In Derek H. Brown and Fiona Macpherson (eds.) The Routledge Handbook of Philosophy of Colour, Oxford: Routledge, 2021, pp. 269 - 284.

16
COLOUR CONSTANCY

Derek H. Brown

At first pass, colour constancy occurs when one sees a thing in one's environment to have a stable colour despite differences in the way it is illuminated at a time (simultaneous constancy) or over time (successive constancy). The phenomenon is intuitively grounded in everyday experiences in which something is partly shadowed but, in some sense, looks to be uniformly coloured (simultaneous constancy), and in which one views something indoors then outdoors and the thing looks to be stably coloured (successive constancy). The concept is established in vision science, where it traditionally refers to any efforts visual systems take to separate surface and incident illuminant contributions to the colour signal (i.e. the light reaching the eye) and stably represent the surface contributions. However, one can broaden the concept in various ways (e.g. Brown 2003), for example to include efforts visual systems take to stably represent colour across changes in viewing angle, the medium through which one is viewing, and other aspects of scene composition. I largely adhere to narrower conceptions, but comment on broader ones where appropriate.

After a brief introduction to the colour constancy concept (§0) and the science of colour constancy (§1), my focus is on the significance of colour constancy for two intertwined philosophical issues. The first is colour ontology, where constancy has been used to argue for the objectivity of colour, and in particular for a reductive form of it (see §2). The second is colour constancy's complicated connection to colour experience and colour epistemology. Colour constancy is a subtle phenomenon: it is situated at the intersection of perceptual experience and judgement; it is influenced by myriad forces within our visual-cognitive systems; and is likely a composite of interestingly disparate phenomena. I approach this suite of issues from the perspective of the *given* in colour perception (§3). As will become plain, the ontological and epistemic issues are related in important ways. This does not, however, detract from the value of focusing on each individually.

§0 What is colour constancy?

Here is a brief sample of various ways to develop the rough idea that colour constancy occurs when one sees a thing in one's environment to have a stable colour despite differences in the way it is illuminated over or at a time. Note that the *thing* in question is typically taken to be a material surface in one's environment (though see §0.4 below).

§0.1 What is the relevant sense of 'sees'? Some one or more of 'visually represents', 'experiences', 'is acquainted with', or 'judges' may be appropriate instead of or in addition to a success-verb notion of 'sees'. If colour constancy is a perceptual phenomenon then it arguably cannot merely arise through judgement. By default, I will assume it is perceptual, and that 'visually represents' and 'experiences' are both acceptable terms for describing the relation between perceiver and the target stable colour. Typically, I default to 'experiences'. This issue is central to (§3).

§0.2 What is meant in saying that the target thing's colour is experientially 'stable' across the illumination variations? Is the experienced colour (a) exactly the same, or (b) of the same colour category (say blue) but perhaps of a different shade within that category (say darker blue)? Alternatively, is colour constancy (c) the ability to experience stable relations between colours across illumination variations, regardless of stabilities of experienced colour itself (Foster 2003)?

Option (a) has intuitive appeal, but it is difficult to defend. Instead, colour constancy is typically regarded as a *tendency* to experience a thing as having a stable colour despite the presence of illumination variations, a fact sometimes summarized by referring to colour constancy as *approximate* (e.g. Bradley & Tye 2001, Byrne & Hilbert 2003). This implies that some degree of colour sameness (at least categorical sameness) is experienced during constancy, but typically not *exact* sameness. The latter is usually associated with *perfect* colour constancy, a rare occurrence (see also §1, esp. note 1). As we will see, there are means of trying to preserve (a), but I will in general assume that colour constancy involves (b), that the degree of sameness that is experienced during colour constancy at least stays within a colour category. I briefly remark on (c) in §3, but otherwise leave this intriguing option in the background.

§0.3 The target object is differently illuminated over or at a time. What contribution do these illumination variations make to colour experience? If perfect constancy were the norm, then one could assume that the illuminant contribution to the colour signal is generally *discarded* and hence neither experienced nor represented during colour constancy. All else being equal this would result in colour constancy experiences only involving experiences of perfectly stable surface colours. This was perhaps the hope of early constancy theorists, arguably an assumption in Land's pioneering work on the topic (e.g. Land & McCann 1971), and central to early idealized computational models (e.g. Wandell 1989). However, given that imperfect constancy is now recognized as the default kind of colour constancy, it is generally understood that various aspects of the illuminant contribution, though often not all, are experienced and represented in colour vision. As pretheoretic evidence think of the above examples, of shadows, scattering effects, the visual difference between seeing the world in daylight versus twilight, et cetera. Two qualifications are important.

First, it remains to explain the *manner* in which illumination enters into colour experience during colour constancy. For example, is it a modifier of the experienced surface colour (and if so what form of modification takes place), or is illumination experienced to have its own colour? This question will return throughout. Second, the idea that colour constancy involves discounting or minimizing the experiential impact of illumination changes is not bankrupt. For example, low level light adaptation mechanisms exist and to some extent *do* discount illumination variations. Intuitively, one experiences this form of adaptation when moving from dimly (/brightly) illuminated areas to brightly (/dimly) illuminated ones: the initially sharp contrast in experienced illumination is to some degree normalized by our visual systems. This thus remains a viable form of colour constancy.

One challenge is to build a colour constancy model that incorporates both the partial discounting of illumination variations via mechanisms like adaptation, and the vivid experiential impact of illumination variations familiar from shadowing and the like. As a first step, adaptive effects typically apply to changes in *overall* illumination (i.e. illuminant changes that bound toward spanning one's field of view). This is entirely compatible with more local illumination variations due to partial shadowing, scattering, et cetera being left largely perceptible under changes in overall illumination. In this vein, Wright (2013) has usefully suggested that we recognize two fundamentally different kinds of colour constancy. I will be focusing on the constancies in which illumination variations are experienced (in some manner or other). This is in part because it follows standard practice and in part because such cases are the most illuminating and challenging to work with.

