UNIVERSITÉ DE FRIBOURG
UNIVERSITÄT FREIBURG

Antoine Rappaz-Bonzon

On the evidential relationship between color constancy theories and color ontology

DOI : https://doi.org/10.51363/unifr.lth.2025.039

Antoine Rappaz-Bonzon https://orcid.org/0009-0001-8872-9853

On the evidential relationship between color constancy theories and color ontology

Rappaz-Bonzon Antoine
Ormont-Dessous, Vaud
2024
Thèse de doctorat présentée à la Faculté des lettres et des sciences humaines de l'Université de Fribourg (Suisse)

Sous la direction de
Prof. Martine Nida-Rümelin, Université de Fribourg, CH
Et
Prof. Fiona Macpherson, Université de Glasgow, UK

Approuvée par la Faculté des lettres et des sciences humaines sur proposition des professeurs Dr. Martine Nida-Rümelin, Dr. Fiona Macpherson et Dr. Derek Brown, Fribourg, le 29 novembre 2024, Le Doyen Prof. Dr. Dominik Schöbi

Acknowledgments

I am deeply grateful to my parents for their unwavering support throughout the years. Their steadfast belief in me, which never faltered, has been a cornerstone of my journey. To my parents-in-law, I extend my heartfelt thanks for their genuine efforts to understand the work I have undertaken. I am equally indebted to my friends – Anna, Gaël, Estelle, Raoul and Lucille – for standing by my side during my examination and offering their unyielding support. Their presence made a challenging moment all the more bearable.

Above all, I owe an immeasurable debt of gratitude to my wife, my better half, Céline – my little Snowflake. Her enduring support throughout the years of writing this dissertation has been nothing short of extraordinary. She stood by me through both triumphs and trials, guiding me through the difficult moments and helping me find the light at the end of the tunnel.

Table of contents

Chapter 1: Introduction

Introduction

Science and philosophy

The interactions between the sciences and philosophy throughout history are deep and numerous. This is evidenced for instance by the various historical figures who contributed to both fields. Famous for his philosophical approach and theories, Descartes was also an accomplished physicist in his time and developed influential laws of motion. According to the first of these laws, rest and motion are opposite and primitive states of bodies so that a body will only change its state if acted upon by an external force. A body in motion will thus always try to go back to moving and similarly for a body at rest, which will always try go back to a state of rest. According to Slowik, “it is not an exaggeration to claim that Descartes helped to lay the foundation for the modern theory of dynamics” (Slowik, 2023).

As another example, Pierre Duhem is well-known for his work in the philosophy of science and his contribution to the Quine-Duhem thesis, which states that theoretical statements cannot be experimentally infirmed or confirmed independently of overarching hypotheses. However, Duhem was also a famous physicist and contributed to the Duhem-Margules and the Gibbs-Duhem equations. The second of these equations describes how the chemical potentials of components in a mixture change with composition. The first, in turn, deals with the relationship between chemical compositions and activity coefficients, where the latter measure how the behavior of a component in a mixture deviates from the ideal behavior.

Fortunately, the interactions between science and philosophy are not restricted to historical figures doing both at the same time so that each part of their work is influenced by the other. Theories themselves can also influence the development of theories in the other research field. Scientific theories can have theoretical import in philosophy. For instance, the special theory of relativity claims that there is no fact of the matter as to whether two spatially distinct events are simultaneous. That is, it entails the relativity of simultaneity. The latter can in turn be shown to entail that there is no fact of the matter as to whether an event is present or not. But this seems to directly contradict the ontological theory of time known as presentism, which claims that the present is ontologically special in some sense, as compared to the past and future. We therefore have an example of a scientific theory informing us regarding which of our philosophical theories can be true or not[1].

We have examples going in the other direction too. Philosophical theories can also influence the development of scientific theories. Atomism began as a family of philosophical theories which claimed that there are indivisible, smallest parts of matter which can *inter alia* explain the behavior of macroscopic bodies. Most notably, these theories were mostly found in Jaina, Vedic and Ancient Greek philosophies (Berryman, 2022). Importantly, the various theories of atomism in these times were metaphysical theories established through a priori arguments. On the contrary, modern atomism was developed through the use of a posteriori, scientific methodology. But this does not eclipse the fact that the elaboration of the various kinds of atomism in ancient philosophy represented fertile grounds for the development of the scientific versions of the theory. As Chalmers puts it, "twentieth-century atomism in a sense represents the achievement of the Ancient Greek ideal insofar as it is a theory of the properties of matter in general in terms of basic particles, electrons, protons and neutrons, characterized in terms of a few basic properties" (Chalmers, 2019).

In the case of the research on color, we can also observe fruitful interactions between philosophy on the one hand and science on the other. In the Iraqi medieval times, Ibn Al-Haytham developed a theory of perception which had an impact on his views on color by essentially implying that colors are properties found in objects (El-Bizri, 2005, 2009)[2]. According to his theory of perception, the color of an object is emitted by the object alongside the light reflected by it. Rather than being a consequence of the light hitting the object, the color is already there in the object and only requires the object to be illuminated to be carried to the observer. During the Enlightenment, Newton famously argued that his theory of optics entailed that we have to endorse a dispositionalist theory of colors, according to which colors are powers of objects to appear in a certain way to observers and not primary properties which explain the behavior of matter.

In contemporary philosophy, the interest of philosophers for scientific data and theories was renewed by Hardin (Hardin, 1988). In his seminal book, Hardin delves into the variety of phenomena studied by color science such as cone opponency, metamerism and perceptual variation in color. His aim was to determine what we could learn about color by focusing on what color science had to say about it. He tried and showed that many of the

---

[1] The same could be said in the case of quantum mechanics and our theory of logic, even though the matter is much less clear (Putnam, 1974, 2005).

[2] Although it is unclear whether in that theory colors ought to be seen as mind-independent or dispositional and mind-dependent properties of objects.

phenomena studied by color science represent an insurmountable issue for color realism, according to which colors are mind-independent properties of ordinary objects such as lemons and grass.

To do so, he proceeded in the following way. He enumerated the properties colors are thought to have by what we may temporarily call our commonsense conception of colors. For instance, each color is usually thought to be either unique or binary. Certain colors are binary in that they seem to "contain" others. Turquoise contains blue and green and orange contains red and yellow. In contrast, yellow, blue, red and green appear to admit of shades that are unique in that they do not seem to contain or be composed of other colors[3]. For instance, unique yellow is neither reddish (unlike orange, which is both yellowish and reddish) nor greenish (unlike chartreuse, which is both greenish and yellowish).

The unique-binary distinction is thought to be extremely hard to explain by some theories of the nature of color. According to a certain type of color realism of the reductive variety, colors are identical with certain physical properties. However, if we accept that colors are either unique or binary, then those physical properties should also be either unique or binary. The issue according to Hardin is that there is no meaningful sense in which the candidates for identification with the colors can be said to be unique or binary. But if reductive color realism cannot explain the unique-binary distinction, then this represents a reason to reject it.

After having proceeded in the same way for many of the supposed properties of color, he goes on to conclude that we have to endorse *color eliminativism*. That is to say, we have to accept that colors, although they may at first glance appear as mind-independent properties of objects (i.e. although some form of color realism appears to be true), do not in fact exist. Hardin thus deals a swift blow to any theory which stipulates that colors are mind-independent, physical properties of objects[4].

Perhaps a welcomed addition to the recent philosophy of color was the development of variants of reductive realism which try to tackle those issues. As a form of realism, they stand opposed to eliminativism, which as we saw is defended by Hardin, in claiming that colors actually exist. But it is also opposed to what we can call *relationalism*. The latter family of theories insists that colors are relational properties of objects which usually require an object,

---

[3] Note that this claim is usually distinguished from the claim that painters can mix colors to produce other colors. For more on the binary-unique distinction, see chapter 5.

[4] Note that, methodologically speaking, Hardin is arguing that our theories of color ought to respect the evidence gathered thus far in color science, which includes sciences as diverse as physics, biology or psychophysics. However, and this will be relevant for what comes next, he is not arguing that ontologies of color might in fact be directly motivated by the empirical success of certain theories in color science.

a perceiver and some circumstances of viewing. According to this view then, an object is not yellow *simpliciter*, but indeed yellow-for-that-particular-perceiver-under-those-particular-circumstances-of-viewing (Cohen, 2004, 2009). Reductive color realists argue that colors are not relational and do not require anything beyond the object itself to be instantiated.

Finally, reductive color realism is opposed to *primitivist* forms of color realism (that is, non-reductive color realism), according to which colors are distinct from any of the properties studied by the sciences (Gert, 2006, 2008). They are *sui generis*, irreducible properties of objects. Therefore, reductive color realism rejects color primitivism's thesis that colors are irreducible properties of objects and instead claims that they are amongst the properties studied by the various sciences (such as physics or chemistry).

So reductive realists hold that colors are identical with or reducible to physical or chemical properties of objects. Perhaps most notable amongst reductive realisms is Byrne and Hilbert's theory, *reflectance physicalism*, which identifies colors with the surface spectral reflectances of objects (Byrne & Hilbert, 2003, 2020b; Hilbert, 1987, 2008). The latter refers to how much energy a surface will reflect as a function of wavelength. A surface might reflect a high proportion of wavelengths around e.g. 450 nm, but very few around 600 nm or the opposite[5]. In a nutshell, the surface spectral reflectance can be thought of as the chromatic signature of an object.

Answering some of the challenges described by Hardin are amongst the tasks of Byrne and Hilbert[6]. One of the challenges they tried to tackle was the following. As we already touched on, colors stand in certain qualitative relations with each other (unique/binary, degree of similarity, incompatibility/compatibility). In addition, one can hold that those relations are essential to the nature of colors. That is, one can hold that orange being more similar to yellow than to blue is essential to orange being the color it is. If it were not so similar, it would not be orange. The challenge results from the claim that there are qualitative relations between the colors and the claim that those relations are essential to the nature of colors[7]. That is, if some properties did not stand in those relations, they could not be the colors. If reflectance physicalism is true, then the SSRs that *are* the colors should stand in the very same relations with respect to one another as the colors do (Boghossian & Velleman, 1991;

---

[5] As it happens, the SSR of natural objects is pretty smooth. That is to say, natural objects will not reflect, e.g., many 500-nm wavelengths but no 501-nm ones.

[6] And all researchers who want to reject eliminativism.

[7] The latter claim could be rejected if we thought that those relations emerge because of the particular conception of our color vision. In that case, they would be merely accidental relations between colors. See chapter 5 for an elaboration of this idea.

Chirimuuta, 2015; Hardin, 1988). However, we have to accept that SSRs do not appear to stand in such relations.

Byrne and Hilbert have developed an answer to this challenge called the magnitude proposal. It begins by reconstruing the content of color experiences. It is usually thought that the latter simply ascribes color properties to objects. That is, one color experience's content might be that the object in front of oneself is blue. However, Byrne and Hilbert argue, it would be more correct to hold that it represents objects as having proportions of hue magnitudes. In that way, redness, e.g., is attributed to an object as a proportion of its total hue. For instance, an orange object will be attributed a high proportion of redness, much more than a yellow object. Objects are thus represented by color experiences as having given hue-magnitudes. In that way, orange and red would indeed be qualitatively similar since they both contain an important amount of red magnitude, whereas blue is distinct from both of them as it has none of it.

If there is a physicalistically acceptable way of understanding what it is to have a hue-magnitude, then reflectance physicalism is not threatened by the challenge emerging from the existence of the phenomenal structure of hues. By "physicalistically acceptable" is meant an account which appeals to the physico-chemical properties of objects. In other words, it must be explained what properties of the objects, which are appealed to in scientific explanations, explain their having a given hue-magnitude. This is one example of a debate surrounding reductive color realism.

Justifying reflectance physicalism using color constancy research

The present work will be in part centered around reductive realism in the form of reflectance physicalism. More precisely, we shall focus on the evidence which can be marshalled in favor of reflectance physicalism. The kinds of evidence which can be gathered for ontologies of color in general can appeal to e.g. empirical data, appeals to intuitions, folk psychology, conceptual considerations and phenomenological considerations[8]. The present work will focus on the empirical data we have concerning the perceptual phenomenon of color constancy. In the following, we will be concerned with a description of this phenomenon. I shall first present a situation in which common color constancy occurs and then try to show why the latter represents such a challenge for our visual system. Finally, we

---

[8] Although it could be argued that this list is redundant if appeals to intuitions are reducible to conceptual considerations or vice versa.

will see how the scientific theories put to work to explain it have been used as evidence in favor of reflectance physicalism.

Consider the following scenario. You are walking in an orchard under a clouded sky and see an apricot tree with what appear to be deliciously ripe fruits[9]. You pick one and take it back home. It is dark now and you are working at your desk under the light of an incandescent lightbulb. You look at the apricot next to you, noticing its vibrant reddish orange. The fruit, which used to sit under the light of the sun, filtered through the clouds, does not appear to have changed in color when it was brought under the light of the incandescent lightbulb. It appears to have conserved its particular shade of orange. Nonetheless, there seems to be a sense in which the color of the apricot under the lightbulb is not quite the same as the one it had in the orchard.

To understand why this situation represents a challenge for your visual system, we have to understand an aspect of the physics of light. Each light which strikes objects in our environment has a certain composition of wavelengths, which we call its spectral power distribution (for short, SPD). Some lights might contain more short wavelengths than long ones, others the opposite. Some have an SPD which is pretty much equal across the visible spectrum, i.e. they reflect roughly the same amount of photons, no matter their wavelength. Others have a more complicated SPD, with peaks and troughs. In addition, objects have a characteristic way of reflecting the light which hits them, which represents their surface spectral reflectance (or SSR). They might reflect a lot of long wavelengths and very few short ones, plenty of middle and short ones but no long ones or, again, have a pretty complicated SSR.

Most importantly, the light which is finally reflected by the object is *inter alia* the result of the SPD of the light that hit the object *and* of the SSR of the object itself. Consider a surface with a SSR such that it reflects 80% of photons with a wavelength of 500 nanometers (at a given light intensity). If this surface is hit by a light which contains, let us say, 10 500-nm photons (obviously, the real number will be much higher), it will reflect 8 photons. In contrast, if it is hit by a light which contains only 5 photons of the same wavelength, it will emit only 4 photons back. Though the reflected light will be distinct in each case, it is always the product of the SSR of the surface and the SPD of the light involved.

The consequences of this phenomenon cause mayhem for our visual system. Take an achromatic surface (i.e. reflecting wavelengths equally across the visible spectrum) such as a

---

[9] Of course, as every Swiss citizen knows, this orchard has to be somewhere in Wallis.

white wall. Let us illuminate this light with an SPD skewed towards the shorter wavelengths. The light which will be reflected by the wall, as a product of the SPD of the light which hit the wall and of the SSR of the wall itself (which, as the wall is achromatic, will reflect wavelengths equally along the visible spectrum), will contain more shorter wavelengths than longer ones. However, note that this reflected light, the *color signal*, is silent as to its origin. It could have been produced by a myriad of different pairs of SPD and SSR. In our case, it has been produced by an achromatic wall under a light with an SPD containing more shorter wavelengths. But it might as well have been the result of a light containing equal proportions of wavelengths across the visible spectrum hitting a wall proportionally reflecting more shorter wavelengths than longer ones.

The problem for the visual system is then apparent. If the same color signal could have been produced by spectrally different pairs of light and surface, how does it come to the right conclusion regarding the nature of the surface it receives light from? Is it a white wall under a chromatic light (in that case, skewed in favor of shorter wavelengths) or a chromatic wall under an achromatic light? As Maximov has shown, a yellow dandelion blossom in the north skylight reflects a color signal with the same chromaticity (i.e. reflecting the same kinds of wavelengths in the same proportions) as grass in direct sunlight (Maximov, 1984, in Foster, 2011). However, even though the light emitted is the same in both situations, we still see the blossom and the grass as wildly different in color. The “apricot” situation is diametrically opposed to this situation. The apricot under the clouded sky reflects a light vastly different from the one it reflects under the light of the incandescent lightbulb and yet we do think that we see it as having the same color in both situations.

But we typically do see the grass as green and the dandelion blossom as yellow. We see that the wall is white and, in the case of color constancy[10], we keep on seeing the apricot as orange (roughly) no matter what illumination falls upon it (clouded sky or incandescent lightbulb). The visual system must therefore have a way of resolving the ambiguity inherent in the light reflected from the various surfaces in its environment. In the case of color constancy, it must have a way of doing so which allows it to represent the color of the surface viewed independently of the variation in the SPD of the illumination. That is, the visual

---

[10] More precisely, this is an example of successive color constancy, where two objects are judged to have the same color across a timespan. In the case of simultaneous color constancy, two identical objects appear to have the same color even though they are viewed in different circumstances. A very natural example is seeing a wall which is partly shadowed by another building. Although one of the sides of the wall is in shadow, we judge the wall to have a uniform color throughout.

system must be able to recover something constant from the surface seen under lights with varying SPDs.

The above represents one way of describing color constancy. In the following section, I will try and explain how Hilbert appeals to a particular class of theories trying to explain color constancy in that exact way to support his branch of color objectivism, reflectance physicalism.

Reverse optics

Theories regarding how our visual system performs such a feat of strength abound. However, for the time being, we shall be concerned with only one class of such theories, namely, those falling under the umbrella of reverse optics (also called inverse optics). According to reverse optics strategies, the general aim of the visual system is to inverse the processes that gave rise to the retinal image. In other words, reverse optics try to determine the solution to a particular inverse problem.

An inverse problem is the opposite of a forward problem, which is the process of calculating the effect of a set of causal factors. That is, a forward problem starts with the cause and calculates the effect. An inverse problem therefore goes in the other direction. From the observation of a phenomenon, one tries to determine the factors which caused it. Forward problems are mathematically well-posed, that is, given the specifications of the causes, it is possible to determine a unique and stable solution to the problem, i.e. it is possible to single out one solution which does not change. In contrast, this is not so for inverse problems. Inverse problems are typically ill-posed and their solution is thus not unique or not stable.

For instance, consider the example of a murder. If one knows the killer, the motive and the circumstances of the murder, it is quite straightforward to determine who is going to be killed. This is a forward problem, with a unique and stable solution. In contrast, a detective will face an inverse problem in figuring out who the killer is given the murder. In the same way, an addition is a forward problem but figuring out which addition yielded a given number is an ill-posed problem. There are just too many additions which are going to yield that exact number[11].

In our context, the forward problem of optics consists in determining the image resulting from a particular light hitting a particular surface which then reflects a particular light in a particular direction to finally reach the back of a retina or a camera sensor. If one

---

[11] In fact, an infinity of additions although it would represent an inverse problem even if there were merely two possible solutions.

knows the relevant factors (including the illumination of the scene, the SSRs of the surfaces involved, the nature of the eventual haze involved, etc.…), it is straightforward to calculate the resulting image. The inverse problem corresponds here to the determination of the causes of the retinal image if one knows the retinal image. In particular, given the retinal image, one aims at recovering knowledge of the surface properties of the scene, i.e., the scene's objects' SSRs.

Inverse optics strategies therefore propose to try and solve the inverse problem of (retinal) image formation in a way which could be implemented by the visual system. Given the image and potentially prior information regarding general scene properties, they will propose an algorithm which, when implemented, actually recovers information about the distal causes of the image, thereby solving the inverse problem. The information recovered will thus concern *inter alia* both the illuminant and the reflectance of surfaces. The type of algorithm proposed and the scene information it needs to operate correctly will vary according to which theory is being considered.

