, and spectrum inversion.

The final chapter in this section, Breckenridge’s contribution (Chapter 26), concerns the perplexing topic of the looks of colours. It is straightforward that the difference between “I see x to be y” and “x looks y” has been important to philosophy of perception. There is a difference between seeing something to be blue and something looking blue. Indeed this difference is often associated with the riddles surrounding colour variation. It is also well known that it is difficult to interpret “x looks y”. What precisely does one mean when one says that something looks blue? On one influential analysis, there are comparative, epistemic and phenomenal uses of “looks” statements. On comparative uses, to say that “x looks y” is to say that x looks as y’s do (for example, that thing looks as cellphones look, or that thing looks as blue things do). On epistemic uses, “x looks y” means roughly that one is acquiring evidence (or justification), by looking, that something is some way (for example, it looks like that thing is a cellphone, or is a blue thing). On phenomenal uses, “x looks y” means roughly that x is presenting a y appearance (for example, that thing is presenting a cellphone appearance, or a blue appearance). These different senses of “x looks y” are fairly clear, but there are other analyses available, most of which distinguish between two or more senses of “x looks y”. Breckenridge discusses the many analyses that have been offered, and proposes a theory according to which all uses of “x looks y” can be reduced to a single, new sense of “looks” introduced in the chapter. Readers are invited to read Breckenridge’s chapter for detail.

## **§6 Language, categories, and thought**

This section considers select topics about the vast number of relations between colour on one hand, and language, categorization, and thought on the other. There is one essay each on colour language, colour categorization, and cognitive penetration. Here is a brief summary.

Much work has been done outside of philosophy—in anthropology, literature, history, and psychology—on the cultural variations of colour language. One question is whether there are such variations or not. Another is whether either of the answers that could be given to the question could or should be explained by the physiology of colour vision, colour experience, culture, thought, language alone, all of these, some of these, or none.

This is the topic that Dedrick (Chapter 27) turns his attention to. Given that the different shades of colour form a visual continuum, and given that we divide that continuum into different sets of colours—colours like red and orange (“colour2s” as Dedrick labels them), as opposed to very specific shades of colour (“colour1s”), like the very different particular shades of red that there are—Dedrick asks, “where do the divisions originate? [And w]hat is the basis for our classifications?” (p. 437 of his chapter). Colour1s can be ordered in a natural way using judgements of similarity and difference to yield a colour space. But are colour2 divisions in any way natural? Dedrick explains that Quine thinks not: they are arbitrary and depend on language and culture. But this has proven to be a very contentious answer that has been greatly disputed over the last 50 years or so. Dedrick lays out the arguments for and against the naturalness of the divisions of colour2s. Some of this debate turns on the question of whether language and words on the one hand, or ideas and concepts on the other, of colour2s (presumably based on common perceptual saliences) are more basic. And some of the debate turns on the highly related question of to what extent the divisions of, or names for, colour2s vary across cultures. The greater the lack of variation across cultures, the more reason there is to think that the divisions of colours2 are not arbitrary.

The state of the debate now seems to suggest that there is a subtle interplay between colour language, colour concepts, human culture, and colour vision on our subsequent perceptual experience and/or perceptual judgements involving colour categories. This debate is further examined by Briscoe (Chapter 28), who focuses on the several decades of psychological research on the exact nature of colour categorical perception. He discusses in detail the notions of perceptual salience and linguistic relevance, and what work they might do to explain human categorization behaviour. And, finally, he examines whether colour perception is indeed categorical.

Lastly, the question of whether beliefs and desires can have an effect on the nature of colour perception and/or experience is covered in an essay on cognitive penetration—a re-emerging topic in philosophy of perception and in cognitive psychology in which colour plays a central role. Stokes (Chapter 29), explains why this debate is so important. It concerns the functional and modular organization of the mind, the role of experience in justifying beliefs, and the nature of perceptual experience. As Stokes emphasizes, debates about cognitive penetration sometimes concern what exactly the nature of the phenomenon is that is called “cognitive penetration”, and that debate often alerts us to which features of the mind are most interesting or important to us.