§0.4 As mentioned at the outset, it is conceptually straightforward to broaden our understanding of colour constancy beyond "experienced stabilities in surface colour across illumination variations". First, one can simply invert the traditional model and ask about experienced stabilities in *illumination* conditions across changes in *surfaces*. Second, in theory a surface can be experienced to have a fairly stable colour not only across illumination variations, but also for example across variations in the medium through which one is looking (air versus water; one pair of sunglasses versus another pair; etc.), and across variations in surfaces surrounding a target surface. Surely these modest feats are "approximately" achieved in human colour vision, there is no reason to disregard them as insignificant, and it is likely that the same or similar mechanisms are involved in them as are involved in securing approximate surface colour stabilities across illumination variations (see again, e.g. Brown 2003). To capture these generalizations, one might conceive of colour constancy as efforts visuals systems make to disentangle and track stable contributions to the colour signal (be them stable surfaces, illuminants, or whatever) as other contributions vary.

To be sure, colour constancy *purists* will resist such deviations from the traditional "experienced stabilities in surface colour across illumination variations" picture, though from my perspective there are few to no philosophically important reasons for doing so. I will nonetheless focus on the traditional conception, as this contains enough to chew on for unfamiliar readers. With luck, what follows will generate some interest in exploring these generalizations of colour constancy in future work.

§1 Colour constancy science

The science of colour constancy has developed immensely in the last 30 years. The details are complicated, and I can only provide a broad, selective outline.[1] First, it is important to note that some degree of colour constancy has been observed in many species, including not only in various land animals but also in birds, insects, and water creatures. The phenomenon is therefore widespread and seems critical to understanding colour perception. Neurally, there is evidence of processing relevant to colour constancy occurring in the retina, in V1 and in V4. The emerging picture is thus that colour constancy isn't achieved by a single mechanism at a distinctive neural location, but instead that different facets of constancy are achieved by different mechanisms at differing levels of neural processing (see also §0.3). In many respects we are at an early stage of research with regard to how these pieces fit together. Of most relevance to our discussion is the psychophysical evidence drawn from human subjects.

There are four main types of psychophysical experiments, those involving: asymmetric colour matching; achromatic adjustment; colour naming; and surface and illuminant change attribution (see Foster 2011, 681–90). Before remarking on them, it is worth highlighting that within each type are tokens that vary in numerous ways. For example, the experiment can test for simultaneous or successive constancy, or it can utilize stimuli that are briefly presented (adaptation minimal) or stimuli that presented for an extended duration (adaptation significant). The nature of the stimuli can also vary greatly. Stimuli can be depicted on computer monitors or actually presented to subjects. Stimuli can be presentations/depictions of natural or artificial kinds, of recognizable things or unfamiliar geometric structures, et cetera. The most influential "measure" of colour constancy is the Colour Constancy Index (CCI), which provides a decent measure of the extent to which what a subject identifies as matching surfaces, despite being differently illuminated, deviates from a perfect physical match (introduced in Arend et al. 1991, p. 665). The index ranges from 0–1, with 1 representing perfect colour constancy, and 0 representing the opposite (i.e. something like perfect constancy failure or perfect colour variation—there is no standardized term).

The asymmetric colour matching paradigm is arguably the most influential. It involves two stimuli either simultaneously or sequentially presented under different illuminants (i.e. simulated on a monitor). The stimuli are traditionally collections of abstract colour patches (*Mondrians*) though have also been more natural scenes. Subjects are asked to adjust a patch on the test stimulus (*test patch*) so that it matches a corresponding patch on the target stimulus (*target patch*). Since the test and target patches are presented under different illuminants, a difference that is apparent in experience, the task is non-trivial. Interestingly, the task is subject to significant instructional effects. An influential design (Arend & Reeves 1986) involves two separate instructions:

> *Appearance Match Condition*: adjust the test patch to match the target patch's hue and saturation.
>
> *Surface Match Condition*: adjust the test patch so that it looks like it is cut from the same piece of paper as the target patch.

While Surface Match responses contain evidence of approximate colour constancy (often ranging from 0.6–0.8 CCI), Appearance Match responses contain evidence of substantive colour change (often ranging from 0.2–0.4 CCI).[2] This has led to difficult interpretational questions (see §2 and esp. §3).

Like the asymmetric colour matching paradigm, the surface and illuminant change attribution experiments typically present to subjects (either simultaneously or sequentially) Mondrian stimuli under different illuminants. However, a crucial facet of these experiments is that subjects are asked to assess whether a given colour-perceptual change is due to a difference in illumination or surface conditions. This task can be paired with Appearance and/or Surface Match instructions and the results compared. Perhaps not surprisingly, changes identified as being due to surface conditions correlate well with Surface Match results, and are linearly separable from Appearance Match results (Reeves *et al.* 2008).[3]

§2 Colour constancy and ontology: the path to colour objectivism and beyond

One can naturally seek justification for the objectivity or mind-independence of colours in folk psychology ('spinach is green') or in naïve experience (green looks to be *of* or *in* spinach and not for example in or part of 'the perceiver'). Colour constancy is a tempting additional anchor for colour objectivists, especially those seeking a more naturalized source of justification. The idea that colour experience is to a reasonable degree invariant across illumination conditions is an instance of something being stable as some variable changes values. This makes colour ripe for study through controlled experimentation, setting at least the tone for a naturalized colour epistemology. Ontologically, if experienced colour is fairly invariant across illumination changes, then, assuming colour experience is broadly veridical, colour is an illumination-independent property. Further, if colour is experienced as a property of things in one's environment (often on/in their surfaces), then, again assuming colour experience is broadly veridical, colour is an illumination-independent property of things in one's environment.

The step from here to colours being mind-independent is tempting. After all, light gets to directly strike things in one's environment all the time, certainly more directly than vision does. Thus, if colour is an illumination-independent feature of those things, it is independent of us. Also tempting is the reductive and more precise postulate that the nature of colour is critically tied to the ways surfaces reflect light (their SSRs) and more generally to the ways things interact with light. SSRs are fairly stable features of things, and are directly relevant to the causal chain from light source to surface to eye that gives rise to colour perception. Colour constancy is therefore a naturalistically respectable anchor for colour objectivism and in particular a reductive form of it like reflectance physicalism (Chapter 17).