In our case, we must focus our attention on the work of Maloney and Wandell (Maloney, 1986; Maloney & Wandell, 1986). Their work consists in the development of an algorithm capable of recovering estimates of SSRs. The latter appeals first to the notion of linear models[12]. Linear models decompose information into a small number of fixed basis spectral reflectance functions (or basis functions, for short). SSRs are thus decomposed into a weighted sum of basis functions. As the basis functions are fixed, distinct SSRs will be distinguished by the weights they attribute to each basis function. As those basis functions are considered known by the visual system, knowledge of a particular SSR is equivalent to knowledge of the respective weights of each basis function. The same analysis is applied to the illuminant of the scene. The visual system thus needs to recover from the scene two sets of weights, one for each distinct set of basis functions.

Using linear models to analyze spectral information essentially acts as constraints on how an SSR or an SPD can vary. In addition, Maloney and Wandell's strategy includes what they call the *subspace method*. Assuming that the surfaces of the scene are lit by a common illuminant, the color signal is analyzed to determine under which illuminant it is more likely to have been observed. Some color signal might produce e.g. too much L-cone excitation to have been produced by an illuminant containing a lot of shorter wavelengths and very few

---

[12] For a more detailed explanation of linear models as applied to color constancy, see appendix A.

longer ones. It would then be eliminated from the possible candidates[13]. Having identified the probable illuminant of the scene, and thus the weights associated with its basis functions, the algorithm can then deduce the SSR at each location in the scene by setting aside the influence of the illuminant's SPD on the color signal. In that way, recovery of the weights of the basis functions describing SSRs can be achieved.

The end product of Maloney and Wandell's algorithm is the recovery of estimates of SSR. And this is where the link with reflectance physicalism can be made. An additional step to add to the algorithm is to hold that, when a color experience attributes a color to an object, it is actually attributing an estimate of SSR to that object. This means that the colors our experiences represent are nothing but estimates of SSR. Reflectance physicalism thus appears to follow from an endorsement of Maloney and Wandell's model and some other theoretical assumptions about the nature of perception. If the model is empirically successful, as it seems to be, then this is an argument in favor of reflectance physicalism.

Chirimuuta's objection

Hilbert's appeal to this particular work in reverse optics[14] has not been unchallenged (Brill, 2003; Funt, 2003). Most importantly for the present work, Hilbert's appeal to Maloney and Wandell's work to argue for reflectance physicalism has been criticized very strongly by Chirimuuta (Chirimuuta, 2008). Chirimuuta's criticism is twofold. First, she argues that Hilbert's argumentative strategy works only if there is a consensus in vision science to the effect that SSRs are indeed recoverable by the visual system. She presents evidence which concludes that they are not so recoverable but also evidence which concludes that they would not be recovered even if they were recoverable. In a nutshell, Hilbert has been cherry picking at the data.

Secondly, Chirimuuta goes further in claiming that there is a conceptual issue at the heart of the research on color constancy. More precisely, she believes that the absence of consensus in color constancy research regarding whether reflectance is recovered by the visual system (that is, whether *Recovery* is true) boils down to researchers adopting distinct conceptions of the nature of vision and not to the collection of sets of data pointing in divergent directions respectively. This issue needs to be addressed before any appeal to the

---

[13] Note that this strategy could not be applied to a scene containing distinct illuminants, i.e., cases of simultaneous constancy. In that case, the visual system would need another strategy to determine that two distinct illuminants are present.

[14] For other researchers working on reverse optics, see Aston et al., 2019; Brainard, 1998; Brainard & Maloney, 2011; Buchsbaum, 1980; Radonjić & Brainard, 2016.

results of this research can be marshalled in favor of any ontological theory of the colors. The thesis Chirimuuta wants to defend here is that the interpretation of at least some of the data in color constancy research is being influenced by the conceptual positions of the teams of researchers and that the nature of this influence prevents us from drawing any ontological conclusions (regarding, for instance, the nature of colors) from the data. Those conceptual positions include assumptions regarding the nature of perception and, in our case, of vision. This difference in conceptual positions explains at least in part if not fully why there is no consensus regarding the question of reflectance recovery in color constancy research.

In the case of Maloney and Wandell, their conception of vision is based on the work of David Marr and this is why Chirimuuta dubs it the *Marrian framework*. According to the Marrian framework, vision is the recovery of object properties by the visual system through the computation of an inverse problem. If this is true, then there must be a property of objects which is recovered and labelled as color and this property is none other than surface spectral reflectance. In other words, if this is true, then reflectance physicalism is true. Chirimuuta believes that it can be said that the Marrian framework entails the truth of reflectance physicalism. The framework at issue presupposes that reflectance physicalism is true. But this means that interpreting the evidence as favoring the reverse optics framework, and thus as supporting reflectance physicalism, first requires endorsing reflectance physicalism. However, if this is correct, then it seems that appealing to Maloney and Wandell's work to justify reflectance physicalism, as Hilbert appears to do, is viciously circular.

Aim and structure of the present work

The general purpose of the present work is first to nuance Chirimuuta's conclusions. I will argue that the situation is not as dire as she portrays it. Even though Hilbert's case is much shakier than one might have thought before Chirimuuta came along, it is not accurate to say that the general process of finding evidence for reflectance physicalism or other forms of color realism in the literature on color constancy is bound to fail.

More specifically, I shall argue for two distinct main theses. First, it will be argued that, even though there is no consensus regarding *Recovery* in color constancy research, this fact by itself does not constitute a serious threat to reflectance physicalism (hereafter, RP) or other kinds of color realism. It would be such a threat were reverse optics the only family of views in a position to potentially justify RP. However, and this is a central point of the present work, there are other theories which fit adequately with RP and may be able to justify it. That is to say, for a theory to represent a justification for RP, that theory does not need to endorse that reflectance recovery is an actual step of visual processing.

The second thesis I will defend is that it is much harder to argue for Chirimuuta's conclusion that the endorsement of *Recovery* is more of a matter of prior theoretical allegiances than an empirical question. First, Chirimuuta believes that the question of whether *Recovery* is true cannot be settled by appeal to experimental results and that this is so because it is more of a conceptual issue than a question of data. The different teams of researchers interpret the same data in different ways because they endorse different conceptions of the nature of vision. Furthermore, one of those conceptions, endorsed by Maloney and Wandell, actually entails or is at least well-paired with a theory akin to RP.

There are various ways to dispute this argument. The way which will be favored here is to show that the question regarding whether *Recovery* is true or false is much more of an empirical issue than what Chirimuuta believes. If it is then thought to be a central element of some of the conceptual positions regarding the nature of vision, then its being amenable to empirical (dis)confirmation entails that those conceptual positions are themselves, though indirectly, amenable to such evaluation. However, this is so only to the extent that it is a central component of those conceptions. If those conceptions can have a broader notion of tracking properties, as some of them may have, then the attack on *Recovery* does not harm them. Nevertheless, this does not impact the general point that the conceptual positions color researchers endorse regarding the nature of perception have empirical consequences and it is thus possible to reject or validate them through empirical examination. I will thus show that certain experiments actually compare distinct models of color constancy and tease out their respective empirical consequences to see which one holds up in the end.

In chapter 2, I start by presenting Chirimuuta's argument against Hilbert's appeal to color constancy as evidence for his particular brand of color realism, reflectance physicalism. The latter being laid out, we move on to Chirimuuta's two-pronged criticism. She begins by pointing out that there is no consensus regarding *Recovery* in color constancy research, neither concerning whether it is achievable by the visual system nor whether it is actually achieved even if achievable. That is, it is questioned whether the visual system could be able to estimate reflectances but also whether it would recover them were it indeed possible.

We then try to clarify what Chirimuuta believes is at least part of the reason why there is no such consensus. She argues that the very same data is interpreted in contradictory ways (either in favor or against *Recovery*) because the different researchers hold distinct conceptions of the nature of perception and, in particular, of vision. In addition, she maintains that one of such conceptions actually entails or is at least well paired with RP, giving rise to a potentially vicious circle.

Doubts concerning Chirimuuta's strategy will first be laid out in chapter 3. When it comes to the consensus issue, I will ask whether the absence of such consensus regarding the status of *Recovery* represents a nail in the coffin of Hilbert's appeal to color constancy research to justify RP. I will present two distinct theories of color constancy which do not adopt a reverse optics approach to constancy. First is Foster's relational color constancy (Amano & Foster, 2004; Craven & Foster, 1992; Foster, 2003, 2011; Foster et al., 1997; Foster & Nascimento, 1994; Linnell & Foster, 1996; Nascimento & Foster, 1997, 2000). According to this theory, the visual system does not attempt to compute estimates of reflectances (that is, SSRs), which is regarded as an unnecessary complex and probably useless task. It rather pays attention to color relations in the retinal image to reach conclusions regarding the relational colors of objects in the field of view.

Secondly, I present a theory which Chirimuuta herself already discusses, namely, Zaidi's heuristic approach to color constancy (Robilotto & Zaidi, 2004a, 2006; H. Smithson & Zaidi, 2004; Zaidi, 1998, 2000). According to the latter, the visual system proceeds heuristically to solve the tasks it is presented with. That is, it tries and finds the simplest computation possible needed to match surfaces across illuminants. In that case, the authors argue that the visual system uses brightness ratios to determine which objects are similar or different[15]. As reflectance recovery is seen as an algorithmic step that is neither necessary nor simple by Zaidi, his theory eschews such a computation.

Both of these theories are examples of color constancy theories which reject the reverse optics approach characteristic of Maloney and Wandell's work. I will try and determine what their relations may be with regard to RP. That is, is it really true, as Chirimuuta appears to presuppose, that a theory of color constancy must endorse reverse optics to count as justification for some form of color realism? To anticipate, I will argue that it is not and that at least Foster's theory can be seen as providing evidence for a theory similar to RP. As we shall see, things are little more complicated when it comes to Zaidi's view, which in all likelihood could not represent evidence for RP.

Even if some theories which do not involve reverse optics or reflectance recovery can be seen as providing justification for Hilbert's theory, this does not counter Chirimuuta's second worry, namely, that whether one endorses the thesis of *Recovery* is more a matter of

---

[15] Brightness here refers to the total amount of light reflected by an object, rather than its propensity to reflect a given proportion of light.

one's conception of the nature of vision than a matter of what the evidence dispassionately indicates. If this is true, it becomes impossible to argue that there are data which unequivocally speak for this thesis, and thus for RP[16], as the interpretation of the data which favors the thesis requires prior endorsement of one particular conception of vision. In addition, if this conception of vision requires endorsing RP or something similar, it becomes circular to appeal to data 'favoring' the thesis that reflectance recovery occurs in visual processing to justify RP.

I address this worry in chapter 4. To begin with, I will argue that, if it is true that adopting the thesis of *Recovery* is due more to conceptual presuppositions than a straightforwardly empirical matter, then it ought not to be the case that there could be data that would speak directly or indirectly for or against this thesis, independently of one's conception of vision. I then argue that, actually, we already have such data and that Chirimuuta's argument does not go through. To anticipate, I argue that reflectance recovery is usually conceived as an essential part of what have been called *two-stage models of color constancy*. According to those models, the visual system proceeds in a two-step fashion to produce color constancy. First, it computes an estimate of the illuminant and then, secondly, uses this estimate in conjunction with the light signal to estimate the SSRs of objects in the field of view. The first step of this process, which has been called *the albedo hypothesis*, is actually a thoroughly studied hypothesis. Most interestingly for our purposes, the thesis of *Recovery* requires it and would therefore stand or fall along with it. There are thus ways, at least in an indirect manner, to empirically (dis)confirm the thesis that reflectance recovery is a central computation of the visual system. It is not the case that *Recovery* is more of a theoretical posit than an empirical issue.

Each part of the argumentation serves a different purpose. At first, arguing that there are various theories of color constancy which could be seen as justification for some form of color realism such as Hilbert's serves to counter Chirimuuta's claim that, since there is no consensus in color constancy research regarding the status of reflectance recovery, there can be no appeal to the field for support of some form of color realism. Note that this does not mean that I will be arguing for some form of color realism. The point I am interested in is less which color ontology is correct than the kinds of relations may hold between theories of color constancy on the one hand and color ontology on the other. In that respect, considering Hilbert and Byrne's theory is a tool I use to do so. The second part of the argumentation is

---

[16] Or even that other kinds of data speak for RP, either directly or indirectly.

there to hopefully dispel any doubt regarding the empirical status of conceptual assumptions in the field of color constancy research.

Even if I mainly center my discussion around Chirimuuta's criticism of Hilbert's argumentative strategy, this must be seen first and foremost as a pretext to discuss the general relations which stand between on the one hand theories of color constancy and the conceptual assumptions which are brought along with them and varieties of color realism on the other. This work should thus be seen as an exploration of a specific instance of appeals to scientific theorizing to justify ontological theories, as we discussed in the introduction to the present chapter.

This matter is addressed in chapter 5. There, I finally defend a methodological point with regard to color ontology which focuses on color science, more specifically, on color constancy research. The thesis which will be defended is that color constancy theories provide empirical justification for particular claims about color ontology. To motivate my case, I present three distinct theories of constancy and try and show what they could tell us about color itself were they to be respectively correct. This can be seen as a continuation of the work presented in chapter 3, where we focused on the features theories of constancy ought to have to justify reflectance physicalism. I then focus on the question of the exact nature of the link between the color constancy theory on the one hand and the particular ontological view justified on the other. I explain why that link cannot be thought as a logical entailment from the former to the latter. The empirical success of a particular theory of color constancy would represent empirical justification of certain ontological claims about color. I end up by reviewing some objections to the methodological point.

Chapter 2: Chirimuuta's challenge to reflectance physicalism

Introduction

In the previous chapter, we were introduced to the way data from color science has been appealed to in order to motivate certain ontologies of color. Most importantly, I concentrated on reductive color realism or reductive color objectivism, according to which colors are objective, mind-independent, physical properties of the objects we see in our environment. Apples are red, frogs green, wood brown and these facts do not depend on us being able to see those colors. More specifically, we saw that one particular kind of reductive color objectivism has been particularly influential during the last thirty to forty years of discussion. According to reflectance physicalism, colors are simply identical with the preferential ways objects have of reflecting certain kinds of wavelengths over others, that is to say, with their surface spectral reflectances (or SSRs). In the present work, we shall mostly be concerned with this particular form of color objectivism.

In particular, we shall be concerned with the relationship between said objectivism and a particular perceptual phenomenon called color constancy. As we saw, it has been argued that the results of the research on color constancy can be marshalled in favor of reflectance physicalism. Hilbert and, later on, Byrne have advanced the thesis that their view is supported by color constancy research. Careful attention to its results yields reasons to endorse their view. Chirimuuta, on the other hand, very much doubts that this can be made to work. Not only is there no consensus in the field of color constancy research, consensus which would be necessary for Hilbert's argument to go through but it is also not clear that the question regarding whether reflectance is recovered by the visual system is a thoroughly empirical question to begin with.

The purpose of the present chapter is the following. We shall be concerned with Chirimuuta's criticism of Hilbert's appeal to color constancy as support for his theory. To begin with, we shall consider her claim that Hilbert's appeal essentially fails because there is no scientific consensus regarding whether the visual system recovers or is even theoretically able to recover SSRs. She points to results diverging from Hilbert's selected few. In turn, she questions why we observe such a lack of consensus regarding reflectance recovery. She proposes that the latter is observed because researchers interpret the same data in different ways. This divergence of interpretation is then argued to be due to researchers adopting distinct conceptions of the nature of vision. It is finally argued that Maloney and Wandell's conception of vision requires the endorsement of some theory akin to reflectance physicalism, creating a potentially vicious circle.

Hilbert's case

Historically, the initial motivation for reflectance physicalism does not come from science but from commonsense. Hilbert argues that commonsense is "robustly realist about colors" (Hilbert, 1987). However, the realism we finally endorse should not conflict with our empirical knowledge about the physics of light and the science of perception. In particular, color realism should not entail that colors are perceptually unavailable. That is to say, colors should be recoverable visually. And if, as argued, colors just are surface spectral reflectances (henceforth, SSRs), then it becomes a necessary condition of the viability of reflectance physicalism that the human visual system be able to recover SSRs – hence the focus on color constancy research. With this requirement, Hilbert makes his theory empirically falsifiable[17].

In his writings on reflectance physicalism (Byrne & Hilbert, 2004, 2003; Hilbert, 1987, 2008, 2012), Hilbert focuses on a particular branch of color constancy research, namely, computational color constancy. The avowed goal of the latter is twofold. First, it aims at constructing algorithms that the human visual system could actually be using to recover an estimate of SSR (and also to an extent of SPD, although the interest in the latter is less than in the former). Second, it designs algorithms which could aid in developing machine vision. We are here concerned with the former objective.

In order to recover an estimate of SSR, researchers working on computational color constancy have proposed a variety of algorithms. Remember that the light emitted from the object to the eye conflates the information pertaining to the surface (the SSR) and the one pertaining to the illuminant (the SPD, or spectral power distribution). In order to estimate the SSR, the visual system should, in a sense, discard the influence of the SPD on the light reflected to the eye (or, as we called it, the color signal)[18]. In other words, the visual system needs to estimate the illuminant before it can estimate the SSR. However, to do so, the visual system only has access to the cone signals of the scene (maybe across a short period of time) and to the various operations it can perform on them. The problem is, in a mathematical sense,

---

[17] Regarding the question of whether it is empirically confirmable, please see chapter 5.

[18] Of course, the visual system does not need to discard the illuminant in particular the sense that it forgets about it. It should rather use it to compute an estimate of SSR and then might consider it on its own, as information regarding the illuminant might be useful for a variety of purposes.

ill-posed. That is to say, there is no unique solution to it as long as it is not further constrained.

The various algorithms encountered in the literature make different assumptions in order to solve this issue, that is, in order to constrain the problem further. A very common one is that the average chromaticity of a global scene is achromatic, that is, gray, and that this accurately represents the color of the illuminant. This is known as the Gray World assumption (Buchsbaum, 1980). A different assumption is that the brightest patch in the scene stimulates maximally and uniformly the cone channels. That is to say, it is assumed that the brightest patch is white, which yields the name of the "Bright-is-White" assumption (Land & Mccann, 1971). Of course, those assumptions are often violated in natural scenes and many variants thereof exist. Determining which constraints could be most useful to our visual systems in a natural environment is an important part of computational color constancy research.

Yet another famous proposal is that of low-dimensional linear models. Perhaps the seminal paper is Maloney and Wandell's (Maloney & Wandell, 1986). The central idea is that, if the SSR of surfaces and the SPD of illuminants can be described by a weighted sum of a few basis functions, and if those functions are known to the visual system, then the SSRs estimates can be computed exactly. Maloney and Wandell proved that, if the number of photoreceptor classes is higher than the number of reflectance basis functions and if the number of surfaces in the scene is higher than the number of illuminant basis functions, then the SSR and SPD of surfaces and illuminants, respectively, can be recovered exactly[19].

Hilbert concludes that his theory is not falsified by the results of the research on color constancy. SSRs are not unrecoverable and it is thus not forlorn to identify them with the colors we see. Note that in Hilbert's mind the importance of such research for reflectance physicalism cannot be overstated. Were it to be shown that SSRs are not recoverable, this would be equivalent to showing that reflectance physicalism is obviously false. If SSRs were not recoverable, then their identification with the colors would entail that we never see colors. Yet the very fact that we think we see colors is the reason we want to postulate their existence in the first place.