## **Omissions**

Despite the length of this volume, and for all of the ground that it covers, we were not able to include essays on all of the topics that might have proved interesting or useful. As previously stated, we have not included a separate chapter on colour blindness, however, several of the other chapters discuss it and Chapters 8 and 11 cover it in some detail. We have no essays that cover colour in philosophical history, before the 20th century. We have no chapter that covers traditional colour dispositionalism in the same sort of detail as some of the other topics in colour ontology, although we hope that this introduction, and the other essays in the ontology section, given the reader enough of an introduction to that topic. Sadly, we have no essay on the

development of colour vision. This is quite a neglected topic in philosophy and psychology, and that in part speaks to our inability to cover it. No doubt there are other areas of research on colour that will not have received as much attention as their importance deserves. However, we hope that this volume will sufficiently inspire the philosophical study of colour, and motivate people to further explore this rich and exciting area of study.

## Notes

- 1 Spectrum inversion scenarios (Chapter 13) are ones that philosophers imagine in which someone experiences colours on the opposite side of the colour wheel to another person when looking at the same objects.
- 2 “Purple Haze” is representative of conscious experiences of colour that Joe Levine (2001) thinks cannot be given an explanation.
- 3 Mary is philosopher Frank Jackson’s (1982) imagined neuroscientist who knows all the physical facts but, as she has experienced only black and white, allegedly does not know all the facts about colour.

## References

- Allen, K. 2016. *A Naïve Realist Theory of Colour*. Oxford University Press: Oxford, UK.
- Brown, D. H. 2006. On the Dual Referent Approach to Colour Theory. *Philosophical Quarterly* 56: 96–113.
- Brown, D. H. 2012. Losing Grip on the World: From Illusion to Sense-Data. In A. Raftopoulos & P. Machamer, eds., *Perception, Realism and the Problem of Reference*, pp. 68–95. Cambridge: Cambridge University Press.
- Byrne, A., & D. Hilbert, eds. 1997a. *Readings in Color, vol. 1: The Philosophy of Color*. MIT Press: Cambridge, Mass.
- Byrne, A., & D. Hilbert, eds. 1997b. *Readings in Color, vol. 2: The Science of Color*. MIT Press: Cambridge, Mass.
- Chalmers, D. 2006. Perception and the Fall from Eden. In T. Gendler & J. Hawthorne, eds., *Perceptual Experience*, pp. 49–125. Oxford University Press: Oxford, UK.
- Chirimuuta, M. 2015. *Outside Color: Perceptual Science and the Puzzle of Color in Philosophy*. MIT Press: Cambridge, Mass.
- Cohen, J., & M. Matthen, eds. 2010. *Color Ontology and Color Science*. MIT Press: Cambridge, Mass.
- Gatzia, D. 2010. Colour Fictionalism. *Rivista Di Estetica* 1(43): 109–23.
- Gert, J. 2017. *Primitive Colors*. Oxford University Press: Oxford, UK.
- Giere, R. 2006. *Scientific Perspectivism*. University of Chicago Press: Chicago, USA.
- Hardin, C. L. 1988. *Color for Philosophers, expanded edition*. Hackett: Indianapolis, ID.
- Jackson, F. 1982. Epiphenomenal Qualia. *Philosophical Quarterly* 32: 127–36.
- Kuehni, R., & A. Schwarz. 2008. *Color Ordered: A Survey of Color Order Systems from Antiquity to the Present*. Oxford University Press: Oxford, UK.
- Levine, J. 2001. *Purple Haze: The Puzzle of Consciousness*. Oxford University Press: Oxford, UK.
- Macpherson, F. 2014. Is the Sense-Data Theory a Representationalist Theory? *Ratio* 27(4): 369–92.
- Maund, B. 1981. Colour—A Case for Conceptual Fission. *Australasian Journal of Philosophy* 59(3): 308–22.
- Nagel, T. 1974. What Is it Like to Be a Bat? *The Philosophical Review* 83(4): 435–50.
- Nida-Rümelin, M. 1996. Pseudonormal Vision: An Actual Case of Qualia Inversion? *Philosophical Studies* 82(2): 145–57.
- Schellenberg, S. 2014. The Relational and Representational Character of Perceptual Experience. In B. Brogaard, ed. *Does Perception Have Content?*, pp. 199–219. Oxford University Press: New York, NY.
- Sellars, W. 1956. Empiricism and the Philosophy of Mind. In H. Feigl & M. Scriven, eds., *Minnesota Studies in the Philosophy of Science, vol. I*, pp. 253–329. University of Minnesota Press: Minneapolis, MN.
- Silva, M. ed. 2017. *How Colors Matter to Philosophy*. Cham: Springer.
- Wilson, M. 2006. *Wandering Significance: An Essay on Conceptual Behavior*. Oxford University Press: Oxford, UK.