Although there are various means of formalizing this reasoning, the following simple rendering is useful for our purposes:

1 *Colour Constancy:* experienced(/represented) colour is fairly stable across illumination variations and is experienced(/represented) as a feature of things in one's environment.
2 *Veridicality:* our colour experiences(/representations) are generally veridical.

Thus,

3 *Illumination-independence:* colour is an illumination-independent property of things in one's environment.

And more precisely,

4 *Reflectance physicalism*: the ways objects reflect light, their SSRs, are plausible candidates for colours.

The simplicity and elegance of the argument are impressive.[4] For centuries colour subjectivists have used colour variation, illusion, and hallucination as a basecamp from which to defend their view. If nothing else this straightforward argument provides a much-needed counterweight, grounded in a phenomenon—colour constancy—that has roots in common sense perception and is naturalistically respectable. The argument can thus serve as a basecamp from which to defend colour objectivism, in this case reflectance physicalism.

§2.1 Objections and replies

Many challenges to reflectance physicalism have been put forward that do not focus on the role of constancy in what motivates the view (see again Chapter 17). Others focus precisely on it. Cohen (2008) contends that colour constancy data do not support the illumination-independence of colour in any straightforward way, an idea discussed in §3. Chirimuuta (2008) argues roughly that "ontologically neutral" experimentation on colour constancy is unavailable. Instead, constancy data that seem to support reflectance physicalism do so because antecedent theoretical assumptions to that effect are built into the experimental approach, and similarly for different assumptions/experimental approaches. Though these are excellent contributions, I focus on evaluating the above argument.

To begin, the two premises contain important qualifiers: experienced colour is *fairly* stable and colour experience is *generally* veridical. The premises thus provide at best defeasible justification for (3) and (4). Much hinges on how to interpret the impact of illumination variations on experienced colour and on the veridicality of colour experience. For example, an opponent might argue that, via (1), illumination variations are often experienced to at least some degree, and that (2) the general veridicality of colour experience implies that colour is to some degree illumination dependent. The defender of the above argument thus needs to justifiably "contain" illumination's impact on colour ontology.

One strategy is to prevent Veridicality from applying to colour experiences of illuminants on grounds that such experiences are rare or anomalous. This is roughly the strategy of Byrne and Hilbert (2003): we sometimes experience light to be blue, say that of a blue laser, but these are exceptional situations and the blueness we experience is illusory. They propose that the experienced blue is actually the blue of the light source, not of the light beam itself. However, even if this strategy has merit in some cases, it is difficult to generalize given the widespread presence of shadowing and scattering effects and the fact that lights drastically change the 'tones' of experienced colours in everyday scenes (e.g. more bluish tones at sunrise and more reddish at twilight). Illuminants impact colour experience in seemingly every scene. Thus, if Veridicality is in force, colours are routinely illumination dependent and so colours cannot be confined to SSRs.

A different strategy concedes that illuminants regularly impact colour experience, but maintains that they do so in such a way as to preserve illumination-independent colour. This requires developing an interesting account of colour experience, the latter being the topic of §3. In this context it is worth rehearsing how the strategy has been used to help bolster the argument for colour objectivism. As a start, one could offer the plausible idea that illuminants *illuminate* colours in various ways, and thereby treat experienced illuminants as an additional dimension in colour experience beyond the traditional three of hue, saturation, and lightness (Hilbert 2005, 2011; Jagnow 2010). Thus, when one experiences a given surface colour one experiences the colour's hue, saturation, and lightness, plus a manner of illumination, where the latter is distinct from the former. How should we understand the *way* in which illuminants enter into colour experience? One option is to suppose, as is plausible, that illuminants can and often do serve an epistemic role, making surface colour more or less visible (e.g. very visible in bright light versus barely visible in low light). Suppose in addition that, controversially, this epistemic role is "marked" in experience. One can explicate this "mark" in numerous ways, but for simplicity (setting aside obvious confounds) assume it is through the vivacity or faintness of the experienced colour, where a vivacious colour is indicative of its being highly visible, and a faint colour is indicative of its being poorly visible. On this account, the only colours typically experienced are surface colours, and thus Veridicality can perform its original function of securing them. In addition, Veridicality can be extended to the new machinery. Thus, experiencing the degree of vivacity/faintness of a colour is both experiencing the level of illumination of the colour and experiencing (in some sense of the term) the degree of epistemic access one has to that colour.

A third strategy combines the first two, allowing illumination to enter colour experience in two distinct ways. Sometimes illumination is experienced as coloured (e.g. laser case)—these experiences are non-veridical. Other times illumination makes surface colours more or less visible by making us experience those colours more or less vividly. In theory illumination can serve both functions in a given perception, generating an illusory colour perception at one location in the scene, and modifying the vivacity of the experienced colours at other locations. Perhaps both functions can even be served by the same sample of light. Suppose a blue laser strikes a wall. The experienced blueness of the laser could be illusory. At the same time the surface colour "behind" the laser might nonetheless be faintly experienced and thus to some degree veridically experienced. This hybrid strategy can thus be a powerful view.

One worry about the hybrid strategy stems from the thought that illuminants regularly—not rarely—impact experienced colour itself. Illuminants often directly impact experienced hue, and arguably saturation and lightness. As mentioned above, sunrise often generates more bluish tones throughout the scene, and twilight more reddish ones. These illuminant colours can arguably vary in richness or saturation (e.g. an intense red twilight versus a pale red one), and can be lighter or darker (e.g. a light red sunset versus a somewhat darker one). More scientifically, recall that in the Appearance Match condition subjects are instructed to adjust the test patch to match the hue and saturation of the target. It is hard not to read this as an instruction to match colour in some basic sense. Yet the reports deviate significantly from surface colour constancy (CCI = 0.2–0.4). Presumably this is because the illumination variation impacts experienced colour proper, not merely how "vividly" colours are experienced.

The problem for the hybrid strategy is straightforward. If illuminants only have occasional impact on experienced colour proper, then that impact can be classified as non-veridical by appeal to something like statistical irregularity. But if illuminants have widespread impact on experienced colour proper, then we need a different justification to classify that impact as non-veridical. The truth of reflectance physicalism would provide such justification, were reflectance physicalism not what we are trying justify in this context. Therefore, if studying colour constancy teaches us that illuminants have widespread impact on experienced colour proper, then constancy suggests that colours depend on illuminants (Cohen 2008). Fortunately, there is another route.