Hilbert therefore argues that Maloney and Wandell's results provide powerful evidence in favor of reflectance physicalism. In a nutshell, the visual system can and probably does recover SSRs and identifying the estimates formed through that process with the colors we see is the very next natural step. As we shall see shortly, Chirimuuta agrees that Maloney

---

[19] For a more thorough explanation of the notion of linear models as applied to color constancy, see appendix A.

and Wandell's theory is linked with reflectance physicalism, however not in the sense which Hilbert has in mind. In addition, she disagrees that this represents the whole story.

First, she argues that those results are not representative of the literature on color constancy as a whole. There is no consensus regarding whether SSRs are recovered. It therefore cannot be assumed that the scientific results of the field as a whole support reflectance physicalism. If this is true, Hilbert's appeal to the research becomes quite wobbly indeed. However, Chirimuuta intends to go further. Not only does she argue that there is no consensus in color constancy research regarding the question of reflectance recovery, but she also wants to propose an explanation as to why there is no such consensus.

Although the color constancy researchers are considering the same sets of data regarding reflectance recovery, they reach different conclusions concerning whether it is an actual step of visual processes because they embrace distinct conceptions of the nature of vision. That is, their reaching a positive or negative conclusion regarding reflectance recovery is more a matter of how they conceive of the nature of perception, more specifically, of vision, than a matter of what the evidence dispassionately indicates. In that case, the evidence or the use we make of it is not neutral and appealing to it in support of reflectance physicalism is problematic.

In addition, some researchers, including Maloney and Wandell, endorse a conception of the nature of vision that entails something like reflectance physicalism. However, if this is true, Hilbert's appeal to their results may be viciously circular. I now turn to the consensus issue.

Reflectance recovery in principle and in practice

In response to Hilbert's claims, Chirimuuta argues that color constancy scientists have not settled on whether reflectance recovery is either theoretically feasible for the human visual system (i.e. whether it is possible in principle) or actually done by the visual system if it is indeed possible (i.e. whether the visual system does recover reflectance in practice). In other words, she centers the first part of her argument around whether it can be shown that there is a consensus, at either a theoretical or practical level, amongst color constancy researchers. I shall first present the theoretical grounds one might have to doubt that reflectance recovery is in principle feasible which Chirimuuta presents and then move on to the "in practice" part of the debate.

The theoretical rationale for reflectance recovery

Chirimuuta's objections to the in principle possibility of reflectance recovery come from both considerations of the challenges faced by linear models frameworks and methodological choices currently made by active researchers. In order to understand the objection regarding the linear models frameworks, it is important to clarify some minor points. As we saw, linear models frameworks constitute a distinct approach to modelling color constancy. It is founded on the idea that one can use a limited amount of weighted basis surface spectral functions to describe SSRs and SPDs. Basis functions do not vary across scenes. That is, the differences between different SSRs (or SPDs) merely reside in how the basis functions themselves are individually weighted. As the basis functions are deemed known by the visual system, knowledge of the respective weights of the basis functions for a given SSR constitutes knowledge of the SSR in question. The same is true for the linear model representing the illuminant.

The issue which follows is that, if the models employed for the illuminant and the surface are too complex, the visual system cannot hope to recover the respective weights of the basis functions. Consequently, the models employed should not have too many basis functions, i.e. have too many degrees of freedom. An obvious question at this point is what it is for the models to be too complex, that is, what is "too many degrees of freedom". According to the important paper of Maloney and Wandell (Maloney & Wandell, 1986), perfect color constancy is possible only if the linear model representing the SSR does not contain more than two degrees of freedom (i.e., two basis functions). If there are more than two, errors will necessarily be made in the estimation of SSRs, preventing perfect color constancy.

Unfortunately, it has been found repeatedly that a linear model describing properly SSRs must have more than two degrees of freedom (Foster et al., 2005; Oxtoby & Foster, 2005; Westland et al., 2000). If those analyses are correct, then the human visual system is not able to use linear models to achieve color constancy and it therefore cannot recover an estimate of SSR by using them. Recovering SSRs, that is, colors according to Hilbert's reflectance physicalism, appears to be an impossible task for the human visual system. Given the identification, we would not be able to see the colors[20] [21].

---

[20] Note that this reasoning does not take into account the fact that color constancy is rarely if ever found to be perfect (but see Granzier et al., 2009; Ling & Hurlbert, 2008). If approximate rather than perfect color constancy is the goal, the requirements on the degrees of freedom may change and the argument may not go through.

[21] However, as Chirimuuta emphasizes, applying linear models to the problem of reflectance recovery is but one of the constraints we might apply. Recovering reflectance through the use of linear models is only one of the

Chirimuuta also argues that many researchers have now abandoned the standard two-dimensional, simplistic paradigm to adopt more complex scenes, either rendered or real. Those scenes contain more cues which could be used to estimate the illuminant. Maloney and his team now use rendered scenes which contain (a simulation of) directional lighting and are three-dimensional (Boyaci et al., 2003, 2004; Maloney et al., 2005; Yang & Maloney, 2001). In contrast, Brainard has advocated for the use of natural scenes, which allow the visual system to access cues we might not even know about yet (Brainard et al., 2003).

I am at a loss to understand why the methodological choice to use more complex scenes, either rendered or real ones and the reasons motivating it, are relevant to this part of Chirimuuta's argumentation. Is it that there are various cues which could be used in reflectance recovery, some of which are still unknown? But this state of affairs can hardly be unexpected. We already knew that linear models were but one possible constraint the visual system might use in its recovery of SSRs. Alternative constraints and alternative cues have been studied at a theoretical level for several decades (Buchsbaum, 1980; D'zmura & Lennie, 2001; Golz & MacLeod, 2002; Land & Mccann, 1971), which seems reason enough to reject this reading of Chirimuuta's objection.

Another possibility is that scientists have now lost interest in linear models and moved on to greener pastures, so to speak. But remember that the general question at issue here is whether there is a consensus amongst scientists as to whether it is possible to estimate reflectance. A move away from linear models by the majority of researchers would seem to indicate that there is such a consensus and that it is that linear models are a poor strategy for in principle recovery of reflectance. This speaks against this reading of Chirimuuta's point as she agrees that many researchers are still concerned with reflectance recovery.

My final option for the interpretation of that point is that researchers are not clear on what cues the visual system uses. But this is neither here nor there. The matter we are concerned with in this section is the in principle debate as to whether the visual system could use linear models to estimate SSRs. As we know that those cues found in more complex scenes could help in estimating the illuminant, it seems that we should not mind yet whether the visual system does actually use them. This sentiment is reinforced by the fact that all the articles mentioned by Chirimuuta in this section of her article are concerned with whether the visual system does use a certain cue, not whether it could use it.

---

way our visual system could proceed. In other words, Hilbert's theory only requires that reflectance be recoverable, not that linear models be applicable. For more regarding this line of thought, see chapter 3.

I do not believe that this last point is of crucial importance to Chirimuuta's case against Hilbert and I therefore advise moving on to the in practice debate as to whether the visual system actually estimates the SPD of the illuminant in order to estimate SSRs.

The actual use of reflectance recovery

Leaving the in principle debate aside, Chirimuuta moves on and presents results which put pressure on the idea of a consensus as to whether the visual system actually tries to recover estimates of SSRs in practice. The focus is on what the visual system in fact does, rather than on what it could theoretically do. Chirimuuta's argument centers on the respective works of two teams of researchers centering around the phenomenon of lightness constancy. Before I go any further, a little introduction to experimental research on lightness constancy is thus in order.

The phenomenon of lightness constancy is the achromatic counterpart to color constancy. Whereas color constancy mechanisms arguably try and recover SSRs of the scene viewed, lightness constancy refers to the constancy of the lightness of objects seen, that is, to whether they display the same total reflectance across distinctly illuminated scenes. In other words, lightness refers to the amount of light, irrespective of wavelength, that is reflected by the object compared to how much light it did receive. For instance, if an object O reflects .4 of the light it received and an object P .8 of the light it itself received, then O is twice darker than P. In contrast, the surface spectral reflectance of an object specifies the amount of light reflected per wavelength of light. Compared to color constancy, the problem of lightness recovery is not necessarily ill-posed.

As we saw, lightness is not related to which wavelength gets reflected, contrarily to color constancy. An object might reflect a lot of light in the middle of the visible range or half that amount in the shorter end of the range and the same in the longer end of the range and still preserve the same lightness whilst having distinct SSRs. It is therefore not dependent on the color of the object in question. A blue object might be as light or as dark as a red object, even though both reflect light in a non-homogeneous manner across wavelengths. What matters here is that they reflect the same amount of light relative to the amount of light they received, i.e., relative to the strength of their respective illuminants. However, as the estimation of the lightness of chromatic objects is a pretty arduous task (consider comparing the lightness of a red patch to that of a blue patch), experiments on lightness constancy restrict their stimuli to achromatic colors (i.e. white through all the grays to black).

The first work Chirimuuta reviews is Zaidi's team's (Khang & Zaidi, 2002; Robilotto & Zaidi, 2004a, 2004b, 2006; Zaidi, 1998, 2000; Zaidi & Bostic, 2008). The latter has worked most importantly with actual scenes rather than rendered ones. In one experiment (Robilotto & Zaidi, 2004a), observers were presented with two scenes, one of which received twice the amount of achromatic light than the other. Two cups wrapped in gray paper stood in each scene, yielding a total of four cups. One of the cups, however, was wrapped in a different paper than the other three, either of a lighter or a darker shade of gray, i.e. of higher or lower (total) reflectance. Observers had to choose which cup was the odd one, i.e. the one with the unmatched paper, by first determining which scene contained the odd object and then identifying which of the cups in the scene had the different paper on. They thus first had to discriminate within an illuminant and then identify which paper was different than the other three across illuminants.

The first task was a simple one of brightness discrimination ("which scene contains cups which do not reflect the same amount of light?") whereas the second one required observers to compare lightness across illuminants ("given that the scene on, say, the left contains the odd object, which one of its two cups has the same lightness as the two cups in the scene on the right?"). Observers could therefore encounter four different kinds of situation settings. The odd object could either be in the less illuminated scene (with half the illumination intensity as the other scene) or in the more illuminated scene and it could either be of a lower or a higher reflectance than the standard objects. For instance, it might have been in the half illumination scene but with a higher reflectance than the standards. In this situation, it would have been brighter than the standard in its scene, that is to say, it would have reflected more light than the standard in its scene. This would have made it more similar in brightness to the standards in the fully illuminated scene.

This observation motivates the first of Zaidi's team's modelling of the possible data. According to a photometer-based strategy, the object chosen as the odd one is the object most different in brightness, more specifically, the object most different in amount of reflected light in absolute terms, that is, unrelative to the amount of light received. The visual system would determine the brightness average of the four cups and pick as the odd one the one furthest from this average. In the situation above, the object chosen would erroneously be the standard in the half illumination scene, which would have reflected a lot less light than the standards in the full illumination scene and a little less light than the odd object in the half illumination scene. Therefore, in situations in which the odd object was of higher reflectance set in half illumination or in which it was of lower reflectance set in full illumination, a photometer-based strategy would have systematically resulted in below-chance level of correct detection of the odd object (i.e. it would have resulted in systematic error). In the opposite situations (higher reflectance in full illumination or lower reflectance in half illumination), it would have resulted in systematically correct detection given correct side detection.

The second modelling of the data also included a photometer-based strategy but added mechanisms of adaptation on top of it. The effect of the latter would here be to diminish the brightness value of a stimulus as a function of the mean luminance of the scene's background. The visual system would start with a brightness value set according to the luminance of the object (i.e. the amount of light it reflects not relative to how much light it receives) and then diminish it as the luminance of the background increases. Depending on the gain of the adaptation (i.e. whether it has a steeper curve or not), the results would range from those of the photometer-based strategy without adaptation (systematically incorrect or correct detection in certain situations, respectively, see above) to the other direction until it reaches perfect identification of the odd object given correct side identification[22].

The final modelling of the data is an inverse optics strategy, according to which the visual system first tries to compute an estimate of the illuminant and then of the reflectance. As we are here dealing with achromatic reflectances and illuminants, this is a more straightforward computation than if it involved SSRs. According to Robilotto and Zaidi, the visual system has enough information to calculate the mean luminance of the background and of the cups and then take the ratios. For instance, let us say that the luminance of the full illumination scene is twice as intense as the luminance of the less illuminated scene, as was done in Robilotto and Zaidi's study. A standard in the full illumination scene would thus reflect twice as much light as a standard in the half illumination scene. After having identified the scene which contains the odd object, the visual system would just have to see which of the two cups in this scene does not stay in such a luminance ratio with any of the other two supposedly standard cups in the other scene. For instance, for a luminance ratio between backgrounds of 1 to 2 (that is, one background reflecting twice as much light as the other), let us say that we have cups A and B in the half illumination scene and cups C and D in the full illumination scene, and B is the odd cup (e.g. a darker one). The visual system might find that the luminance ratio between B and either C or D is, say, 1 to 4 rather than 1 to 2. It would

---

[22] It is unclear to me why Robilotto and Zaidi think that a mechanism of adaptation could have any role to play in their experiments. Their observers were not given the time to adapt as they were frequently switching their gaze to the other scene. Mechanisms of adaptation are usually thought to require some time before they can effectively have an impact on light sensitivity.

then judge accordingly that B is the odd cup. This strategy presupposes that the visual system is able to compute the illuminant ratio between the two scenes[23].

According to Robilotto and Zaidi, the pattern of results they found is best modelled by the second strategy, namely, the photometer-based one with mechanisms of adaptation[24]. A visual system using such a strategy with a certain degree of adaptation (i.e. the gain of the adaptation, see above) would exhibit asymmetries of results that would be of a similar shape and magnitude as the ones displayed by Robilotto and Zaidi's actual observers. The authors performed a similar experiment a couple of years later (Robilotto & Zaidi, 2006) using patterned rather than plain paper on the cups. Observers thus had access to an additional cue compared to the 2004 experiment, that is, they could evaluate the brightness contrasts on the cups due to their pattern of darker and lighter areas. Results were once again deemed to support a photometer-based strategy and led Robilotto and Zaidi to explicitly reject an inverse optics strategy. Zaidi and Bostic (Zaidi & Bostic, 2008) replicated the experiment using colored stimuli and concluded again that a model of reflectance estimation did not accurately describe their participants' results.

As Chirimuuta points out, this is not so when it comes to the work of another central researcher in lightness and color constancy, Laurence Maloney. Maloney and colleagues performed an experiment on the effect of slant on lightness constancy (Boyaci et al., 2003). They use rendered scenes in which simple geometrical shapes were simulated. Viewed binocularly, those scenes were three-dimensional and the orientation of the central test patch could be altered. Other geometrical objects populated the scenes and had varied surface properties (such as shiny, matte or transparent). The illumination involved was a complex mixture of a diffuse element and a punctate one. A punctate illuminant can be exemplified by a spotlight, whose light will have a direction (rather than being diffuse) and thus create shading, shadows, specular highlights and so on. The addition of a punctate element therefore allows the use of many more cues to the illuminant by the visual system.

The observers' task was twofold. First, they would adjust a stick-and-circle gradient probe positioned on the test surface to estimate its orientation. Such an estimation would ultimately allow comparisons between the observer's performance on the subsequent task and his or her orientation settings. This latter task was a sort of asymmetric lightness matching[25].

---

[23] For readers who doubt that this corresponds to a reverse optics strategy, please read on!

[24] For a critical evaluation of that conclusion, see chapter 4.

[25] I specify "a sort of" because I would argue that, normally, an asymmetric matching has both the test and the reference patches embedded in scenes. One might argue that the reference chips were embedded in a black background and that this constitutes a scene but this is at best doubtful. In addition, usual asymmetric matching

After the orientation setting, the probe would disappear and an array of lightness chips would appear on the right side of the scene on a black background. Observers would have to choose the chip which corresponded the most in lightness to the test patch. The chosen chip could either be identical in lightness with the test patch, in which case the identification is perfect, or it could be further away from it, indicating that the observer's lightness identification is less than ideal.

The authors modelled the possible data according to a Lambertian geometric discounting function. In order to understand the gist of the proposal, it is necessary to go through the geometrical notions at play. Various factors are involved in producing the light reflected from a Lambertian surface (a matte surface)[26]. The lightness of the surface (i.e. its achromatic reflectance, that is, its reflectance averaged over wavelengths) will obviously contribute but so will the intensity of any diffuse light source and, if there is a punctate light source, its intensity and its angle relative to the surface in question. As all of those variables stand in relation, it is possible to calculate one if one knows the others. For instance, knowing the lightness judgments of the observers would allow one to calculate their estimates of the other values, and especially, the geometric discounting function, composed of the estimates pertaining to the intensity and angle relative to the test patch of the punctate light source. If the observers' geometric discounting function is similar to the ideal discounting functions (with the real values), then they would appear to be discounting the punctate light source in their judgments of lightness.

Maloney and colleagues did indeed find that their observers' geometric discounting function was similar to the ideal one, though they did misestimate the intensity of the punctate light. However, the observers' judgments of its orientation were nearly devoid of errors. The results of Boyaci, Maloney and Hersh therefore appear to support the hypothesis that our visual system performs reflectance recovery (or *Recovery*) in performing this task.

It therefore seems at first glance that the literature contains both evidence against and for a model of color constancy which incorporates reflectance recovery. There does not seem to be a consensus regarding this matter. But Hilbert's appeal to color constancy research to support reflectance physicalism is then threatened if it requires such a consensus. As Chirimuuta further argues, this lack of consensus can be at least partly explained by the fact

---

has the observer manually adjust a patch and not choose a set of patches with varying lightness. Ultimately, this is not an important point but one that we must still keep in mind when comparing experiments with diverging methodologies.

[26] When presenting color constancy, I called this light the color signal.

that the science on this issue is not as conceptually neutral as we might want it to be and this puts even more pressure on Hilbert's strategy. I now turn to this question.

The conceptual assumptions at play in color constancy research

We have seen that different teams studying lightness and color constancy actually interpret what appears to be, according to Chirimuuta, "fairly similar psychophysical data" (Chirimuuta, 2008, p. 577) in very different ways. On one side, some argue that these data show that observers actually engage in reflectance recovery in certain tasks. On the other side, other researchers have argued that such strategies are incompatible with their results. According to Chirimuuta, this discrepancy in the reading of the evidence originates in different theoretical allegiances pertaining to the nature of visual processes.

As Chirimuuta points out, there are various ways to deal with evidence which is being interpreted in different incompatible ways. One might say that the evidence is ambiguous. Researchers therefore tend to interpret it in a way that is consistent with it and which they favor, preferably, for other theoretical or empirical reasons. At this point, the only thing left to do is to design new experiments to add evidence to the one already existing and determine whether these new data are compatible with all the proposed interpretations of the old data. As this progress of accumulation goes on, the hope is that all but one interpretation will eventually get discarded so that we will discover which researcher, if any, was right to begin with.

However, it sometimes seems that the issue lies not so much with the state of the evidence but with the preexisting conceptions brought to the table by the various researchers. Those conceptions might amount to nothing less than distinct research programs. In that case, relying on more and more data might not alleviate the issue. According to Chirimuuta, this is what we found in the case of color constancy. In order to understand her proposal, we have to understand the core notion of 'research program' as she uses it.