§2.2 A more liberal "dual-colour" alternative

A simple way forward supposes that we experience both surfaces *and* illuminants to be coloured. If these two colour types are distinctly experienced, then when the illuminant (/surface) colour varies and the surface (/illuminant) colour remains stable, and one experiences precisely that, one experiences a constant colour while also experiencing a variable one. A natural model for developing the thought is in terms of colour layering (Brown 2014), and the rough idea is defended by Mausfeld (2003), Tokunaga and Logvinenko (2010), and Davies (2016). This approach can accommodate the ideas from the account just discussed. Illuminant colours can be viewed as ways of experiencing surface colours, and we can have better or worse access to surface colours depending on illumination conditions. We can also posit the converse, that we have better or worse access to illuminant colours depending on surface conditions.

By positing surface and illuminant colours, this view undermines the argument for reflectance physicalism. This is an unfortunate outcome if one seeks an argument for reflectance physicalism. It is not an unfortunate outcome if one seeks an argument for some form of colour objectivism, for illuminant colours are in theory as objective as surface colours.

Independently of one's ontological interests, views admitting surface and illuminant colours can maintain a crucial role for colour constancy, when we conceive of colour constancy as efforts visual systems make to disentangle and track stable surface and illuminant contributions to the colour signal. This requires departing from the traditional conception of colour constancy as efforts to isolate stable surface colours across illumination variations, but the departure is minimal. The departure is certainly not a rejection of the spirit of the traditional conception colour constancy, as would for example obtain if illuminants and surfaces collectively impact a single type of colour, say relational surface colour (as in Chapter 19). We will shortly see an analysis of colour constancy that proposes this kind of more radical break.

The dual-colour view receives anecdotal support from naïve perceptual judgements that distinguish surface and illuminant parameters, for example claims that we can experience different surfaces as similarly illuminated, and uniform surfaces as differently illuminated. But such reports must, under scrutiny, buttress the idea that both illuminants and surfaces impact distinctly experienced colours. Unfortunately, traditional colour constancy experiments are not designed to test this idea. They perhaps come closest in the surface and illuminant change attribution experiments (§1). There is additional evidence that we are very good at identifying changes in overall illumination, as opposed to not noticing it (e.g. because it is "adapted out") or conflating such changes with changes in surface colour (Gerhard & Maloney 2010). However, further empirical work would be needed to assess the viability of the dual-colour proposal.

§2.3 Other colour ontologies

Independently of whether or not colour constancy can be used to *justify* some form of colour objectivism, one might wonder how *compatible* constancy is with other colour ontologies. It is very compatible. While some evidence of this can be drawn from the differing analyses of constancy covered below, it is perhaps useful to reflect on the general idea that experiencing colour to be constant across illumination variations only demands a stable colour *experience* [or *representation*] across changes in the illumination parameter. There is no antecedent barrier to our visual systems *constructing* such experiences [/representations], regardless of whether or not they are veridical. This opens the door to consistency between constancy and various subjectivist ontologies (Chapters 21 and 22), although some form of error-theory about colour perception is plausibly required. There is also no antecedent barrier to the experienced stable colour being a *sui generis* property, and thus colour primitivist views are consistent with constancy (Chapter 18). It is much trickier to incorporate constancy into colour relationalism (Chapter 19), but we will soon consider a way of closing this gap. None of this diminishes the interest of the above argument, but it is worth keeping in mind that above argument uses colour constancy to *justify* colour objectivism, and in particular a reductivist form of it, and doesn't try to show that constancy requires objectivism.

§3 Colour constancy and epistemology: the given in colour perception

What do subjects receive from the world via perception? With a bit more precision, suppose a subject reports on a perception in speech, thought. or non-verbal behaviour with the intent simply to state or act on what it is that she perceives. Call these *perceptual reports*, and what they express the *contents* of those reports.[5] How much of the content of a report reflects what is *given* to her in perceptual experience, and how much reflects information or hypotheses that outstrip what perceptual experience itself makes available?[6] This is a central question in philosophy of perception, a key issue in perceptual epistemology, and it informs our understanding of the mind-body-world relation. It is also a very difficult question to assess.

Colour constancy is a fascinating case study of this basic struggle. Suppose a thing has a uniform surface. When we see the thing to be partly in shadow, in some sense we experience it to have a constant colour, in some sense we experience the shadowed part to be darker in colour than the unshadowed part. How can the same experience elicit tendencies to report both colour constancy and colour variation? Two broad proposals are found in the literature, at their core differing on what we are perceptually given in this kind of case. One proposal asserts that one of the reports (variation or constancy) reflects what is experienced and the other (constancy or variation) reflects something else, for example a non-experiential perceptual output or a post-perceptual judgement. This is a two-component approach to the puzzle (§3.1). The other proposal asserts that each constancy and variation report reflects a different part of the same experience. This approach seeks to resolve the puzzle by complicating our account of what is perceptually given (§3.2).

The core psychophysical evidence for colour constancy—data from asymmetric matching experiments—informs this discussion in two direct ways. First, how should we interpret the divergent data elicited by the Appearance Match and Surface Match instructions, given that both data sets emerge from subject efforts to perceptually match the same stimulus? According to the first 'two-component' proposal, one of these reports reflects what is perceptually experienced, and the other reflects something else. According to the second proposal, both reports pick up on different aspects of the same perceptual experience. Second, it is important to note that neither set of responses is near "perfect": Surface Match responses deviate significantly from perfect constancy (CCI = 0.6–0.8; 1 being perfect), and Appearance Match data deviate significantly from perfect variation (CCI = 0.2–0.4; 0 being perfect). Taken at face value, this means that subjects experience neither a surface colour that is fully separated from illumination contributions (considering Surface Match data), nor a colour that fully blends the surface and illuminant contributions (considering Appearance Match data). So, what is given to subjects?

There is no simple way out of this quandary. Nonetheless, the following sketch outlines some solutions that contain fruitful insights. The first two accounts propose a two-component—phenomenal and non-phenomenal—analysis of colour constancy, while the second two aim for a more purely phenomenal analysis.

§3.1 Two-component solutions

Two component accounts of colour constancy propose one component to capture colour experience and the other, non-experiential component, to explain residual constancy data (Cohen 2008; Wright 2013). In theory, the second component might be perceptual or cognitive. If cognitive, the second component is something akin to a post-perceptual judgement. If perceptual, the idea is that our perceptual systems, in addition to facilitating colour experiences, generate non-experiential outputs that impact subjects' dispositions to formulate constancy reports (i.e. dispose subjects to formulate thoughts, words, actions with colour constancy contents). Both two component accounts I consider opt for a perceptual (but non-experiential) interpretation of the second component. Most likely this is because they wish their accounts of constancy to preserve constancy's status as a perceptual phenomenon. This means that post-perceptual judgement is not used to fill any gaps between perception and perceptual reports.

Cohen (2008) argues that constant colours are not given in perceptual experience. Instead, what is given are variable, illumination-dependent colours. A second, *counterfactual* component, explains constancy reports. This content is of the form <surface$_1$ would match surface$_2$ were both similarly illuminated>, and is a non-experiential, perceptual output. The rough idea is as follows. Colour experience varies across illumination changes (in-line with Cohen's colour relationalism—see Chapter 19), and thus there is no experienced colour constancy. Appearance Match data reflects these variable experienced colours. Evidence of colour constancy from both Surface Match and pretheoretic sources is explained by appeal to counterfactual contents. Interestingly, *even here* colour attributions are illumination-dependent, that is, reports of constant colours are reports of colours matching (experientially or counterfactually) under similar conditions of illumination. Cohen thus generally rejects the very idea that constancy involves illumination-independent colour. Although this provides a straightforward analysis of the divergent data sets from asymmetric matching experiments, it substantively departs from the traditional conception of colour constancy—indeed one could argue that it is an attempt to redefine colour constancy from scratch. If the justification for colour relationalism is strong enough, this redefining may be warranted. It is worth noting that to this point the account does not illuminate why both Appearance and Surface Match data are imperfect. However, in theory Cohen's relationalism does not commit him to colour outputs being perfectly variable across illumination variations, it only commits him to colour outputs being variable across those variations.

Wright (2013) argues for two perceptual constancy mechanisms: a phenomenal and a non-phenomenal (in his terminology *projective*) mechanism. Both can contain approximate colour constancy, though to different extents, via different internal mechanisms, and in response to different kinds of stimuli. Phenomenal approximate constancy occurs most notably when low-level adaptive mechanisms are able to factor-out some of the shift due to changes in overall illumination conditions. We experience this when sunshine changes to cloud cover and our visual systems partially accommodate to the new illumination conditions. However, within these experiences many sharp changes in illumination conditions are still experienced and experienced as variations in colour. Shadowing effects are a prime example. Thus, while there is some degree of constancy in colour experience (e.g. regarding changes in overall illumination conditions), there is still robust colour variation (e.g. regarding local illumination contrasts). This colour variation is used to explain the Appearance Match data. However, even here Wright believes that there can be a powerful sense of surface colour constancy—this is what Surface Match data illustrates. Projective constancy is postulated to explain this phenomenon. It is a non-experiential, perceptual output of an approximately constant colour that explains why the partially shadowed thing in some sense looks to have a constant colour. Thus, like Cohen, Appearance Match data is explained by appeal to colour experience, and Surface Match by appeal to a non-phenomenal perceptual output. Since both of Wright's constancy mechanisms yield at best approximate constancy, the colour "outputs" from both are to some degree illumination-dependent. This provides a direct account of the imperfect responses from both Appearance and Surface Match conditions.

Both Cohen's and Wright's accounts propose two perceptual outputs, a phenomenal and a non-phenomenal output. Subject reports can be informed by one or both of these, and are not readily dissectible to determine which or how much of each component informs a given report. It is in the first instance the task of empirical researchers to try to tease out these disparate contributions to perceptual reports. However, the task is extremely difficult. Part of Cohen's argument is that asymmetric matching data doesn't and likely won't straightforwardly favour a traditional conception of constancy over a relationalist-friendly counterfactual one. Instead the matter is arguably to be settled by appeal to broader considerations. I don't detect a similar commitment in Wright to the underdetermination of theory by evidence, but he certainly agrees that teasing apart the various contributions to perceptual reports is horribly challenging.

Interestingly, although both authors propose two colour outputs that are to some degree illumination-dependent, there is an important difference. Cohen's approach is designed to minimize if not excise our commitment to illumination-independent colours in perception, for both experienced colour and counterfactual contents ascribe illumination-dependent colour to things. By contrast, Wright's account is designed to recover a considerable degree of illumination-independent colour in perception, since both phenomenal and projective components express different forms of approximate constancy (i.e. different forms of colour stability across illumination variations). It would be worthwhile to work through this difference in future work.

§3.2 Solutions from phenomenology

The second type of account tries to explain colour constancy via perceptual experience, as opposed to in part via a non-experiential component. This is achieved by making colour experience more complex in the following respect. It is often tacitly assumed that in typical visual perceptions a uniform surface is simply experienced to have a single, straightforward colour (perhaps with a determinate hue, saturation, and lightness value). By contrast the views under consideration hold that when one perceives a uniform surface, the purported surface colour is experienced, not *simpliciter*, but in some *way* or from some *perspective*. These perspectives qualify, in some manner, the experienced colour, creating the conceptual space for illumination variations to enter into experience without thereby altering the underlying surface colour. Although views of this sort are in theory compatible with a two-component approach, what is distinctive of them is their attempt to explain more with colour experience than their rivals. Numerous approaches fall into this broad category.[7] Some of them treat perspectives on colour as primitive (arguably Kalderon 2008 and Gert 2010). Others explicate perspectives on colour by appeal to more basic elements. I briefly outline two options of the latter sort. Both views were introduced in §2, but here the emphasis is on their approach to colour experience, as opposed to their approach to colour ontology.