The notion of research programs was developed by Hungarian philosopher Imre Lakatos in an attempt to remedy the flaws of Popper's criterion of demarcation between science and non-science, namely, the falsifiability of a thesis[27]. As interesting as the notion of

---

[27] According to Lakatos, the element of evaluation which should be at play is not a single thesis but what he calls a research program. The latter comprises an ensemble of theses. Some of those theses constitute what he called the 'hard core' of the program: they articulate the fundamental tenets of the program and do not, by themselves, predict any observation. The other theses in the program are 'auxiliary hypotheses' and, joined to the hard core, predict observations which are, if possible, novel and surprising. When an experiment is done, and its results seem to suggest that the relevant research program is flawed, the blame usually is and ought to be put on those auxiliary hypotheses and the latter modified in order to account for the results. The hard core can thus continue

research program as understood by Lakatos might be, it is not the one Chirimuuta proposes to use here (personal communication). We have to understand it by looking at what happens in research departments. When a student finally gets a position as a PhD student in, e.g., a psychology lab, she inherits more than a job. Her supervisors will teach her the tools of the trade, so to speak, and, doing so, they will probably convey to her how it is that a given issue should be solved. More precisely, they will convey to her the necessary methodology to solve this issue. Usually, this methodology will come alongside a certain conception of the problem at hand.

For instance, let us consider a team of researchers working on color constancy and construing it, as we saw in the previous chapter, as a reverse optics problem. According to this understanding, the visual system, in order to be color constant, must solve an inverse problem. That is, from a given light signal (or, equivalently, cone signal), it must find its way back to the real-world properties which were (partly) responsible for this light signal. In the case of color constancy, the problem is construed in such a way that those properties are thought to consist in the specific ways objects have of reflecting light, that is to say, in their surface spectral reflectances (or SSRs).

Once color constancy is understood in this way, a methodology naturally imposes itself. If the visual system aims at solving an inverse problem, then the first step is to identify an algorithm whose output would consist in a specification of the objects' SSRs[28] and whose behavior actually mimics the behavior of our visual system. Once this is done, we need to identify exactly how the visual system could implement said algorithm. Finally, if all went well, one will end up trying to specify where exactly in the visual cortex the algorithm is implemented[29].

As Chirimuuta rightly observes, conceiving of perception in general, and vision in particular, in this way is at the core of what has been called the "Marrian" school of research which originated in the work of David Marr, among others, in the latter half of the previous century. According to this line of thought, perception can be understood as the algorithmic process by which the brain responds to a series of ill-posed inverse problems. A prime example of such problems is to understand how the visual system can construct a three-dimensional representation of the world from a bidimensional array of light hitting the retina.

---

to be developed unless the whole research program starts to encounter an increasing number of insoluble problems and is superseded by another research program which fares better.

[28] Of course, in the process of doing so, the visual system will also have to generate an estimate of the illuminant's properties. More on this later in chapter 4.

[29] Although they may not be exactly co-extensive, note that this understanding of research program is very similar to Lakatos'.

Another example is of course to understand how to move from cone signals conflating information about the illuminant's SPD and the surfaces' SSR to a representation of the latter which does not take into account the influence of the illuminant. What is important here is that adopting such a conception of vision entails a particular methodology. In other words, this conception of vision comes with its own map of accepted and contraindicated methodological moves.

Prominent researchers in color constancy such as David Brainard, Ana Radonjić, and, closer to home, Laurence Maloney, have all sided with the Marrian train of thought in their designing of experiments and in their readings of results (Brainard & Maloney, 2011; Delahunt & Brainard, 2004; Radonjić et al., 2015; Radonjić & Brainard, 2016; Radonjić & Gilchrist, 2021; Rutherford & Brainard, 2002). As we saw, according to Chirimuuta, such train of thought can be equated with a particular research program. In contrast, researchers like David Foster and, especially, Qasim Zaidi have eschewed such program and taken a different route. It is harder to determine exactly the conceptual position of the latter researchers. However, it is clear that they believe that recovery of physical properties from sensory signals is neither a central nor a necessary step in visual processes. In other words, they reject reverse optics approaches.

These scientists instead propose to understand perception as a set of tools designed to help the organism in question navigate its environment. This research program, according to Chirimuuta, is characterized by a few traits. First, it assumes a distinction between sensation and perception. Secondly, it evaluates proposed perceptual strategies according to how well they allow the organism to execute a given task, ignoring whether an accurate representation of the mind-independent world is involved. That is, it is performance-oriented.

The distinction between sensation and perception has been famously defended by Thomas Reid (Reid, 1764). Chirimuuta refers herself to a standard textbook on perception:

> "The study of sensation, or sensory processes, is concerned with the first contact between the organism and the environment. Thus, someone studying sensation might look at the way in which electromagnetic radiation […] is registered by the eye. This investigator would look at the physical structure of the sense organ and would attempt to establish how sensory experiences are related to physical stimulation and physiological functioning. These types of studies tend to focus on less complex […] aspects of our conscious experience. For instance […] how we perceive brightness. (S. Coren, L. H. Wood, and J. T. Enns 2004 in Chirimuuta, 2008)"

And she follows with this quote:

> "[T]he sensory question might be ''How bright does the target appear to be?'' whereas the perceptual questions would be ''Can you identify that object?'' ''Where is it?'' ''How far away is it?'' and ''How large is it?'' In a more global sense, those who study perception are interested in how we form a conscious representation of the outside environment and in the accuracy of that representation. (S. Coren, L. H. Wood, and J. T. Enns 2004 in Chirimuuta, 2008)"

Sensations are thus conceived as the materials used by perceptual processes to create a representation of the mind-independent world. They are usually simple qualities, like brightness. Chirimuuta argues that Zaidi and his colleagues are concerned with the sensory side of the problem. They want to understand how our visual system uses sensory information in order to solve the problems it is presented with. Consequently, they are not interested in whether reflectance recovery occurs later on in the chain, that is, not on the sensory side.

In contrast, Zaidi's team is interested in whether simple algorithms could account for observers' performance in tasks such as their "finding-the-odd-one" one. In an article which replicates their findings but with chromatic cups this time, they announce that "the second purpose of this paper is to identify a simple observer strategy that accounts for both correct and incorrect identifications" (Zaidi & Bostic, 2008). In an earlier article, Zaidi claimed that "the problem for the visual system to solve is not to bring about stable color appearance under different illuminants by discounting them but to recognize that objects are indeed being viewed under different illuminants and to discover what the illuminant properties are" (Zaidi, 1998). Zaidi believes that bringing about "stable color appearance" is only one way of achieving color constancy and a very pricey one, computationally speaking, in addition. It is better to look for simpler strategies the visual system could use to identify objects across scenes. It is likely that the visual system has indeed favored those simpler strategies over more resource-intensive ones.

If it is true that different theoretical conceptions of vision are at play in color constancy research, then it might be asked what influence they bring to the table. According to Chirimuuta, Robilotto and Zaidi's conception of perception influences their interpretation of the data "in such a way that what is a crucial stage in vision in the Maloney interpretation (illumination estimation) becomes a side issue for Robilotto and Zaidi, not really part of the visual process even if it does take place" (Chirimuuta, 2008, p. 579). On the other hand, Maloney and likeminded researchers view vision as nothing but the recovery of object properties, which is not even seen as an essential part of visual processes by Robilotto and

Zaidi. Chirimuuta concludes that those differences "in the conceptualization of vision […] amount to changing the subject or explanandum" (Chirimuuta, 2008, p. 579).

Subsequently, Chirimuuta wants to argue that the same set of data is interpreted quite differently by the teams of researchers at issue and that we observe this difference of interpretation because those teams adopt different schools of thought, that is to say, different conceptions of what vision amounts to. However, it is unclear why exactly she stresses that it is the *same* data that is interpreted differently. In a very trivial sense, the data being interpreted is the data of the respective team's experiment and in that sense is constituted by distinct sets of data. As far as Zaidi's team is concerned, this last statement is admittedly false. Zaidi does indeed comment on Maloney's results and argues that the differences with his data is that subjects tried to do conscious corrections in Maloney's experiments, leading to individual differences (Robilotto & Zaidi, 2004a, p. 792). He therefore does think that Maloney's results are not irreconcilable with his favored framework. In contrast, Maloney and his team do not comment on Zaidi's results. Since they do not, it cannot be said that the *same* data is being interpreted differently.

What I believe can be safely said is that the same *kind* of data is being interpreted in distinct ways. That is, the data of both experiments show a particular pattern of responses. In Chirimuuta's own words:

> Zaidi's and Maloney's groups present somewhat similar psychophysical findings in that both demonstrate that observers have partial lightness constancy. (That is, as with colour constancy, judgment of the lightness of a surface is to some extent stable with respect to illumination level changes, but the stability breaks down in certain situations.) (Chirimuuta, 2008, p. 578)

Chirimuuta's point is therefore that the same pattern of behavior is given different explanations by the two teams of researchers. In the case of Zaidi, this pattern is interpreted as evidence for a heuristics strategy which does not involve reverse optics. In the case of Maloney, however, the same pattern is apparently explained by the idea that the visual system uses inverse optics.

This is already quite problematic on its own for Hilbert's argumentative strategy. If the data we have amassed so far do not simply speak in favor or against reflectance recovery on its own but need additional assumptions, then they can speak in favor of reflectance recovery only if we have independent reasons to believe in those additional assumptions. And it is unclear exactly which reasons we may have to endorse one research program over others given the current state of the research. Both of these explain the data relatively well and it seems arbitrary to opt for one at the expense of the other.

However, Chirimuuta goes even further than this. She argues that, in the case of Maloney and Wandell, their conception of vision is profoundly linked with something very close to RP. More specifically, it is possible that their conception actually entails a theory at least very similar or even identical with RP. If, in achieving color constancy, the visual system estimates reflectances and those estimates correspond to the colors we see or dictate which colors we see, then the identification of color with reflectance is essentially a free meal. As she puts it herself,

> "Interestingly, the theoretical commitments of a scientist such as Maloney map smoothly onto a worked-out philosophical doctrine such as reflectance physicalism. That is, it can be said that the Marrian framework implies colour realism, by positing that, since vision is the recovery and representation of external world properties, there must be a physical property that we see as colour. And this is entirely consistent with Maloney's explicit statements about the objectivity of colour." (Chirimuuta, 2008, p. 579)

However, it then becomes dubious that one can appeal to the success of Maloney and Wandell's work to support RP. That work already embodies a commitment to RP and that commitment appears to motivate a particular reading of the data. If that is true, then the "success" of Maloney and Wandell's work reflects not so much the data themselves as the previous commitment to reflectance physicalism. This appearance of success cannot then be made to work for reflectance physicalism as it seems that one must first embrace RP to then read the data as supporting it.

More formally, it seems that Chirimuuta's reasoning is the following. Consider the claim that the results of Maloney's experiments are evidence that the visual system estimates the illuminant to then recover SSRs. The question then becomes: how would one be motivated to accept this claim? According to Chirimuuta, one can be so motivated only if one already leans towards an inverse optics conception of vision. That is to say, one already accepts (or is tempted to accept) the claim that the visual system's processes are inverse optics processes.

But it is not clear exactly how our visual processes being inverse processes ought to lead us to the claim that Maloney's results are evidence that the visual system estimates the illuminant to then recover SSRs. I believe that Chirimuuta's case for the vicious circularity at hand then consists in the following argument:

1. If our visual processes are inverse optics processes, then the outputs of visual processes are estimates of physical properties.
2. Color is an output of visual processes.
3. Therefore, if our visual processes are inverse optics processes, then color is an estimate of a physical property.

The antecedent of premise 1 can be understood as the core of the reverse optics approach. That visual processes are essentially inverse processes is the central point of the framework. Obviously, the consequent of the premise is that the goal of those processes is the estimation of the physical properties at the source of the signal at hand, that is, the color signal in the case of color constancy. Premise 2 claims that color is not a mere byproduct of our visual processes. It is rather an intended output of our visual processes and therefore stands alongside other outputs such as shape, relative distance, and size.

If color is an estimate of a physical property, then this pressures us to find which physical property it is an estimate of. It then becomes quite tempting to say that Maloney's results are evidence that the visual system estimates SSRs because this yields an answer to that question while at the same time explaining the data quite well. That is, interpreting the evidence in this way answers one of the central questions we need to deal with if we endorse an inverse conception of vision.

Therefore, the endorsement of inverse optics leads one to accept the claim that color is an estimate of a physical property. But such claim is already tantamount to accepting a form of color realism. If color is an estimate of a physical property, then the blues, reds and greens I see are merely physical properties of the objects in front of me. It then becomes apparent that the interpretation of Maloney and Wandell's data in such a way that it supports the claim that the visual system estimates SSRs, since it requires the endorsement of reverse optics, requires a form of color realism. Hilbert's use of the same interpretation to bolster his own form of realism thus appears to imply a circular element.

Subsequently, we cannot see the relationship between Hilbert's theory and Maloney and Wandell's approach as one of simple evidential support of the latter to the former. The alternative picture Chirimuuta wants to motivate is to see (some theory like) RP as a theoretical presupposition of a particular research program in color constancy research (under her conception of research program). Were the research program to fail, by not being able to account for an important number of perceptual phenomena, RP would be refuted, *modus tollens*. In contrast, its success would increase our confidence in the truth of the theory. On that picture, the research program of Maloney and Wandell and similarly minded researchers can be seen as including an operationalization of RP.

Summary

In this chapter, we have seen how Chirimuuta criticizes the notion that there is a consensus which Hilbert can appeal to in order to argue for RP. As she rightly points out, many researchers disagree with Maloney and Wandell's approach to constancy and it is thus incorrect to speak of a consensus regarding the status of reflectance recovery. Zaidi and colleagues, in particular, have starkly rejected the idea that the visual system even cares about the nature of the reflectances in its environment.

This lack of consensus is due, according to Chirimuuta, to the fact that the same evidence is interpreted in different ways because the researchers at issue have different theoretical opinions. The evidence therefore speaks in favor of *Recovery* only if one already adheres to a particular theory about the nature of perception or, more specifically, of vision. Chirimuuta wants to argue against the thesis that color constancy science is suitably empty of this theorizing for appealing to it to justify reflectance physicalism. The research programs of all researchers working on color constancy are loaded with theoretical presuppositions which are akin to philosophical positions such as reflectance physicalism. This nullifies Hilbert's appeal to it to support his own position for fear of vicious circularity. We have seen that Chirimuuta proposes reconstruing reflectance physicalism as a theoretical posit of a particular line of research which stands or falls with it.

In the fourth chapter, I shall examine Chirimuuta's claim that the evidence cannot be interpreted in favor of *Recovery* if one does not already subscribe to a particular research program and the accompanying theory of perception. I will try and show that, actually, the evidence accrued in color constancy research is much more neutral than what Chirimuuta believes by focusing on one particular example and the empirical consequences which can be seen to follow from it. The fifth chapter will be devoted to an exploration of what we could say about the methodology of color ontology if we focus on the link between theories of color constancy on the one hand and the distinct color ontologies (that is, ontological accounts of color) on the other.

However, in the next chapter, I will start by discussing the notion that Hilbert requires a consensus regarding *Recovery* in order to appeal to the science to support his theory. We shall see that it is not clear that he does require such a consensus. The array of perceptual theories he can appeal to is broader than what we may have once thought. Consequently, whether there is a consensus or not regarding *Recovery* is not as important as one may think beforehand. The material presented in the third chapter will represent an example of the methodological point we shall consider in the fifth chapter.

Chapter 3: On the lack of scientific consensus concerning reflectance recovery

Introduction

In the previous chapter, we have been introduced to Hilbert's argumentative strategy regarding reflectance physicalism (hereafter, RP). According to Hilbert, and later on Byrne, there is evidence for RP since the visual system performs reflectance recovery. Since our visual system recovers reflectances (that is, SSRs) and we represent colors in our experiences, the natural conclusion is that those colors we experience are nothing but estimates of the SSRs of objects.

We have also learnt about Chirimuuta's doubts regarding this strategy. First, it is by far not clear that there is a scientific consensus regarding whether the visual system employs reflectance recovery (henceforth, whether *Recovery* is true) in its quest for color constancy. There are many researchers who explicitly eschew strategies involving reflectance recovery (that is, strategies involving reverse optics) and others who propose theories which ignore it entirely. Furthermore, this absence of consensus is partly explained, or so Chirimuuta argues, by researchers adopting distinct conceptions of the nature of vision. Some view vision as the solving of a series of inverse problems and thus as a recovery of worldly properties, as in David Marr's approach, while others conceive it as a set of quick-and-dirty solutions to particular issues presented by our environment, without caring whether those solutions really involve recovering real, "mind-independent" properties.

In the present chapter, I want to discuss the first part of this criticism, namely, the one regarding the scientific consensus about *Recovery*. I will argue that, *contra* Chirimuuta, it is not necessary that there be a consensus to the effect that our visual system recovers SSRs in order for color constancy research to support RP. In other words, a successful theory of color constancy does not have to subscribe to *Recovery* to be in a position to justify the endorsement of RP. This entails that the consensus as to whether reflectance recovery is a strategy used by the visual system, though interesting in its own right, is not a determining issue when it comes to the viability of RP.

I shall first present a famous theory of color constancy, namely, Foster's theory of cone-excitation ratios. According to that theory, the visual system does not estimate non-relational properties of objects (such as their SSR) but a particular kind of relations the objects enter into. After proposing a slightly modified version of the theory, we will see that there is no reason to believe that its success could not represent evidence for reflectance physicalism. This will represent an example of a theory of color constancy which eschews reverse optics but is consistent and even maybe supportive of the identification of colors with SSRs.

Finally, to nuance my position, I will revisit one of the theories discussed by Chirimuuta which explicitly rejects reflectance recovery as a step in visual processing. According to Zaidi's theory, the visual system uses simple heuristics to solve the practical issues emerging from the complexity of the optics of our world. We will try and determine whether it can, not unlike Foster's theory, be seen as support for reflectance physicalism. To anticipate, I will argue that it cannot. In this way, we will be able to see that not all theories of color constancy could be seen as justifications for RP.

The moral of this chapter is that it is wrong to think that a theory of color constancy has to postulate *Recovery* if that theory is to represent evidence for RP, were it to be successful. I argue that this is not the case and that some theories can support reflectance physicalism without endorsing *Recovery*. The distinction between theories that can support RP and those that cannot is orthogonal to the distinction between those that embody reflectance recovery and those that do not.

As will become clear in the fifth chapter, the material presented below represents an example of the methodological point to be described in the final chapter. For the time being, however, I invite the reader to keep in mind that what we are investigating is the evidential relationship there may be between an ontology of color (such as Hilbert's reflectance physicalism) on the one hand and a theory of color constancy (such as Foster's) on the other. Note that, even though it may look like we are defending RP in this chapter, the focus is actually on the relations that may hold between it and various theories of constancy. In that regard, RP and the theories of constancy which may or may not justify evidence are but a pretext.

Talking or not talking past one another

Before I can move on to the heart of the matter though, a small misunderstanding that can arise upon reading Chirimuuta's argument needs to be cleared out. At many points in her discussion surrounding the issue of scientific consensus as to whether *Recovery* is true, Chirimuuta seems to imply that Zaidi's and Maloney's teams actually do not talk about the same thing in their respective works. She states that they are "using different lightness constancy paradigms which, arguably, amounts to an exploration of different problems" (Chirimuuta, 2008, p. 572). In addition, she argues that the differences "in the conceptualization of vision […] amount to changing the subject or explanandum" (Chirimuuta, 2008, p. 579). A weak reading of this would encourage us to see it as hyperbole and discard the idea that there are really different subjects or explananda at play. However, a strong reading would imply that Zaidi's and Maloney's teams are really only talking past each other.