One approach (Hilbert 2005, 2011; Jagnow 2010) proposes that illumination often enters into colour experience as a non-coloured entity that helps reveal colours to us (e.g. perhaps by making colours more or less vivid). Thus, during colour constancy perceptions, when one perceives a uniform surface that is differently illuminated, one experiences that surface colour to be differently illuminated. The sameness in colour is part of what is experienced, but the same colour is experienced from different perspectives, specifically under different illuminants. Notably, while the experienced illumination variation can impact our epistemic access to colours (seeing them better or worse), it is not typically a variation in colour *proper* (e.g. hue, saturation, lightness).

Since there is a sameness in experienced colour across illumination variations, this approach can explain constancy reports, including Surface Match data, in terms of experienced colour. That said, the degree of imperfection in Surface Match data isn't straightforwardly explained. It is also unclear how the account explains Appearance Match data, where there is robust constancy failure. In Appearance Match conditions subjects are asked to match the hue and saturation of the target and test colours. Thus, it seems straightforward that in at least this regard illumination variation impacts experienced colour.

Another alternative proposes that illumination typically enters into colour experience and enters it as a coloured entity alongside surface colours (Brown 2014). During surface colour constancy perceptions, one experiences a stable surface colour and variable illuminant colours. Similarly, during illuminant colour constancy perceptions, one experiences a stable illuminant colour and variable surface colours. It is supposed that in some sense we experience surface colours through illuminant colours, a natural conceptualization being that we experience colour layers along a line of sight. Regardless, one's experience of each is *interdependent* on one's experience of the other, and thus during colour constancy subjects are perceptually given two distinct colours along a given line of sight.

This yields several notable consequences. First, an experience of a surface colour under two distinct illuminant colours can be interpreted as one having two different *perspectives* on that surface colour, each perspective facilitating distinctive means of accessing that colour. In this regard the epistemic role of illuminants familiar from the Hilbert-Jagnow approach can be incorporated into the view. Second, since during colour constancy subjects are perceptually given two distinct colours along a line of sight, any reports about a single experienced colour require that subjects perform a dissection of what is given that seeks to isolate the colour of interest. Such dissections invariably involve some form of post-perceptual activity (e.g. judgement, mental imagery) on the part of subjects and, without training, are likely fallible. This permits an interesting interpretation of Surface and Appearance Match conditions.

In Surface Match conditions subjects are asked to select the surface contribution to experienced colour and match for it, ignoring the illuminant contribution. However, because the illuminant colour is nonetheless experienced on both the test and target patches a kind of guess is required, and the post-manipulation test patch (presented under one illuminant) should not be expected to perfectly match the target patch (presented under a different illuminant). This provides a means of explaining the imperfection in Surface Match data and the fact that subjects often report being unable to perfectly perform a surface match.

Appearance Match conditions effectively ask subjects to adjust the test patch so that it matches a fusion or mixture of the surface and illuminant colours from the target patch. However, it is not obvious that a fusion of the surface and illuminant colours from the target patch is perceptually given to subjects—indeed if the target patch is inducing a colour constancy experience then on this view such a fusion is *not* experienced. The same holds for the test patch. In this regard the Appearance Match instructions presume a perceptual given that may not obtain.[8] This account thus also provides a means of explaining the imperfection in Appearance Match data.

The view can be developed in various ways, including by allowing for non-constancy experiences that fuse aspects of surface and illuminant contributions, and by allowing for experiences of *partial colours* (see Brown 2014). Such developments create the conceptual space to consider that there may be varying extents to which illuminant and surface contributions are experienced, and are distinctly experienced, in response to a stimulus. This is useful because it may be that the extent to which we experience, and distinctly experience, surface and illuminant contributions is sensitive to a host of factors, including not only the nature of the current stimulus, but also previous experiences, stimuli, cognitions, et cetera. Postulating that our colour experiences are sensitive in such ways might help explain the perplexing data that has arisen from constancy research. As stated above, while such an account is consistent with there being non-experiential perceptual outputs influencing subject reports, the account is distinguished by creating numerous avenues within perceptual experience to explain disparate phenomena.

§3.3 Additional considerations

There are important issues about colour experience in the vicinity, some of which were highlighted in §0. To what extent are constancy reports of determinable or *coarse-grained* colours as opposed to determinate or *fine-grained* ones and what does this tell us about what is given in colour experience? Evidence for the existence of *unconscious* colour constancy tendencies has recently been offered (Norman *et al.* 2014). To what extent do these tendencies impact constancy reports offered in response to conscious colour constancy perceptions?

Also of interest is the idea of *relational* colour constancy, roughly that colour constancy only reflects a perceived stability in some relation(s) between colours across illumination variations, as opposed to a stability in colour itself (Foster 2003). For example, under one illuminant one might perceive the colour in one region to be (say) 20% lighter than the colour of another region. That lightness relation might be preserved under a different illuminant, even though the perceived colour of each region has changed. In this case variable colours are perceptually given in constancy experience, but those variable colours maintain at least some stable relations between other (variable) colours in a scene. Physiologically, the idea is grounded in a stability in cone excitation ratios across illumination variations. It is also readily applicable to constancy reports, where <x and y match in colour> is interpreted to mean the weaker claim <x and y relate to other perceived colours in the same way>. This can be offered as the correct interpretation of our reports, as a rational reconstruction of them, or of what our reports should be.

As plausible as this might sound, there are key limitations. We typically regard the "stability" in colour constancy as at least involving experiencing a thing's colour across illumination variations to be in the same colour category. This is arguably supported both by reflection on everyday experience and empirical data. On its own relational constancy does not mandate this constraint and so cannot explain what seems to be a central feature of colour constancy. Nonetheless, the relational conception is a potentially important ingredient to constancy processing and maybe to constancy experience.[9]

§4 Conclusion

Colour constancy is of interest to colour ontologists largely because it can be used to construct an argument for colour objectivism that is particularly attractive to naturalistic philosophers. While there are difficulties with the argument, it serves as an important basecamp to counter familiar arguments from variation, illusion, et cetera that have created a safe space for subjectivist colour ontologies. It is noteworthy that a more liberalized objectivist ontology, one that admits not only surface but also illuminant colours, overcomes the core difficulty the argument faces from the routine impact illumination has on experienced colour.

Colour constancy is also of interest to perceptual epistemologists, as it illustrates how difficult it is to capture what is given in colour experience, how that informs colour perceptual reports, and how to interpret the meanings of those reports. The plethora of disparate accounts on offer is testament to these difficulties, and we have only partially come to grips with the depth of the challenge.

Beyond this, there is much to be gained from exploring wider conceptions of colour constancy as opposed to merely focusing on the narrow conception ('experienced colour is fairly stable across illumination variations'). Colour constancy mechanisms are not hostage to illumination variations and surface stabilities, but can to some extent be triggered via variations in media/filters (e.g. looking through air versus tinted glasses versus water) and variations in background/surrounding objects. The concept *colour constancy* as it is employed in vision science is typically not understood this broadly (an exception is Brown 2003). However, if the same or similar mechanisms are triggered in these broader contexts, philosophers should at least be open to the *phenomenon* demanding a broader treatment.

There are also worthwhile connections outside the microcosm of philosophy of colour. For example, little work has been done to examine how colour constancy impacts debates between rival theories of perceptual experience (e.g. naïve realism, representationalism, sense-datum theory). There is ample to chew on here, as is revealed by reflection on the different one- and two-component accounts of constancy. In addition, the relation between colour constancy and other perceptual constancies (e.g. shape and size constancy) is understudied, despite being important for broader projects. Consider two brief examples. The idea that colour and other constancies are substantively similar is critical to some of Gert's (2017) arguments for his preferred account of colour. Outside colour, the idea that constancies can be treated *en masse* to motivate significant philosophical theses is central to Smith (2002) and Burge (2010). Yet, given our discussion, colour constancy is difficult to interpret, and many of its perplexing features at least seem to be peculiarities about colour. Thus, any presumption in favour of there being a uniform treatment of constancies that reflects a consensus view should give us pause. Suffice to say, colour constancy is a rich phenomenon that we understand well-enough to use to illuminate core question in philosophy of colour and perception more widely, and that we understand poorly enough to motivate future endeavours.[10]

Notes

1 Helpful recent review articles include Foster (2011) and Shevell & Kingdom (2008). An earlier but insightful review is Jameson & Hurvich (1989).

2 Recall (§0.2) that colour constancy is rarely perfect and instead typically viewed as a tendency of visual perception. CCIs ranging from 0.6–0.8 in the Surface Match condition are a good empirical indication of this. Perfect constancy (CCI = 1) is rarely observed. Instead, we observe reports that fall far short of perfect constancy, but nonetheless imply that some degree of experienced colour stability is achieved. Thus approximate (as opposed to perfect) colour constancy emerges as a robust, empirically supported phenomenon. The CCI = 0.2–0.4 results from Appearance Match conditions demonstrate a robust failure of colour stability in these conditions.

3 A brief remark about the other two kinds of experiments. In achromatic adjustment experiments subjects are presented with a stimulus (e.g. a Mondrian) and asked to adjust a test patch so that it appears achromatic. This is not a measure of surface colour constancy, but can provide useful information about the locally experienced illuminant colour, and about the influence of memory on colour experience. Colour naming paradigms ask subjects to identify colours in different illumination conditions using a selection of colour names. There is a naturalness to the appeal to colour names, but this is at times counteracted by struggles with defining adequately fine-grained terminology.

4 See Chapter 17, Bradley and Tye (2001, 479–80), and Byrne and Hilbert (2003, 8–9) for this kind of argument. Two qualifications might be helpful. The first concerns Byrne and Hilbert's colour ontology. They maintain that surface colours are identifiable with *classes* of SSRs. Colours more broadly, which they take to include surfaces colours, volume colours, and the colours of light sources, are identifiable with classes of what they call *productances*. Light itself, as opposed to light sources, is explicitly excluded from their colour ontology. Second, a crucial premise in Hilbert (1992), also mentioned in Chapter 17, ties colour constancy to the *function* of colour vision, and seeks to reach conclusions about colour ontology from said function. I will set this issue aside (see Brown 2017, for commentary).

5 Although it might initially seem odd to talk about the contents of non-verbal behavioural reports, as opposed to the contents of thought or speech reports, it is fairly common to do so. For example, a subject might believe that blue things contain food and when hungry reach for the blue container. This is a reasonable expression of the content <that is blue> and that expression may but need not coincide with the formation of a thought or utterance to that effect. More directly, if a subject is asked to match two colours then push 'Enter', her pressing of 'Enter' following her manipulation of colours is naturally taken to be an expression of <those colours match>. Readers should at minimum admit that it is rational for us to ascribe such contents to these actions. This is enough to motivate what follows.

6 The classic discussion of the perceptual given is Sellars (1956). Recent work includes Gupta (2006), Crane (2013), and Montague (2016). I take the *given* in perception to refer to what a subject perceptually *experiences*. I do not mean to imply that what is given in perception is non-inferentially knowable or yields infallible knowledge, but merely that the perceptual given is what perception itself makes consciously available to perceivers for use in perceptual reports. There are other valuable ways of construing the idea that are not as tied to experience. For example, one could take the perceptual given to be what perceptual systems "deliver", via experiential or non-experiential channels, to subjects for use in reports, or what perceptual systems make available to other neural/cognitive systems for use in their processing. Such construals make no critical link between the perceptual given and perceptual experience, and thus I set them aside. My commitment to a given-experience link doesn't prevent there from being non-experiential contributions to perceptual reports (see below). It merely prevents such contributions from being part of the perceptual given.

7 E.g. Hilbert (2005, 2011), Kalderon (2008), Gert (2010), Jagnow (2010), Brown (2014), Davies (2016).

8 In slightly more detail, in Appearance Match conditions subjects are potentially asked to mentally construct a blend of the surface and illuminant target patch colours, and manipulate the test patch to a state presenting surface and illuminant colours that, if blended, would match the first construct. That is, they are asked to match for two surface-illuminant colour blends, neither of which are perceptually given.