This strong reading would be reinforced by the fact that, as we saw, Chirimuuta mentions the sensation-perception distinction, arguing that Zaidi and his colleagues concern themselves exclusively with the sensation side of the divide. Let us assume that there is indeed such a distinction (or a similar one) to be made within the cortex. Zaidi would be concerned with what happens on the sensation side, whereas Maloney would presumably view reflectance recovery as a process happening on the perception side of vision. However, note that it is then entirely possible for both researchers to be right. Zaidi argues that there is no reflectance recovery on the sensation side and Maloney agrees. Maloney holds that reflectance recovery happens after the processes responsible for sensations have passed their information on and Zaidi does not really care what happens at this point.

However, if this were the case, it would appear to wreak havoc for Chirimuuta's argument to the effect that there is no consensus regarding the status of reflectance recovery. Her argument works only if Zaidi and other similarly minded researchers speak about the same thing as Maloney and other proponents of the reverse optics strategies. If Zaidi does not care whether reflectance recovery happens at a later stage of processing (later than sensation, that is), then he does not really have anything to say as to whether there is a consensus regarding whether reflectance recovery happens at such a later stage. And it then becomes entirely possible and even likely that there is such a consensus, which is opposite to Chirimuuta's opinion.

I believe that we ought to resist this strong reading and assume that it was hyperbole on Chirimuuta's part which led her to assert such strong statements. Therefore, I would argue that we opt for the weak reading of those assertions and assume that she understands the situation as one of theoretical conflict between two groups of researchers, one believing that reflectance recovery occurs at *any* point in the normal visual processing of the brain or is at least a question worth researching and the other believing that it does not happen at any such point and that we have other paradigms worthier of our investment. Having clarified that point, we can move on to the presentation of relational color constancy.

Relational color constancy

The original theory

I would now like to present a theory of color constancy which is distinct from Maloney's whilst also being a potential candidate in our search for theories which can support reflectance physicalism. I want to focus on the notion of relational color constancy, which has been developed by Foster and his colleagues (Amano & Foster, 2004; Craven & Foster, 1992; Foster, 2003, 2011; Foster et al., 1992, 1997; Foster & Nascimento, 1994; Linnell & Foster, 1996; Nascimento et al., 2004; Nascimento & Foster, 2000) upon consideration of the nature of asymmetric matching experiments.

Asymmetric matching involves the comparison and adjustments of stimuli across scenes with spectrally different illuminations. Usually, collections of flat (2D), matte rectangular surfaces of varying sizes and colors, called "Mondrians" are presented across illuminations, either side by side or successively (and thereby testing either simultaneous color constancy or successive color constancy). Their peculiarity resides in the fact that a central patch in the "test scene" does not match in color its corresponding patch in the "standard scene" (i.e. the scenes are identical except for this central patch). The participants' task is to adjust this central patch in the test scene so that it matches in color the corresponding patch in the standard scene. If the adjustment's coordinates in a color space are close to the coordinates that the central patch in the standard scene would have had if it had been placed under the test scene illumination, this means that the participant exhibits a high degree of constancy. In that case, the participant's settings are driven by the (simulated) SSR of the surface, not by the color signal itself.

Foster argues, however, that what is actually measured by asymmetric matching is not color constancy but a very close cousin, namely, relational color constancy. Relational color constancy (RCC) refers to the constancy of perceived chromatic relations across scenes illuminated with different lights or, more generally, across changes of scene configuration. RCC is therefore not or not directly concerned with the constancy of the color of a given surface itself. For instance, imagine a yellow surface neighboring a blue surface under a neutral illumination. If the illumination becomes more greenish, a visual system displaying traditional (i.e. non-relational) color constancy will judge that the first surface is still yellow and the second as still blue (of the respective same shades)[30]. In contrast, a visual system displaying RCC will judge that the first surface is yellower to a certain extent (presumably,

---

[30] Note that talk of the visual system *judging* that something is the case of course requires that we have an analysis of what judging comes down to in this context.

much yellower) than the second surface and this irrespective of whether the illuminant is neutral or greenish.

Foster argues that, in an asymmetric matching experiment, subjects actually use color relations in order to make their matches. That is, asymmetric matching is a test of relational color constancy, rather than color constancy as traditionally understood. Independently of whether this is an accurate reading of asymmetric matching experiments, it must be admitted that the notion of color relations at play in RCC is not entirely clear. We do not appear to explicitly engage in reasoning to determine the nature of color relations between colored surfaces across illuminant changes in those experiments.

Perhaps an intuitive way to understand what they consist in is to look at standard color relations such as x being bluer than y or as red as y. Subsequently, for any given color relation such as these in the left Mondrian, a perfect paper match entails that the right Mondrian displays the same color relation. For instance, if the patch to the left of the test patch in the left Mondrian is bluer than the test patch itself to some extent, then, in a perfect paper match, the corresponding patch in the right Mondrian will be bluer than the corresponding test patch to the same extent.

Though this conception of color relations has the merit of being intuitive, it is unclear that it is the right way to think about them. For one, some of such color relations are much harder to judge than others. Suppose you have to determine whether a light blue is bluer than a dark blue to the same extent than a dark blue is bluer than a much darker blue. This might be easy enough. In contrast, if you have to judge whether a blue is bluer than a brown (sic) to the same extent as another shade of blue is bluer than another shade of brown, it might be too complicated. Secondly, Foster mentions but once such a kind of example to motivate his case (Foster, 2011) and in an unsystematic way. It is thus fair to say that his proposal actually does not rest on such an intuitive framing of color relations.

But we can consider the theory of cone-excitation ratios[31]. Foster argues that cone-excitation ratios could compute the necessary operations linked with RCC. Consider a scene with various objects illuminated by a neutral illuminant. To each surface will be associated a given cone signal. That is, each surface will, through the light it reflects, excite each type of

---

[31] Humans usually have three functional kinds of cones and are thus a member of the trichromatic species. Cone types vary depending on the kind of pigment they contain. Cones sensitive to the shorter wavelengths (or "S-cones") of the visible spectrum have a peak absorption around 420 nm. Cones sensitive to middle-length wavelengths ("M-cones") have a peak at around 534 nm, whereas the cones sensitive to long wavelength ("L-cones") peak at around 564 nm. Note that the absorption curves are neither sharp nor narrow. Even though L-cones are maximally sensitive around 564 nm, they will still respond very strongly to wavelengths a bit longer and shorter than this.

cones to a given extent. For instance, a red surface under a neutral light will excite the L-cones to a very great extent, the M-cones to a lesser extent and the S-cones maybe not at all or to a very small extent. However, if the illuminant is changed, this cone signal will change accordingly. That is to say, the red surface will now excite the different cone channels to a very different extent than it did under neutral illuminant. Cone excitations for a given surface are not stable under illuminant change.

This is different when it comes to cone-excitation ratios. Let us reconsider our red surface under neutral illuminant and presuppose that it neighbors a blue surface. Under neutral illuminant, the red surface will stimulate the L-cones to a much greater extent than the blue surface, which will stimulate the S-cones much more than the red surface will. Let us assume, say, that the L-cones will absorb ten times more light from the red surface than from the blue surface. The cone-excitations ratio for the L-cones between the red and blue surfaces will then be 10 to 1. If it were the opposite for the S-cones, then the cone-excitation ratio for the S-cones between the same red and blue surfaces would be 1 to 10. Finally, let us assume that the ratio for the M-cones between the same surfaces would be 7 (for the red surface) to 1. We thus have defined cone-excitation ratios between two surfaces under a neutral illuminant of 10/1, 7/1 and 1/10 for the L-, M- and S-cones respectively. Note that those ratios are within cone channels and not across them.

As it happens, we empirically know that such kinds of ratios are mainly stable across changes of illuminant. Under randomly sampled daylights, ratios for Munsell chips were found to be invariant to within 4% (Foster & Nascimento, 1994). The same results were found for 640’000 reflectances under regular daylights (Nascimento et al., 2002). Going back to our earlier example, if we were to change the illuminant falling on our red and blue surfaces from a neutral illuminant to a bluish illuminant, the individual excitations of L-, M- and S-cones would change dramatically as a result. But this would not be so for cone-excitation ratios, which would probably remain close to 10/1, 7/1 and 1/10 for the L-, M- and S-cones respectively. Cone-excitation ratios are therefore very stable under illuminant changes and might thus provide an interesting cue for the visual system to consider.

Fortunately, it seems that the visual system does indeed pay attention to cone-excitation ratios. Nascimento and Foster (Nascimento & Foster, 1997) designed an experiment in which two different illuminant changes were presented. The first scene underwent an illuminant change and the small deviations in cone-excitation ratios between surfaces were preserved. In the second scene, the same illuminant change was performed but the scene was then corrected so that the cone-excitation ratios were preserved exactly. Even though the latter scenes corresponded to very improbable events, they were systematically misidentified as containing natural illuminant changes. It can thus be argued that the visual system seems to consider cone-excitation ratios preservation as indicative of illuminant rather than reflectance change.

It is important to note that we do not know with certainty that the visual system pays attention to cone-excitation ratios themselves and not to some other ratios closely correlated with them. For instance, ratios can be determined between cone-opponent channels. A red surface will excite the L-M channel to a certain extent and likewise for the S-(L+M) channel. A blue surface in the same scene under the same illuminant will excite those channels to a different extent, allowing us the computation of a ratio of cone-opponent channels' excitations of (L-M) for the red surface to (L-M) for the blue surface and similarly for the S-(L+M) channel. Because the ratios within cone-channels across surfaces are constant across illuminant changes, any ratio involving only the same within-cone-excitations will also be stable across illuminant changes (Nascimento & Foster, 2000), which ensures that the ratios between cone-opponent channels' excitations are also constant.

Moreover, cone-excitation ratios, whether within cone channels or across cone-opponent channels, can be determined not only between two different surfaces but also between a surface and the chromatic average of the scene. Amano and Foster (Amano & Foster, 2004) showed that such a ratio was stable as long as the chromatic average was the result of averaging 48 different surfaces or more[32]. Cone-excitation ratios may even be calculated over a temporal extent rather than a spatial one (Cornelissen & Brenner, 1995)[33].

We are now in a position to ask ourselves the following question: were the theory of cone-excitation ratios to be successful in accounting for our performances in an array of experiments, could this success be construed as evidence for reflectance physicalism? That is to say, would the success of the theory of cone-excitation ratios (henceforth, CER) speak in favor of RP in the same way that Maloney and Wandell's linear models' success would speak in favor of RP according to Hilbert?

---

[32] Note that, in this study, the surfaces in the scene were also permuted meanwhile the illuminant was changed. Relational color constancy in the case of scenes undergoing only illuminant changes might require a number of different surfaces lower than 48.

[33] Note that this explanation leaves open exactly the relation there may be between RCC and the theory of cone-excitation ratios. Is RCC explained by the theory or is it merely that the latter can adequately explain our performance in asymmetric matching experiments and the need to postulate the existence of RCC is thus avoided? For our purposes, we can ignore this question but see Foster et al., 1997 for more on the relation between RCC, color constancy as traditionally understood and the theory of cone-excitation ratios.

At first glance, it seems obvious that the answer to this question ought to be negative. RP on the one hand states the colors we experience are identical with the reflectances[34] of the objects we see. Those reflectances are not properties whose nature depend on their relationships with other objects' reflectances. They are not relational properties in that sense[35]. The theory of CER, on the other hand, states that what the visual system actually computes, at least in order to perform asymmetric matching, are the color relations between the objects in the field of view, understood as CER. Yet those color relations could not be identified with the nonrelational properties that are reflectances. If the colors we think we see are in fact color relations in disguise, it is impossible to advocate for reflectance physicalism. Where the latter entails that some object we see is, e.g., blue, Foster's theory only entails that this object is bluer than some other object or chromatic mean of the visual field. This appears to entail that the theory of CER is not well-paired with Hilbert's theory.

Relational color constancy, reflectance physicalism and the anchoring problem

Barring any further considerations, the moral of the story should be that Foster's theory, if successful, would not speak in favor of Hilbert's. Nonetheless, this ignores an important addendum to Foster's theory which should be considered anyway. In order to understand the issue, it is important to consider what has come to be known as *the anchoring problem*. The anchoring problem refers to an important problem in lightness constancy research. We therefore have to take a slight detour through this research to understand the issue.

As we saw, lightness constancy refers to the achromatic counterpart of color constancy. A visual system that is lightness constant attributes the same lightness to objects, independently of varying illumination, depth or orientation of the objects with regard to the viewer. Theories of lightness constancy try to explain how our visual system manages to achieve its apparent level of lightness constancy. There are distinct types of lightness constancy theories. However, for our purposes, we can restrict ourselves to intrinsic image models. According to those models, the visual system tries to determine the distinct contributions of the reflectances in the scene, its illumination and the depth and orientation of the objects in that scene. Under that understanding, inverse optics models count as intrinsic

---

[34] I use reflectances and SSRs interchangeably.

[35] It is not clear that SSRs are fully non-relational properties. They may possibly be better construed as relational properties of the surfaces in relation to the light emitted on them. In that sense, they are dispositional properties. An object having a certain reflectance is equivalent to this object being disposed to reflect certain wavelengths in certain ways.

image models and vice versa, since both types of models try to disambiguate the influence of the illuminant on the color signal from the influence of the SSRs present in the scene. That is, the color signal is to be parsed into distinct layers, one of those corresponding to the impact of the illumination and the other corresponding to the reflectances.

However, most of the intrinsic image models (thus including inverse optics models) in lightness constancy research actually reach conclusions regarding the relative luminances or the relative lightness of the surfaces in the scene[36]. That is, their outputs are limited to the differences of lightness or luminance between the surfaces which are distinguished by the models. For instance, the output of a given model could be that a difference of luminance across two surfaces is due to a reflectance change rather than, say, an illuminant change.

But intrinsic images models are then incomplete. We do not only see luminance or lightness differences but also absolute values of lightness[37]. A given surface is not only seen as lighter than a neighboring one but also as a specific shade of gray[38]. Gilchrist, perhaps the first researcher to point out the anchoring problem, puts it best:

> "For instance, consider a pair of adjacent regions in the retinal image whose luminance values stand in a 5:1 ratio. This 5:1 ratio informs the visual system only about the relative lightness values of the two surfaces, not their specific or absolute lightness values. It informs only about the distance between the two gray shades on the phenomenal gray scale, not the specific location of either on that scale. There is an infinite family of pairs of gray shades that are consistent with the 5:1 ratio. For example, if the 5 represents white, then the 1 represents middle gray. However, the 5 might represent middle gray, in which case the 1 would represent black. Indeed, it is even possible that the 1 represents white and the 5 represents an adjacent self-luminous region. So, the solution is not even restricted to the scale of surface grays." (Gilchrist et al., 1999)

---

[36] There is a difference between the luminance and the lightness of a surface. A surface can be very luminous ("bright") but very dark if it is illuminated by a sufficient amount of light and another can be very light (e.g. a shade of light gray or white) but illuminated by a very small amount of light and thus be categorized as not bright at all. In contrast, relative lightness and relative luminance actually parallel each other. Since luminance is the product of illumination and reflectance, taking the ratio of luminance entails dividing the illumination by itself, which actually yields the equivalent of a lightness ratio. Luminance ratios are commensurate with lightness ratios.

[37] Note that it is unclear exactly how Gilchrist intends to motivate this. Is it supposed to be an appeal to phenomenology? Or are there certain experiments which show that we have experiences of absolute values of lightness? It is likely that both would yield the same result. A very simple experimental design to test this would be to present to subjects scenes which have the same lightness ratios but distinct absolute lightness. If subjects can distinguish the two scenes, it must be because of a non-relational property of the scene since all the ratios are preserved.

[38] Note that this construal of lightness constancy is actually ambiguous. On one hand, lightness might refer to one of the three color dimensions along which surfaces can vary, amongst hue and saturation. Or it can refer to the specific shades of achromatic color between white and black. It is not clear that this distinction is of much import to our discussion and we will thus not need to adjudicate between them.

Luminance or lightness ratios constrain what the lightness of the surfaces could be. However, they do not constrain what each and every lightness value could be but what some values are, given the others. Another constraint or rule must then be provided so as to "fixate" the luminance or lightness difference patterns (i.e. the patterns resulting from all the luminance or lightness ratios) on a lightness scale. In that way, a particular ratio can be seen to rise from, say, a light gray and a black or a white and a dark gray. Gilchrist called this particular constraint *the anchoring rule*.

An anchoring rule specifies at least one point of correspondence between luminance values in the retinal image and some absolute lightness value. In that sense, aided by the whole ensemble of luminance or lightness ratios, the visual system can determine the exact lightness value of each luminance value. There have been various anchoring rules proposed in the literature[39]. The highest luminance rule, for instance, states that the highest luminance in the field of view is assigned the lightness value for white. It is then possible to specify the value of the other luminance points. For example, we know that white emits 30 times more light than black, if both are under the same illuminant. If we were to find two patches standing in a 30-to-1 luminance ratio (under the same illuminant), we would be in a position to assign black to the patch emitting the lowest amount of light. If, in contrast, we were to find two patches standing in a 300-to-1 luminance ratio, we would know that the two patches are not under illuminants of the same intensity. Other rules include the average luminance rule, according to which the average luminance of the display is gray. This rule is related to the previously mentioned Gray World hypothesis. Once the average luminance is related to the average lightness, that is to say, gray, all the other lightness can be determined using luminance ratios.

Independently of which rule we opt for, if we acknowledge that lightness theories do need an anchor in order to account for all lightness-related phenomena, it is possible to run a very similar argument in the case of color constancy. In the same way that we do not only have perceptual experiences of lightness differences, we do not only have experiences of color differences but also of absolute colors. If we restrict ourselves to intuitive examples for the moment, we can judge not only that this patch is, say, redder than this other patch but also that the first is red and the second orange. However, Foster's theory of CER is able to account only for the first ability[40].

---

[39] Note that it is an anachronism to call those 'anchoring' rules although this is indeed what they are.

[40] Note that this argument may be construed as a phenomenological objection to Foster's original theory. In a nutshell, one could argue that it is a phenomenological datum that we have experiences of absolute colors (and

If the above reasoning is sound, then the theory of CER is at best only one piece of the puzzle[41]. If we want to hold that it explains our performance in all color-related experiments (and Foster surely wants to (Foster, 2003, 2011)), then the theory will anyway need an anchoring rule at some point[42]. Let us now imagine that one anchoring rule has been provided which fits quite well Foster's theory and is well corroborated by the data at hand. We would be in a position to ask again the question as to whether that theory, now supplemented with the anchoring rule, could be seen as supporting reflectance physicalism, were it to pan out experimentally.

I believe it can be argued that the success of the theory of CER, supplemented with a suitable anchoring rule, could be seen as evidence in favor of reflectance physicalism. This could be the case, even if the theory involves neither estimates of SSRs nor estimates of the illuminant's SPD. However, some may argue that it is unclear why a theory that does not involve reflectance theory can be said to support reflectance physicalism, which explicitly states that the colors we have experiences of are nothing but the reflectances of objects. We can motivate this case by reflecting on an analogy.

Consider the case of temperature measurement and the various types of thermometers there are. Although all kinds of thermometers measure temperature, they do not function in the same way. Thermal expansion thermometers use the expansion of certain substances to measure temperature. This type includes the alcohol thermometer and the well-known mercury-in-glass thermometer. Resistance thermometers, in contrast, use the electrical resistance of a metal. As the electrical resistivity of metals (more specifically, of platinum, nickel or copper) varies with temperature, it is possible to determine the temperature of an area by measuring the particular resistance of a given piece of metal at a given time. Finally, infrared thermometers use thermal radiation emitted by an object to determine temperature. Thermal radiation is electromagnetic radiation produced by the thermal motion of particles. This is but a subset of all the types of thermometers which have been invented.

---

not or not only of related colors). A theory of constancy failing to address this datum would be at the very least incomplete. Subsequently, the theory of CER would be at least incomplete. If we had empirical data on this matter (perhaps following the idea expressed in footnote 37), the case would be quite strong indeed.

[41] In the fifth chapter, we shall see what happens for color ontology if the above reasoning is not sound.

[42] Note that the theory also does not have the ability to recover information regarding the nature of the illuminant. Since cone excitation ratios result from chromaticities differences and chromaticities confound the influence of the illuminant and the influence of the reflectance, a visual system displaying only relational color constancy cannot tell whether the illumination is strong or weak on a particular surface, whether it is bluish or yellowish, etc.… This is unfortunate, as many researchers (including, as it happens, Zaidi (Zaidi, 1998, 2000)) have claimed that information pertaining to the illuminant is especially valuable to navigate one's environment. Providing an anchoring rule would also help with that issue, since it is possible to determine the intensity of an illuminant from the lightness of a surface and the color signal.

The conclusion to be drawn here is the following. Thermometers of different kinds use different types of information. However, even though they function in vastly different ways, they are all measuring the same thing, i.e., temperature. Yet they are measuring it by using different *proxies* of temperature. Thermal expansion thermometers use the expansion of a liquid, a correlate of the heat (and pressure) of this liquid. Resistance thermometers use electrical resistivity which is directly affected by the heat of the metal in question. Infrared thermometers, finally, measure the product of an object's heat, its thermal radiation.

I take it that it is not up for discussion whether thermometers can be correctly said to measure temperature. If this is true, we can transpose the example in the case of reflectance physicalism and its theoretical relation with the theory of CER. As we saw, the latter does not include a direct estimation of SSRs. However, we are now in a position to determine whether this is a real issue by using our little exposition of the various kinds of thermometers.

To take the most intuitive example, a mercury-in-glass thermometer does not directly measure temperature *per se*. Rather, it directly measures the expansion of a given liquid, expansion which is directly correlated with the temperature of matter. In the same way, it could be argued that the theory of CER, supplemented with an anchoring rule, is measuring a *proxy* of surface spectral reflectance. Rather than telling us directly about what SSR a given object has, it tells us that it has such-and-such color differences which, when conjoined with an anchor (such as the highest luminance rule), constrains what SSR the object could have[43].

If the above argument is sound, then its conclusion is of import. If the theory of CER, supplemented with an anchoring rule, is able to estimate reflectances through the use of proxies, then this means that a theory does not have to incorporate reflectance recovery (as exemplified in linear models, for instance) for its success to count as evidence for reflectance physicalism. Having discussed a theory which originally eschews reverse optics and seen that it is not necessarily incompatible for that reason with reflectance physicalism and can hypothetically represent evidence for it, it would be very advisable to consider in turn a theory which would prove more difficult to reconcile with reflectance physicalism. If we can do so, we make sure that our thesis is not trivial by showing that not all theories of color constancy could represent evidence for this ontology. Before we do that, however, we have to address some possible objections to our portrayal of the theory of CER as such potential evidence.

---

[43] Of course, one immediate objection would be that the algorithm could not tell you exactly which SSR an object has but rather to which metameric class of SSR it belongs. This is not an issue for Hilbert and Byrne since their theory only requires that colors are metameric sets of SSRs.

Objections and replies

Interpreting Foster's theory (and adding an anchoring rule to it) so that its empirical success could be seen as evidence for reflectance physicalism is not a flawless strategy. There are a variety of objections which need to be addressed. First, one may worry that the analogy between thermometers and color vision supposed to motivate our case is not adequate. If we push it enough, the analogy will break down and we will lose an intuitive motivation. Secondly, a more conceptually-minded thinker might object that Foster would actually reject the proposed modification of his theory and steadily hold that we do not perceive colors in any absolute sense. We are under the illusion that we do and the only things we really perceive are color relations. A final objection concerns the notion of measurement at play when we use phrases such as "measuring reflectances". Whereas Maloney and Wandell and other researchers have given us a way to understand the notion of reflectance recovery or reflectance estimation, it is not clear that the same is true of the notion of measurement of reflectance. I address these objections in turn.

The analogy objection

One may object to our argument on the basis that there is a disanalogy between the case of thermometers and temperature on the one hand and the case of theories of color constancy and surface spectral reflectance on the other[44]. The objection starts with the observation that we accept that distinct kinds of thermometers can be said to measure the same phenomenon, i.e. temperature, because we already have made up our minds on the relationship between mean kinetic energy and temperature. That is, we already believe that temperature is reducible to mean kinetic energy and this is why we accept that the thermometers measure the same phenomenon. However, we do not have such an agreement regarding color and SSR. We therefore cannot use the distinct types of theories of color constancy to argue for the identification of SSR with color.

Note that this objection assumes that we must have identified color with SSR in order for the analogy to function. However, this is plainly not true. The analogy does not rest on identifying color with SSR. The current proposal does not even have to mention colors in order to be successful. What is required is that Foster's theory embody an algorithm which measures SSR. This step of the argument does not even need to add to the theory that the output of the algorithm be cast in terms of color descriptors.

---

[44] Thanks to Derek Brown for pointing out this objection.

The disapproval objection

A minor objection to our proposal to revisit Foster's theory would locate the issue at the level of the anchoring rule. One could argue that Foster would reject the addition of an anchoring rule to his theory, maintaining instead that the latter can explain the entirety of the phenomenon of color constancy. However, without an anchoring rule, it will not count as measuring or tracking reflectance, only reflectance differences and would not support the present argument that there is more than one way for a theory of color constancy to support reflectance physicalism.

An easy answer would consist in holding that it does not matter whether the theory of CER should or should not be supplemented with an anchoring rule. What matters is that it can. Insofar as this is accepted, it can be argued that the theory with an anchoring rule is capable of supporting our argument. A more sophisticated answer, however, would rest on the fact that Foster himself believes that the theory is not a complete theory anyway and that adding an anchoring rule to it is thus not such a dreadful thing. Here he writes, comparing distinct photographs:

> "For example, in Fig. 3, the scene is illuminated by different daylights, with correlated color temperatures (a) 17,000 K, (b) 4000 K, (c) 6500 K, and (d) 4000 K. The color of the light reflected from the sphere in the bottom left corner in a–c is clearly different. Nevertheless, […] it can be seen that the sphere has the same or similar surface color in each image by comparing it with the nearby foliage and by looking over each image as a whole. By contrast, in d, although the color of the light reflected from the sphere is the same as in a, it can be seen that the sphere has a different surface color, now more bluish, again by comparing it with nearby foliage or over the image as a whole. In a–c, the perceived relations between the colors are largely preserved, and in d, they are not." (Foster, 2011, p. 680)

In that citation, Foster does not appear to endorse the heavy phenomenological thesis that what we see are actually only color relations, and not, as we may naively think, colors in a non-relational sense[45].

Even if the current answer to the objection is admittedly lacking, in the fifth chapter, we shall be considering which ontology of color can be motivated by Foster's theory when it is not supplemented with an anchoring rule to more seriously consider the objection.

---

[45] However, note that it is always questionable whether we ought to interpret such quotes in such a way. It is not clear that Foster is here doing serious phenomenology.

The measurement objection

The final and fiercest objection concerns the notion of measurement of reflectance used in the modification of Foster's theory. This notion is meant to apply to both the latter and Maloney and Wandell's theory. It is argued that both theories include an algorithm which can be said to *measure* reflectance and it is because they perform such measurement that they can both be seen as potential evidence for RP[46]. Consequently, this notion does a significant part of the work in the argument. However, one may object that it is too unclear to successfully do that work. Whereas we do have a relatively explicit notion of what it is for an algorithm to recover reflectances (thanks to Maloney and Wandell's work, amongst others), the same cannot be said for the notion of measurement. What we do have so far is an intuitive understanding based on the proposed analogy with thermometers and their diversity in functioning. Nevertheless, it is not enough to justify seeing the revisited version of Foster's theory as measuring reflectance in one way or another.

According to the objection, the unclarity of the notion of measurement has at least two dire consequences for the argument. First is the already mentioned lack of precision of the notion, which implies that some may find it undetermined whether Foster's revisited algorithm can be rightly said to measure reflectance. The second worry is related to the first. If it turns out to be vague enough, the notion of measurement will allow virtually any theory to count as potential evidence for RP, thus undermining the appeal of a thoroughly empirical argument for RP[47].

To answer the objection, then, is to provide a list of criteria which will allow (a) to determine whether Foster's theory can be rightly seen as *measuring* reflectance and (b) to clearly distinguish between theories that can be seen as such and those that do not. I believe that one way to provide such a list is to draw on Akins' desiderata for a sensory system to count as a veridical system (Akins, 1996)[48]. A veridical sensory system can be thought of as a sensory system which tracks, detects or *measures* the occurrence of a given external event or property. Even though measurement is a notion which applies to systems which do not count as biological (such as thermometers), Akins' portrayal of veridical sensory systems can actually also be applied to artificial systems.

---

[46] Indeed, our reconstruction essentially invites the conclusion that it is because Maloney and Wandell's theory involves a measurement of reflectance that it is able to represent potential evidence for RP, not specifically because it involves reflectance recovery *qua* reflectance recovery.

[47] On whether the argument ought to be seen as empirical, please see chapter 5.

[48] Note that Akins presents this conception of veridical sensory systems in order to show that not all of our sensory systems can be characterized as veridical and maybe none even can. In her own terminology, biological sensory systems are *narcissistic*. That is, they care about interpreting the sensory signals in a way which fits the needs of the organism at issue, even though this may entail distorting or misrepresenting those signals.

In the course of assessing naturalistic theories of intentionality[49], as presented especially by philosophers such as Paul Churchland, Dennett, Dretske and Fodor, Akins discusses the relationship between the intentional nature of some of our mental states and the alleged directedness of our sensory systems. To do so effectively, she proposes to examine sensory systems which count as veridical. The latter can be rightly said to detect, track, or measure a property of the environment that is relevant for the organism. In order to count as veridical, a sensory system should have three distinct characteristics.

First, a veridical sensory system should display, in Akins' terms, *constant correlation*. That is, its deliveries should reliably indicate the things it is concerned with. As she puts it, "if a signal is to be informative ("tell the truth"), it must be produced when and only when a particular stimulus (or stimulus set) is present" (Akins, 1996, p. 343). A system firing in a haphazard way would fail to reliably indicate the presence of a relevant stimulus.

However, a sensory system could exemplify constant correlation but fail to register the structure of the stimuli domain. Taking Akins' thermoception example, a thermoception system could send a signal S for a given temperature and T for a higher one but wrongly order them so that it would signal that S is higher than T when it is not. This system would embody constant correlation, that is, each one of its deliveries would be correlated with a given temperature, but it would still fail to correctly map the structure of the stimuli domain, in this case, the temperature range. So, a veridical sensory system must respect the ordering of the relevant domain. This is what is meant by Akins' requirement of *structure.*

The third and final requirement of Akins' proposal is that of *servility*. In her own words:

> "In the service of the brain, [the sensory systems] toil tirelessly to report the "what, when, and where" of the world's events. They do not interject their own opinions; they do not slyly skew the information to reflect their own interests or prejudices. Their job is to state the facts."(Akins, 1996, p. 344)

In other words, veridical sensory systems should be as neutral as possible. Their reports should concern external properties only, nothing more, nothing less. That is to say, a sensory system signalling more strongly for external properties which are dangerous to the organism would fail to be servile in that sense.

Taking stock, a particular signal of a sensory system should be sent when and only when a particular external property or event is present, should correctly mirror the relations

---

[49] Intentionality is the property of (some) mental states to be *about* something. My belief that a lion is chasing me, for instance, is about a lion and is in that sense an intentional state.

present in the stimulus domain and should not contain anything "extra", not relevant to the property or event tracked[50]. It is easy to see that a thermometer is something which tracks a property in that sense. A well-functioning thermometer should assign a reading to one and only one temperature. It should not tell you that 34° C is lower than 23° C. Finally, it usually does not embroider facts by telling you that the current temperature is a dangerous one or one nice for a walk outside. So Akins' analysis can be applied both to sensory systems and to artificial instruments.

With that understanding of veridical sensory systems and thus of systems which can be said, or so I argue, to measure a given stimulus domain, we are in a position to determine whether the theories of color constancy we discussed can be thought of as including a measurement of reflectance. Not only can we see whether Foster's theory's fall under those criteria, we can also consider Maloney and Wandell's. Reflectance recovery can indeed be seen as one way to measure reflectance and it is thus interesting to determine whether it falls under this conception of veridical sensory systems.

Remember that Maloney and Wandell's theory has the color signal decomposed into two sets of weighted functions, one for the reflectance contribution to the color signal and the other for the illuminant contribution. Each set is thus a sum of distinctly weighted *basis functions*. Those basis functions are stable across scenes and are assumed known by the visual system. They represent the dimensions along which SSR encoding can vary. That is to say, SSR differences will be encoded as differences in the weights attributed to the basis function. A given SSR might be encoded as having, say, twice the quantity of one basis function as another SSR. The number of basis functions represents the degree of freedom of the model. Maloney and Wandell's theory rests partly upon the idea that natural SSRs and SPDs can be adequately modelled by models with very few basis functions. In other words, most of the variance found in natural SSRs and SPDs can be captured by models with very few basis functions[51].

We can thus ask how Maloney and Wandell's algorithm fares when it comes to the three criteria above. Would a sensory system implementing this algorithm be rightly regarded as veridical, that is, as tracking, detecting, or measuring a property of the environment? Perhaps the easiest criterion to deal with here is the one of servility. A sensory system which

---

[50] Of course, Akins' conclusion will be that some sensory systems, including thermoception, do not fit those three criteria.

[51] As we saw in chapter 2, even though models with very few basis functions (as low as two or three) can adequately represent most of the variance of naturally occurring SSRs and SPDs, it is not clear that it is enough to recover *enough* (sic) of this variance to reach observed levels of color constancy. For a more systematic explanation of linear models as applied to color constancy, see appendix A.

is servile, understood according to Akins' discussion, is one that neither adds anything to their report regarding external properties, nor skew the information according to the organism's interests. I take it that it is clear that a system implementing Maloney and Wandell's algorithm would not report anything else but estimations of (sets of) SSR and SPD, expressed as weighted sums of basis functions. None of the estimations are singled out as special or of particular relevance to the organism. Therefore, the system would count as a servile one.

Regarding the criterion of constant correlation, Maloney and Wandell's algorithm would not be able to assign one unique signal to, say, one unique reflectance. However, this is a consequence not of the peculiarities of the algorithm itself but of the fact that the system in question uses only three sensors (that is, the three human cones) to recover a multidimensional value. It is thus subject to the issue of metamerism. That is, two different SSRs which excite the cones in the exact same way will be assigned the same estimate, regardless of the differences in SSR. Consequently, the algorithm could only at best assign one unique value to one set of metameric SSRs. It could be argued that this is not enough for the algorithm to fulfil the criterion of constant correlation. In order to count as reliable enough, a sensory system should not conflate distinct stimuli.

However, note that this may be setting the bar too high. Any system, whether biological or artificial, is going to function better in certain contexts rather than others and will "miss" certain values the relevant stimulus can take. The typical mercury-in-glass thermometer cannot operate in temperatures below roughly -37°C and will thus conflate all the lower temperatures. Similarly, there is no device which is sensitive to all possible wavelengths but that does not imply that a spectrophotometer is not a device which, *stricto sensu*, measure wavelength distribution. As we saw, even Akins admits that a veridical sensory system is one which produces a signal "when and only when a particular stimulus (*or stimulus set*) is present" (Akins, 1996, p. 343, my italics). For this reason, I will assume that Maloney and Wandell's algorithm can be rightly seen as fulfilling the criterion of constant correlation.

In order to determine whether it also fulfils the *structure* criterion, we first have to understand what it is to model the structure of reflectances. As SSRs are distinguished by the amount of light they reflect as a function of wavelength, it would be quite natural to assume that to model their structure is to order them according to what amount they reflect of certain wavelengths rather than others. For instance, a given reflectance could be modelled as containing more longer wavelengths than shorter ones, or vice versa. This would amount to ordering them based on a within-individual-SSR basis. Or it could be assumed that to model their structure is to order them according to the amount of certain wavelengths they reflect in comparison with other reflectances. For instance, reflectances could be grouped according to how much longer wavelengths they reflect. Reflectances favoring longer wavelengths would be modelled as closer to one another than other reflectances favoring shorter ones. This would order reflectances on an inter-SSRs basis.

Independently of our choice, it is interesting to note that Maloney and Wandell's theory already has a very intuitive way to model the structure of reflectances. A sensory system implementing their algorithm could simply order reflectances according to the weights they assign to the respective spectral basis functions. That is to say, given a system with, say, three basis functions, the structure of reflectances would be mapped by the location of those reflectances in a three-dimensional space. The closer the coordinates of two distinct SSRs, the more similar those SSRs would be. SSRs would be ranked according to the weights they attribute to each basis function and their similarity and differences would consist in the differences of weights attributed[52].

This way of structuring the space of reflectances would be much coarser than ordering them based on the proportion of light they reflect as a function of wavelength. The latter would assign one location to each SSR in a space with as many dimensions as there are of wavelengths[53]. However, this is obviously overkill. Naturally occurring SSRs and SPDs do not have very sharp peaks and troughs. They are mostly well described by a gentle sinusoidal. For this reason, it would be enough to order them according to their rough tendency: more longer wavelengths than shorter ones, roughly flat, high amount of middle wavelength and not much of the others, or two peaks at the extremes and not much happening in the middle range could all qualify as structural attributes of the encoding of SSRs and SPDs in the visual system. For this reason, it can be assumed that Maloney and Wandell's algorithm correctly maps the structure of the SSR (and SPD) domain.

This algorithm can thus be seen as measuring reflectance, in that particular case, through reflectance recovery. But can the same be said for Foster's theory? To answer this question, we also need to determine whether it fulfills those three same criteria. Beginning with *constant correlation*, note that this is where the distinction between Foster's original theory and the version which would include an anchoring rule becomes important. First,

---

[52] Note that this ordering would be made on an inter-SSR basis. To order them on a within-individual-SSR basis, two distinct surfaces would be grouped closer if and only if they display the same numerical relationships between their respective weights. For instance, other things being equal, two surfaces having twice the weight of the first basis function compared to the second basis function's would be grouped closer together than surfaces which do not exemplify that relationship between their respective basis functions' weights.

[53] Technically, an infinite number of dimensions.

consider the original theory. It only attributes relative properties such as *being twice as short-wavelengths-emitting as x* or *emitting ten percent less longer wavelengths than y* to surfaces in the field of view[54]. In other words, no absolute color property such as *being green* is attributed to anything.

Such a strategy cannot fulfill the criterion of *constant correlation*. To see this, note that, to fulfill it, a sensory system needs to attribute a given signal to one and only one stimulus. But note that there is no one signal which will always be attributed to one single object if we follow Foster's original theory. Given an object in a scene, many relative properties will be attributed to the object. For instance, let us say that the properties *emitting twice as much middle-wavelength light than its neighboring surface on the left* and *emitting half the amount of shorter wavelengths than its neighboring surface on the right* are attributed to the object. Those attributions will remain if we only change the illumination falling on the whole scene. However, some changes will imply that a whole other set of properties will be attributed to the object. For instance, if we then shine another light with a different SPD on half the scene (covering, that is, the central surface and only one of the two neighboring surfaces), the system will stop attributing one of the two properties above. In the case in which we change the whole environment of the object, the properties attributed to it will change massively. Consequently, there is no one and only one property which is stably attributed to the object's surface. If this is the case, then Foster's theory cannot fulfill *constant correlation*.