9 Will Davies (in progress) is working through ways to develop the relational conception of colour constancy.

10 I am indebted to Keith Allen, Mazviita Chirimuuta, and Will Davies for helpful comments on an earlier draft. My sincere thanks to the Templeton Foundation, via The University of Cambridge and

Tim Crane's *New Directions in the Study of Mind* project, for financial assistance. This publication was made possible through the support of a grant from the John Templeton Foundation. The opinions expressed in this publication are those of the author and do not necessarily reflect the views of the John Templeton Foundation.

References

Arend, L., & Reeves, A. 1986. Simultaneous color constancy. *Journal of the Optical Society of America A—Optics Image Science and Vision*, 3, 1743–51.

Arend, L.E., Jr, Reeves, A., Schirillo, J., & Goldstein, R. 1991. Simultaneous color constancy: Papers with diverse Munsell values. *Journal of the Optical Society of America A—Optics Image Science and Vision*, 8, 661–72.

Bradley, P., & Tye, M. 2001. Of colors, kestrels, caterpillars, and leaves. *Journal of Philosophy*, 98, 469–87.

Brown, D. 2014. Colour constancy and colour layering. *Philosophers' Imprint*, 14(16), 1–31.

Brown, D. 2017. Colouring for and colour relationalism: A critical notice of *Outside Colour* (M. Chirimuuta. 2015. MIT Press). *Analysis Reviews*, 77(2), 433–49. doi:10.1093/analys/anx025.

Brown, R. 2003. Background and illuminants: The yin and yang of colour constancy. In R. Mausfeld & D. Heyer, eds., *Colour Perception: Mind and the Physical World*, pp. 247–72. Oxford, UK: Oxford University Press.

Burge, T. 2010. *Origins of Objectivity*. Oxford, UK: Oxford University Press.

Byrne, A., & Hilbert, D. 2003. Colour realism and colour science. *Behavioral and Brain Sciences*, 26, 3–64.

Byrne, A., & Hilbert, D. 2007. Color primitivism. *Erkenntnis*, 66, 73–105.

Chirimuuta, M. 2008. Reflectance realism and colour constancy: What would count as scientific evidence for Hilbert's ontology of colour? *Australasian Journal of Philosophy*, 86(4), 563–82.

Cohen, J. 2008. Colour constancy as counterfactual. *Australasian Journal of Philosophy*, 86, 61–92.

Crane, T. 2013. The given. In J. Schear, ed., *Mind, Reason and Being-in-the-World: The McDowell-Dreyfus Debate*, pp. 229–49. London: Routledge.

Davies, W. 2016. Colour constancy, illumination, and matching. *Philosophy of Science*, 83(4), 540–62.

Davies, W. 2017. Relational colour constancy. Lecture delivered at the *Philosophy, Psychology and Neuroscience Research Seminar* at the University of Glasgow, 27 November.

Foster, D. 2003. Does colour constancy exist? *Trends in Cognitive Sciences*, 7, 439–43.

Foster, D. 2011. Color constancy. *Vision Research*, 51, 674–700.

Gerhard, H., & Maloney, L. 2010. Detection of light transformations and concomitant changes in surface albedo. *Journal of Vision*, 10(9), 1–14.

Gert, J. 2010. Color constancy, complexity, and counterfactual. *Noûs*, 44(4), 669–90.

Gert, J. 2017. *Primitive Colours*. Oxford, UK: Oxford University Press.

Gupta, A. 2006. *Empiricism and Experience*. Oxford, UK: Oxford University Press.

Hilbert, D. 1992. What is color vision? *Philosophical Studies*, 68(3), 351–70.

Hilbert, D. 2005. Colour constancy and the complexity of colour. *Philosophical Topics* 33(1), 141–58.

Hilbert, D. 2011. Constancy, content, and inference. In G. Hatfield & S. Allred, eds., *Visual Experience: Sensation, Cognition, and Constancy*, pp. 199–211. Oxford, UK: Oxford University Press.

Jagnow, R. 2010. Shadow-experiences and the phenomenal structure of colors. *Dialectica*, *64*(2), 187–212.

Jameson, D., & Hurvich, L.M. 1989. Essay concerning color constancy. *Annual Review of Psychology*, 40, 1–22.

Kalderon, M.E. 2008. Metamerism, constancy, and knowing which. *Mind*, 117, 935–71.

Land, E., & McCann, J. (1971). Lightness and Retinex theory. *Journal of the Optical Society of America*, 61, 1–11.

Mausfeld, R. 2003. 'Colour' as part of the format of different perceptual primitives: The dual coding of colour. In R. Mausfeld & D. Heyer, eds., *Colour Perception: Mind and the Physical World*, pp. 381–430. Oxford: Oxford University Press.

Montague, M. 2016. *The Given: Experience and its Content*. Oxford, UK: Oxford University Press.

Norman, L., Akins, K., Heywood, C., & Kentridge, R. 2014. Colour constancy for an unseen surface. *Current Biology*, 24, 2822–66.

Reeves, A.J., Amano, K., & Foster, D.H. 2008. Color constancy: Phenomenal or projective? *Perception & Psychophysics*, 70, 219–28.

Sellars, W. 1956. Empiricism and the philosophy of mind. In H. Feigel & M. Scriven, eds., *Minnesota Studies in the Philosophy of Science*, vol. I, pp. 253–329. Minneapolis, MN: University of Minnesota Press.

Shevell, S.K., & Kingdom, F.A.A. 2008. Color in complex scenes. *Annual Review of Psychology*, 59, 143–66.

Smith, A.D. 2002. *The Problem of Perception*. Cambridge, MA: Harvard University Press.

Tokunaga, R., & Logvinenko, A.D. 2010. Material and lighting dimensions of object colour. *Vision Research*, 50, 1740–7.

Wandell, B. 1989. Color constancy and the natural image. *Physica Scripta*, *39*, 187–92. Reprinted in Byrne & Hilbert, eds. 1997. *Readings in Color Volume 2: The Science of Color*, pp. 161–75. Cambridge, MA: MIT Press.

Wright, W. 2013. Color constancy reconsidered. *Acta Analytica*, 28, 435–55.