This is not so when it comes to his theory when supplemented with an anchoring rule. Whereas the original theory provides a map of color differences, the anchoring rule ensures that this map is *anchored*, that is, located on an absolute color space. A consequence of this is that absolute color descriptors (such as red, green, blue) can now be attributed to the distinct surfaces seen. As long as there is a sufficient variety of information regarding color differences and the system can and does apply an anchoring rule[55], the algorithm can tick the 'constant correlation' box.

This is also true of the *structure* requirement and much easier to see. Foster's theory, even in its original form, will map reflectance differences as cone-excitation ratios. That is, the algorithm will use cone-excitation ratios as proxies for SSR differences. Those differences

---

[54] Note that Foster's original theory will actually use proxies of those properties expressed as some ratios within cone channels as we saw above.

[55] This is just to say that there are conditions under which Foster's revisited algorithm will not operate correctly. Note that this is true of any theory of color constancy. This is why we experience color constancy breakdown or illusions. Note that, in chapter 5, we shall consider Foster's original theory differently. It will be plain that it is not clear that it cannot also fulfill this criterion.

can then be situated on an absolute space using the anchoring rule. Foster's original theory thus maps the differences in SSRs whereas the anchoring rule determines the absolute position of those differences, similar to how we would fix a vector of constant length in a mathematical space. This will not cover differences between metameric surfaces but, as we saw, this is a non-issue. What matters is that the space of SSR differences will be mapped by the space of cone-excitation ratios.

When it comes to *servility*, one may argue that a sensory system employing Foster's algorithm will actually be informing the rest of the brain about the sensory system's own idiosyncrasies. That is, the output of the computation is actually information regarding the receptors' respective excitations rather than information about the state of the outside world. In turn, it could be held that this implies that the sensory system in question is not servile enough. It does not disinterestedly report " 'the "what, when, and where' of the world's events". Rather, it is concerned with itself.

This objection goes astray in an interesting sense. In order to report anything about the outside world, a sensory system will have to encode this information in one way or another[56]. For instance, a system which must detect the occurrence of one of two events could assign distinct neural firings for each event. One of the events would be assigned, say, a rapid firing whereas the other one would be assigned a much slower one. In that way, the system could distinguish between the two events and tell the rest of the nervous system about it. But the same objection can then be presented to that sensory system too. The objection would have sensory systems inform the rest of the organism about something but without encoding that information in any biological form. But this is obviously an impossible task. Note that even the infrared thermometer encodes temperature differences as differences in infrared emissions which are themselves going to be encoded as, say, bits of encoding in the numerical device itself. Yet we still consider that thermometer a paradigmatic case of servility. So Foster's algorithm can and should be regarded as servile too.

The above reasoning shows that both Foster's and Maloney and Wandell's theories include an algorithmic step which is, *stricto sensu*, a measurement of some of the external world's properties. Therefore, the objection does not go through. We have an understanding of the notion of measurement which is clear enough to allow us to determine whether a given theory counts as involving a measurement of reflectance.

---

[56] Note that it is even more general than this. In order to *do anything*, not just report about the outside world, sensory systems must assign distinct signals to distinct things, be they the outside world's events, bodily states, or, in Akins' terms, narcissistic concerns.

So far, it appears we have managed to find a theory of color constancy distinct from Maloney and Wandell's which can *prima facie* be seen as embodying an algorithm that measures reflectance. The theory of CER does not recover reflectance in the sense at issue in Maloney and Wandell's theory. We can thus conclude that there is more than one kind of theory of color constancy which can be appealed to as potential justification for reflectance physicalism. However, it would be useful to determine whether we are being too lax with our criteria, potentially allowing any theory of constancy to count as such justification. Below I discuss Zaidi's approach to color constancy and try to show that it is much harder if not impossible for it to justify reflectance physicalism.

Zaidi's approach to color constancy

The rejection of reverse optics and a performance-oriented view of color constancy

Whereas it has been relatively straightforward to determine whether Foster's theory of CER can be classified in the same way as Maloney and Wandell's, that is, as involving a measurement of SSR, the same cannot be said for the theory of Qasim Zaidi. Perhaps it is easier to start with what this researcher eschews rather than what he favors. To begin with, Zaidi has repeatedly stated that he eschews a reverse optics approach to color constancy and has arguably contributed at numerous occasions to the empirical case against it (see e.g. Khang & Zaidi, 2002; Robilotto & Zaidi, 2006; Zaidi & Bostic, 2008).

For instance, he has maintained the following:

> "It makes little sense for the human visual system to extract high-dimensional spectra because subsequent color computations would be high-dimensional and costly." (Zaidi & Bostic, 2008, p. 2674)

Here, the computational cost of reverse optics strategies is used to motivate the idea that they are not used by the visual system. That is to say, a cost-benefit analysis of the use of reverse optics strategies by the visual system yields the conclusion that the costs far outweigh the benefits and that the visual system would therefore have chosen a different strategy. In other papers, Zaidi and colleagues argued that such strategies cannot account for parts of their data (Robilotto & Zaidi, 2004a, 2006; Zaidi & Bostic, 2008)[57]. Their case against reverse optics therefore appears to include both theoretical and empirical arguments.

However, it is worth pointing out that Zaidi's understanding of the notion of reverse optics is a peculiar one to say the least. Explaining the crux of this approach, he says that:

---

[57] For more on this point, see the next chapter.

> “[many] models of lightness and color constancy assume that the visual system estimates the scene illuminant and uses this estimate to determine material reflectance […] In terms of this illumination-estimation hypothesis, the questions arising about this study are, what information is available in the display, and what is the best any visual system could estimate? Backgrounds in the two boxes have similar statistics. Theoretically, a system could calculate the mean luminance of both backgrounds, then measure the mean luminance of each of the four cups, and take the ratios.” (Robilotto & Zaidi, 2004a, p. 792)

In a later paper, he holds that:

> “If an observer could take precise mean luminance ratios between the backgrounds, this would provide an estimate of the relative intensities of the two illuminants. This estimate could be used to discount the mean luminance differences between the objects.” (Robilotto & Zaidi, 2006, p. 27)

Note that those quotes appear to imply that a reverse optics strategy only employs luminance *ratios*. However, calculating luminance ratios does not entail estimating either the absolute intensity or the SPD of the illuminant. Neither does it involve estimating the SSR of surfaces. But estimating the SPD of illuminants and the SSR of surfaces is an essential step in reverse optics[58], whose main goal is to use the color signal to estimate the causal factors which generated it.

Zaidi therefore appears to sometimes endorse a very peculiar understanding of reverse optics. This is to be contrasted with other passages in which he refers to it as the usually understood recovery of estimates:

> “a number of deterministic and probabilistic inverse-optics models have been proposed to first extract the spectrum of the illuminant and then the spectra of materials under the illuminant for a variety of conditions.” (Zaidi & Bostic, 2008, p. 2674)

And, at some point, both interpretations appear to be mixed:

> “In a reverse optics model of lightness perception, observers first estimate illuminant intensity and then extract relative lightness by discounting the illuminant.” (Robilotto & Zaidi, 2006, p. 33)

In the first part of this quote, Zaidi and colleagues make reference to the estimation of the illuminant[59], a central step in reverse optics strategies as we shall see, whereas in the second part they talk only about relative lightness. This is more reminiscent of Foster’s theory or of

---

[58] In the next chapter, we shall dig deeper into what is essential to reverse optics approaches as applied to color constancy.

[59] In the case of lightness constancy, this will only refer to an estimate of the illuminant intensity whereas in the case of color constancy, an estimate of the illuminant’s SPD is required.

the lightness theories criticized by Gilchrist for their inability to account for the anchoring problem, theories which admittedly only care about relative properties of the scene. However, note that, if the visual system already has access to the illuminant intensity, there is no point in continuing to talk about relative lightness. Once the illuminant intensity is known, absolute lightness of surfaces can be computed with this knowledge and knowledge of the color signal. Zaidi therefore adopts a very particular view of reverse optics.

Keeping this in mind, we can move on to another aspect of Zaidi's approach to constancy. Many researchers will define color constancy using notions such as "perceived object colour" (Buchsbaum, 1980), hold that we possess "a perceptual ability that permits us […] to assign stable colors to objects" (Maloney & Wandell, 1986), talk about "the traditional interpretation of color constancy [as] the invariant appearance of surface colors under changes in the spectral composition of the light source" or maintain that "the visual system processes the retinal image to stabilize object colour against changes in the spectrum of the illumination" (Brainard et al., 2018).

Those ways of framing the issue all seem to contain a phenomenological understanding of color constancy[60]. Color constancy, so conceived, refers to the fact that the colors of objects we experience in our usual, ordinary perception of them are in some sense stable, i.e. they do not depend on illumination conditions or, more generally, on factors extrinsic to the object's surface properties (such as the presence of haze, of transparent objects between the perceiver and the object, or the depth or position of the object in the scene). If I experience a leaf as a very light green under a particular illuminant (or, again, in a given scene), my perception of it is (perfectly) color constant if and only if I see that leaf as the same shade of very light green under any other illuminant (or in any scene distinct in any aspect which causes a significant change to the color signal). If my perception is color constant in this sense, then, as a consequence, which color I see an object as having (in normal circumstances) is always entirely determined by the surface properties of this very object and not by anything else[61] [62] [63].

---

[60] Of course, this is not to say that there are no non-phenomenological elements involved in this understanding. Admittedly, talking of "changes in the spectral composition of the light source" has nothing to do with phenomenology. But the criterion for the possession of color constancy in the face of such changes is the existence of a constant phenomenological element in our ordinary experiences of color.

[61] Or by properties which strictly covary with those surface properties. Note that talking about "surface properties" will include properties such as texture which we may not want to include here.

[62] Some will disagree here. From the fact that my perception of a color is constant in that sense, it is not possible to conclude that which color that color is is entirely determined by the surface properties of the object in question. The argument for that latter claim is supposed to be that the only thing which remains constant across all those different scenes are the surface properties. But this may not be case. The visual system of the observer may also be constant across those scenes and there is therefore no reason to deny that it may play a causal role in

In contrast, Zaidi eschews such phenomenological approaches to color constancy and employs a performance-oriented view of the phenomenon:

> "For many visual tasks the functions of color constancy can be accomplished by establishing correspondence among surfaces seen under different lights across either space or time. The observer can safely infer that corresponding surfaces have similar essential colors (or reflectances)."
> (Zaidi, 1998, p. 1770)

Here Zaidi severely diverges from the inverse optics understanding of color constancy. According to the latter, the chronological ordering of processes when it comes to color constancy is the following. First, the visual system retrieves information pertaining to the objects' surface properties from the color signal. To do so, it applies various constraints which are assumed to apply to natural scenes. In a second time, it uses the information obtained to establish correspondence between objects. Ultimately, the fact that the same object under two illuminants with distinct SPDs is seen as having the same color or as being the same object is a consequence of the fact that the visual system has retrieved information about its surface properties. The surfaces properties, once known, are used to establish correspondence or identity.

This is not so when it comes to Zaidi's view, as far as the above quote is concerned. Here, reaching a conclusion regarding the surface properties of the object appears to be a consequence of the visual system reaching a conclusion regarding correspondence of objects under distinct illuminants[64]. More specifically, the conclusion regarding the surface properties is the conclusion of an inference that actually happens beyond the bounds of visual processing. Knowledge of the surface's properties is the outcome, not the input of the process and this knowledge corresponds to judgments of identity of surface properties across the scene. In other words, it is determined *that* two objects have the same surface properties, not *which* properties they have[65]. The point is also reinforced by the following:

> "The present paper, however, suggests that the problem for the visual system to solve is not to bring about stable color appearance under different illuminants by discounting [illumination changes] but to recognize that

---

at least partly determining which color I see the object as having. That it plays such a role seems to explain why different observers can attribute distinct colors to the same object viewed under the same circumstances.

[63] Note that this does not mean that there may not also be another aspect of our experience which changes along with, e.g., the illuminant. More on this below in this chapter and in the fifth chapter.

[64] And note that this knowledge is not about which wavelength is more or less reflected by the object.

[65] In a trivial sense, which properties they have is also determined. The content of the conclusion could be something like "this object has the surface properties *which* this other object also has". But there is no substantial, independent specification of those properties.

> objects are indeed being viewed under different illuminants and to discover what the illuminant properties are." (Zaidi, 1998, p. 1767)

We can again see the emphasis on the visual system's ability to recognize objects across illuminants with distinct SPDs. Zaidi even defines lightness constancy "in a direct performance-based manner as the ability to identify two objects as having the same lightness across physically different illumination conditions" (Robilotto & Zaidi, 2004a, p. 779).

Note that understanding color (and lightness) constancy in this performance-oriented manner is not without important consequences. First and foremost, adopting this view of color constancy implies that knowledge of reflectance is relegated to a mere "tool in the toolbox". It implies the recognition that estimating SSR is but one of the ways to effectively identify objects across illuminant changes. If it proves too expensive to compute estimates of SSR, then the visual system would have chosen a simpler approach. This is merely another way of saying that knowledge of SSR is not valuable for its own sake (with which researchers adopting reverse optics agree) *and* that there are other, better ways to identify objects across illuminant changes (with which they do not).

Secondly, the question of whether we need to explain an apparent stability of color in our perceptual experience is bypassed and replaced by the question of whether the proposed theory can account for our *ability* to act in a color-constant manner, that is, whether it can account for our recognitional abilities. Those abilities include the capacity studied by Zaidi's papers, namely, the capacity to recognize identical (or sufficiently similar) objects across simultaneous illuminant changes and therefore also to perform asymmetric matching. I take it that Zaidi would also argue that it includes cases of recognition across temporal changes of illumination. For instance, one would be able to re-identify the apple seen on the ground at noon even after a few hours, when the sun is nearly gone and the only light hitting the apple is skylight.

Importantly, understanding color constancy in this fashion has the important consequence that the phenomenological situation need not be one in which the colors attributed to objects in our experiences of them do not vary with factors extrinsic to the objects. That is to say, it is possible to be color-constant in Zaidi's performance-oriented way without also being color-constant in a phenomenological sense. Zaidi hints to that possibility when he says that:

> "[if] such algorithms are implemented by the visual system, they can enable the observer to recognize when the same objects are being viewed under

> different illuminants despite there being a discernible shift in the colors of the objects." (Zaidi, 1998, p. 1774)

And he repeatedly stresses that it would be possible "to identify similar materials across illuminants despite possible appearance changes" (Zaidi, 2000, p. 192). Finally, asymmetric matching, contrarily to Zaidi's experimental method of finding-the-odd-one-out, makes "the assumption that similar materials should appear identical across illuminants" (Zaidi, 2000, p. 193)[66]. According to this understanding of constancy, color-constant behavior does not entail color-constant phenomenology. This means that the colors we think we see may well not be constant across, say, illumination changes at least under certain circumstances. Since SSRs are constant across illumination changes, it would then be impossible to identify the colors with them.

Of course, a philosopher could here agree with Zaidi that the colors do vary across illuminant changes while still believing that there is constancy in a phenomenological sense. For instance, she may argue that, while there is a difference in two differently illuminated scenes with regard to color, it is not a difference consisting in which color is attributed to the objects but rather a difference in how those colors are seen as illuminated. Although the color attributed to the object would here be constant, the way this color is seen as illuminated would not. There would be both constancy *and* constancy at a phenomenological level. In that case, there would be some constancy at a phenomenological level without the overall phenomenology being constant[67].

However, note that Zaidi is most of the time ignoring the distinction between what has been called in the scientific literature on color constancy "perceptually unasserted color" and "surface color" (Arend et al., 1991; Arend & Reeves, 1986). This distinction is supposed to apply to different kinds of judgments an observer might make. On the one hand, she might judge the hue, saturation and brightness of a certain point in the scene viewed. This would correspond to a judgment of perceptually unasserted color (or, for short, unasserted color). She might also judge the color of the surface of an object in the scene as a property of the object. This would be a judgment of surface color. For instance, two oranges may have the same surface color yet differ in perceptually unasserted color (also called "unattributed color" for short) if one of the oranges is seen under a different, say, bluer illuminant. The orange under the bluer illuminant will have a more bluish unasserted color than the other orange. If

---

[66] Note that it is unclear what exactly is meant by the idea that it is asymmetric matching itself which makes this assumption. As an experimental procedure, it seems to be quite free of such assumptions. Zaidi perhaps means that it is the researchers employing the procedure which make this assumption.

[67] Note that this is how most philosophers of color understand the phenomenology of color.

we were to put the two oranges under the same illumination, they would then have the same unasserted color[68].

Zaidi usually avoids the distinction[69] and admittedly similar ones and assume that, if two objects are seen under chromatically very different illuminants (such as skylight and sunlight), then the two objects will appear to have different colors. As we can and do recognize identical objects across scenes and illuminants (i.e. we do have performance-based color constancy) at least under certain circumstances, this entails that the same object will change color across viewing conditions and that there is, strictly speaking, no phenomenological color constancy. Zaidi thus reinterprets what it is for two objects to have the same color across two distinct illuminants as the counterfactual claim that, if they were put under the same illuminant, they would have the same color. He argues that this is a view we can notably find in Lichtenberg:

> "In ordinary life we call white, not what looks white, but what would look white if it was set out in pure sunlight, or in a light whose quality did not differ much from sunlight. It is more the potential to be white and become white, in all its gradations, that we call white in some object, rather that the pure white colour itself. I take this sheet of paper, for example, to be white in the deepest twilight, even at night in the weakest starlight, by tallow, wax or lamp light, in the brightest sunshine, in the red of evening, by snow and rain, in the woods or in a decorated room, etc., for I am convinced that in the clearest sunshine, taken from an alpine peak where the blue sky's reflection is missing, it is nothing less than white." (Joost et al., 2002, p. 302)

Zaidi distinguishes between the color which changes with the viewing conditions, the "sensed" color, and the invariant color which is the "concluded" color. The former is thought to be part of our pre-conceptual experience, akin to an immediate grasping of the information present in the light. The latter in contrast is part of the content of our subsequent judgment, judgment based on the immediate sensation which contains the sensed color.

This is a different distinction than the one between unasserted color and surface color if only for the fact that the latter are thought to be judgments by Arend and Reeves (Arend & Reeves, 1986), whereas the sensed color of an object is supposed to be an immediate part of

---

[68] Note that this distinction can be understood as a phenomenological one only if one holds that the phenomenology of color experiences is constituted by judgments. If the latter is true, then our color experiences would be a conjunction of a judgment that an object O has a particular stable color (i.e., surface color) and a judgment that O has a particular transient color (that is, the unattributed color). Our color phenomenology would therefore exhibit both constancy and inconstancy. However, the constancy and inconstancy would be located at the level of different kinds of properties.

[69] Of course, it is entirely possible that Zaidi would agree with the distinction but argue that the surface color is actually attributed to an object by a cognitive rather than a perceptual or sensory process.

the content of our visual experience. Note that finding out the concluded color of an object implies an inferential process. That is, the color we think is invariant is not something we sense but something we conclude. Furthermore, the concluded color is defined in terms of the sensed color. That is, an object having a particular, constant concluded color only means that this object would appear to have the corresponding sensed color under sunlight. Having a particular concluded color means being able to display the corresponding sensed color when placed under sunlight[70].

Zaidi's theory and its inability to measure reflectance

With all of this in place, we are in a better position to understand Zaidi's proposal regarding what the visual system does when it tries to achieve color constancy understood as the ability to recognize that two identical objects are being seen across distinct illuminants. Obviously, the proposed strategy does not involve computing estimates of reflectance. In contrast, Zaidi proposes distinct strategies which operate on cone signals only to conclude whether some objects under a particular illuminant correspond to some other objects under another illuminant.

Most noticeably, Zaidi proposes to use the fact that cone excitations are affected in a systematic way by illuminant changes. If one plots the cone-opponent excitations of a set of objects under a fairly natural illuminant (say, direct sunlight) against the cone-opponent excitations of the same set of objects under another fairly natural illuminant (say, skylight), the points representing the individual objects will mostly lay on a straight line. In other words, there is a very strong correlation between the respective cone-opponent excitations of an object across illuminant changes. Furthermore, this correlation will actually be very similar to the correlation of cone-opponent excitations corresponding to the two illuminants. The points representing the objects will fall on that very line which runs from the origin (the point indicating no absorption from either cone type) to the point representing the absorption from the illuminants directly hitting our retina (that is, directly projected in the eye). In other words, if for a particular cone type, the excitation corresponding to the first illuminant is x and the excitation corresponding to the second illuminant is y, then the ratio x/y will also describe the correlation of the objects seen under those two illuminants.

More precisely, there is a very simple relationship between the cone-opponent absorptions of one object under one illuminant and the absorptions from the same object but

---

[70] For a more thorough examination of the distinction between sensed and concluded color, see chapter 5.

under another illuminant. Let us take an object O and its signals on the L-(L+M) and S-(L+M) channels under a given illuminant as O_1 and O_2, respectively, and its signals on the same channels under another illuminant as O_3 and O_4, respectively. It has been found that the corresponding shift along the L-(L+M) axis is an additive shift, whereas the corresponding shift on the S-(L+M) axis is a multiplicative one. We can thus write the following:

$$\begin{bmatrix} O1 \\ O2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & \sigma \end{bmatrix} \begin{bmatrix} O3 \\ O4 \end{bmatrix} + \begin{bmatrix} \tau \\ 0 \end{bmatrix}$$

The first step in the visual system processing depends on whether it has access to the same sets of chromaticities under distinct illuminants or whether one of the sets represents a subset of the other one. In the first case, Zaidi's proposal is that the visual system will use the systematic ordering we previously discussed to solve this equation. Remember that cone excitations across two distinct illuminants actually fall on a straight line if you map them. That is, performing rank ordering on the cone excitations will be enough to then match them and solve for the equation, thus recovering information about the illuminants' relative lightness (or relative color).

In the case in which one of the sets of chromaticities is a subset of the other, the visual system cannot just compare them straight away as it does not know which chromaticity corresponds to which other chromaticity in the other set. That is to say, it does not know which object is responsible for which pair of chromaticities across the two sets. Zaidi proposes that it will select the smallest chromaticity (i.e. the one causing the fewest absorptions on the cone-opponent channels) in the subset and match it to one chromaticity in the superset. It will then estimate σ and τ in the equation above (that is, the shifts in illuminant chromaticities). With those estimates, it is able to estimate the transformed chromaticities under the second illuminant. The transformed chromaticities are assumed to belong to the same object as the nearest chromaticity in the superset and the distances between the two respective points are added to give an estimate of the error of the model. The visual system then repeats the whole process by matching the smallest chromaticity to the next chromaticity in the superset. The match which ultimately minimizes the error is chosen as the correct one and correspondences between the respective chromaticities of the subset and superset are then established.

This is in a nutshell how the visual system could operate to achieve performance-oriented color constancy according to Zaidi. A noteworthy point of this perceptual strategy is that it never deals with estimates of reflectances or anything else than plain chromaticities (expressed here as cone-opponent excitations). This entails that the visual system never knows what reflectance properties can be attributed to any single surface but can only establish whether a surface qualitatively identical to another one is being seen under another illuminant. In other words, it never knows what the surface properties of the object are but only whether they reappear under another illuminant. Zaidi's algorithm is therefore as far as possible in spirit and practice from a reverse optics approach.

Note that Zaidi has actually proposed different strategies in subsequent articles. In the case of lightness constancy, it was hypothesized that participants could have used brightness matching after adaptation (Robilotto & Zaidi, 2004a). In the case of color constancy, relative color percepts were argued to underlie the participant's level of constancy (Zaidi & Bostic, 2008)[71]. However, the point of those strategies was to determine whether they could account for the participants' pattern of answers and it was not made clear how exactly those strategies were to be implemented by the visual system. In addition, it was not made clear whether those strategies were subpersonal strategies of the visual system *per se* or whether they corresponded rather to personal-level strategies, that is, consciously used by the participants to perform the task at hand. The latter matters since we are interested in what our visual system does when it seemingly attributes colors to objects, not in what we do with that information once it is available. Consequently, we will focus on the perceptual strategy described above.

But first one astute reader may now ask about our ability to determine the concluded color of an object, as understood above. Zaidi proposes here to use the concept of essential color. The essential color of an object corresponds to the sensor responses from this object under an equal-energy illuminant. However, it is not clear that Zaidi has a definitive answer regarding the process used by the visual system to determine the essential colors of objects apart, that is, for the claim that observers can remember the essential color of an object if they have seen it under different illuminants (Zaidi, 1998, p. 1774). But it is not clear why seeing the same object under different illuminants would help in determining its essential color. Seeing it under an equal-energy illuminant whilst at the same time knowing that the illuminant is an equal-energy one would certainly help. If that were the case, the visual system would just encode that this pattern of luminance is equivalent to the essential color of the object and it would be available for future processing. But this is not what Zaidi proposes and it is not even clear how the visual system would know that the current illuminant is an equal-energy one. As no answer appears forthcoming, I suggest we leave this matter aside for the

---

[71] See the next chapter for a more thorough discussion of those studies.

time being[72] and consider whether Zaidi's theory, if it were successful, could be seen as evidence for reflectance physicalism.

It seems that, at first glance, the answer to this question must be negative. The reason for this appears straightforward. Zaidi's algorithm's output solely consists in judgments of correspondence across distinct illuminants. That is, the output of the algorithm, if it is functioning adequately, is that some object under the first illuminant is identical (or very similar) to another object under another illuminant regarding their surface properties. The algorithm is silent regarding the nature of those surface properties. RP states that the colors which figure in our usual perceptions of the world are nothing but the reflectances of objects.

Zaidi's theory therefore cannot represent evidence for RP for the following reason. Whereas RP requires theories of constancy to contain algorithms whose outputs consist in descriptors of SSRs, Zaidi's algorithm only tells us about correspondence between objects or reflectances. That is to say, the latter is not in any position to tell us about the non-relative color of objects. To see this, consider the following scenario. You are in a forest at dawn, looking at trees and other shrubs. As the sun has still not come up, the only light hitting the vegetation is skylight. As a consequence, there is few or any shadow or illumination change across the scene you are viewing. Let us say that you are looking at the leaf of a particular tree. Most probably, you will see the leaf as having a particular color, presumably, some shade of green. How you came to see it as this color is something that cannot be explained by Zaidi's theory. This is not a case of determining which other leaves on other trees have the same surface color as the leaf you are looking at right now. In this situation, your visual system is able to attribute some color property (some shade of green) to a certain object in the scene you are viewing (a leaf). However, as far as this task is concerned, Zaidi's algorithm is of no use to us here.

To present the point more systematically, the algorithm does not seem to be in a position to explain our ability to perform color naming. Color naming refers to the experimental paradigm in which subjects identify the color of a patch on a neutral background under varying illumination (Troost & De Weert, 1991; Weiss et al., 2017; Witzel et al., 2016). For example, a red patch might be displayed on a neutral background under a spectrally neutral illumination and subjects have to identify its color. The same red patch will then later be represented, after many other trials, under a different illumination, say, a blue one, and subjects try identifying the color of the patch again (without knowing that it is the same patch,

---

[72] But see chapter 5 for a more complete discussion of the distinction between sensed and concluded color.

of course). As the number of different illuminations the patch is presented under increases, it can be statistically determined whether the same red patch is still judged as red under other strongly chromatic illuminations. The same process will be repeated for all the other hues and a measure of constancy statistically developed.

Even though the point of color naming is to ultimately determine the extent to which we are color constant, note that our ability to perform the task seemingly requires us to attribute non-relative colors to color patches, and not to match objects across illuminants, either simultaneously or sequentially. Those tasks are distinct. A theory of color constancy which has the potential to represent evidence for reflectance physicalism arguably will be able to explain how we achieve color constant behavior in both senses. If a color descriptor is attributed by a visual system to a surface independently of the properties of the scene unrelated with the surface's properties, then that attribution will make possible both color naming (since the descriptor will be correlated with the color name) and asymmetric matching (since the descriptor will not change with the illumination), either simultaneously or sequentially. However, it does not look as if Zaidi's theory fits the bill.

It appears we have found a theory which is unable to represent evidence for reflectance physicalism. If Zaidi's theory were confirmed and the algorithm it contains were an accurate portrayal of the perceptual process occurring when our perceptual system behaves in a color-constant fashion, this would not constitute justification for the thesis that colors are actually SSRs. It would actually put pressure on the idea that our visual system is at all interested in the SSRs of objects.

At this point, one might retort that the theory would fit the bill if we included the part concerning essential colors. That is, including the notion of essential colors into the theory would allow for a measurement, as we have understood it, of surface spectral reflectances. However, we have seen that it is unclear how the visual system would find its way back to essential colors. As long as no answer to that issue is provided, it is difficult to argue that the theory so modified would involve a measurement of SSR. Nevertheless, it is true that adding essential colors to Zaidi's simpler theory may represent a viable way to do so.

Another objection to our construal is that the theory only needs an anchor in the same way as Foster's in order to count as involving a measurement of SSR. Once an anchor is provided, the theory will be able to assign colors to distinct SSRs. But note that this objection misunderstands the nature of the output of Zaidi's theory. This output only consists in judgments of correspondence between two surfaces under respectively distinct illuminants.

Even with an anchor, this output will not suddenly signal anything about the SSR of the surface.

Summary

In this chapter, we have addressed the lack of consensus in color constancy research regarding *Recovery* which is argued by Chirimuuta as being an important problem in Hilbert's strategy. I have tried to show that it is not so and that whether a theory of color constancy supports reflectance physicalism does not depend on whether that theory includes reflectance recovery as understood by Maloney's and subsequent work. A color constancy theory needs not entail *Recovery* in order to potentially support RP. To support RP, a theory of color constancy should include some measurement of reflectance, in a sense wider than but including reflectance recovery. This measurement will consist in an attribution of colors to SSRs which fulfills the three criteria of veridical sensory systems we have identified.

On the one hand, Foster's theory can be seen as a theory which does not involve reflectance recovery but still includes a measurement of SSR once an anchoring rule is provided. On the other hand, there is no addendum to Zaidi's theory that allows a measurement of SSR. In that way, we have identified a criterion for inclusion into the set of theories potentially able to support reflectance physicalism.

In the next chapter, we delve into the second part of Chirimuuta's argument, that is, that Hilbert's strategy fails because the opinion in favor of reflectance recovery being a part of visual processing is more a reflection of previous theoretical commitments at play than a dispassionate reading of the evidence.

Chapter 4: The albedo hypothesis and reverse optics frameworks

Introduction

Let us take stock. In the previous chapter, we saw how a color constancy theory distinct from Maloney's could represent evidence for reflectance physicalism if it were to be successful. In contrast, Zaidi's theory is probably unable to provide such support. The crux of the matter is that whether the theory in question was evidence for reflectance physicalism did not hang on whether that theory involved reflectance recovery. Nevertheless, it ought to include some kind of process which can be rightly seen as a measurement of surface spectral reflectance.

However, we have so far not addressed Chirimuuta's second worry. Her main point is not that Hilbert's case is undermined merely because there is no consensus regarding reflectance recovery in color constancy research. It is rather that there is no such consensus because the various teams of researchers weighing in on the issue embrace distinct conceptual assumptions about the nature of vision. The evidence accrued in color constancy research is not, as Chirimuuta puts it, conceptually neutral.

This is an issue at least for the following reason. If one needs additional assumptions regarding the nature of vision (or perception, in general) to read color constancy data in a way which supports reflectance physicalism, then one needs reasons to believe in those assumptions. As there are different conceptual assumptions available, those reasons should distinguish between those which speak in favor of the reading supporting reflectance physicalism and the other readings. In addition, Chirimuuta points out that if the respective conception of the nature of vision adopted by the researchers shapes their interpretation of the data in such a profound way, then there is little hope that acquiring more data will solve the issue. Once these new data come to light, there is no reason to believe that they will not be interpreted in a way which favors the pre-existing conceptual assumptions of the researchers.

However, there is a more dire and more specific issue according to Chirimuuta. The conceptual assumptions which one needs to interpret the data in a way which supports a reverse optics approach to vision are very much linked with a position akin to reflectance physicalism. If this is true, then it seems that in order to believe that color constancy data speak in favor of RP one must already believe in RP. The appeal to the data thus appears to be circular. Even if we are successful in showing that the pool of theories Hilbert can appeal to in order to seemingly justify RP includes but is not limited to the set of theories involving reflectance recovery, this does not help as long as we have not addressed this second worry. If we have not addressed it, any theory included in the larger pool can still be accused of presupposing RP and, if it is true, then appealing to it to justify this color ontology is actually a circular move.

In this chapter, I tackle this second worry. To reiterate, Chirimuuta's central point is that the results of color constancy data do not speak unequivocally for or against the thesis that the visual system performs reflectance recovery (henceforth, *Recovery*) but requires the previous adoption of RP to speak for this thesis[73]. I here try to break the cycle, so to speak. I will argue that there already are data which are able to speak for or against *Recovery* without us endorsing or rejecting either reverse optics or RP. *Recovery* has empirical consequences which can be assessed and whose confirmation or rejection would be relevant to its truth or falsity, respectively. A central part of this chapter will be concerned with the development and explanation of one of those consequences. In order to first set the stage, we have to consider the previously discussed notion of reverse optics. Strategies subscribing to reverse optics involve reflectance recovery as the final step in achieving color constancy. To be in a position to perform this final step, however, the visual system has to first compute an estimate of the illuminant as a necessary first step. In other words, there can be no recovery of reflectance (i.e. *Recovery* cannot be true) unless the visual system estimates the illuminant. This illuminant estimation is the empirical consequence we shall discuss.

In the final part of the chapter, I also present data, both old and relatively new, which speak against Chirimuuta's claim that it is not possible to compare two color constancy models with distinct conceptual ancestry against one another. Those experiments apparently show that this is not the case and that such models are not as experimentally isolated as has been claimed. Before we dive in, however, we have to carefully expose Chirimuuta's worry. We will thus have to determine exactly how she conceives of the relationship between the theories of color constancy and the conceptual assumptions endorsed by the researchers defending those theories.

---

[73] In addition, note that, if the reasoning I developed in chapter 3 is sound, then it may very well be the case that *Recovery* is not necessary to defend reflectance realism. However, this does not mean that it is not a viable way to defend the latter.

Chirimuuta's conception

The first main thesis of this second part of Chirimuuta's project is that it is impossible to interpret color constancy data as speaking in favor of *Recovery* if one does not subscribe to reverse optics first. This is, as we have seen, already an issue in its own right. Color constancy evidence should be neutral in a way which allows to adjudicate between rival approaches to color vision. If it does not unequivocally rule in or out the thesis *Recovery*[74], then it is not in a position to do so. Secondly, Chirimuuta argues, things are even worse for if one subscribes to reverse optics, it becomes impossible not to also subscribe to a form of reflectance physicalism. Were this the case, it would be circular to argue for RP on the basis of color constancy data since one would have to first adhere to RP in order to judge this data as speaking for *Recovery* and therefore for RP[75].

As far as the first thesis is concerned, Chirimuuta argues that it is important to consider the following facts. Both teams of researchers we discussed interpret what appears to be very similar data in incompatible ways. Maloney's team sees its data as speaking in favor of *Recovery*, whereas Zaidi's holds that its data are in conflict with it. But the only thing which differs between the two teams is their respective stance on the nature of vision. On the one hand, Maloney conceives of vision as the recovery of physical properties by the visual system which is trying to solve an inverse problem. On the other, Zaidi endorses a heuristic approach to vision according to which it is nothing but the most pragmatic processes and "tricks" the visual system has found to deal efficiently with its environment[76]. But then, Chirimuuta concludes, the only factor which can explain their divergence in the interpretation of the data is their endorsement of those distinct conceptions of vision. There is nothing else which relevantly varies between the teams.

Consequently, or so Chirimuuta argues, Maloney and his colleagues endorse a particular interpretation of their data, one which speaks in favor of *Recovery*, because they adhere to a reverse optics conception of vision. Similarly, Zaidi and his colleagues reject that interpretation of the data for one which speaks against *Recovery* because they reject that conception and subscribe to a heuristic approach to vision. Chirimuuta motivates this thesis by pointing out the consequences of each conception of vision. As we saw, Zaidi and

---

[74] Or theses which are part and parcel of reverse optics approaches.

[75] As we saw in chapter 3, if it is not the case that the empirical argument for RP requires the truth of *Recovery*, then it cannot merely be said that it is an issue that one needs to endorse RP in order to interpret data as speaking for *Recovery*. However, if the latter is true, this may become an issue if the theory of constancy which is supposed to motivate RP entails that *Recovery* is false. In that case, the argument would entail that *Recovery* is both true and false.

[76] And with the rest of the neural machinery, of course. The visual system has to deal both with inputs from the environment and with the demands of the cognitive system.

colleagues believe "vision to be the processing of sensory information in such a way as to find the set of relations between 'sensory qualities' […] that is most useful to the organism" (Chirimuuta, 2008, p. 576).

On that view, there is no point in looking for a place in the brain where the solution to an inverse problem is given or for an algorithm capable of yielding that solution. Even if there were such a place or algorithm, it would not be a proper part of vision[77]. However, given the cost of the necessary computations and the fact that alternative strategies exist, the heuristic conception of vision denies that we will find proof of reverse optics strategies being implemented by the visual system. Because it is possible to interpret their data in a more parsimonious way, Zaidi and colleagues opt for the heuristic interpretation, which is the more parsimonious one. Hence, they do interpret their data in a way consistent with this interpretation. Maloney's team is looking for a transformation capable of achieving the inversion of the color signal to finally reach a conclusion regarding the physical properties causing this color signal. They see their results as proof that this is exactly what the visual system is doing[78].

Chirimuuta also claims that endorsing reverse optics entails endorsing a theory very similar to RP. In other words, one can subscribe to reverse optics only if one subscribes to an ontology of color which identifies the colors which prominently figure in our perceptual experiences with certain physical properties of the objects we encounter in our environment. To see this, remember that, according to the reverse optics conception, color vision is the visual system's attempt at a solution to a series of inverse problems generated by the ambiguity inherent in the color signal. By solving those problems, the visual system makes judgments about the physical properties causally responsible for this signal, in our case, about the SSR of objects. It then seems quite natural to say that judgments of color are actually judgments of SSRs in disguise. But this is very reminiscent of Hilbert claiming that colors are estimates of reflectance:

> "Our view that reflectance-types are represented in color experience is translated by Funt's terminology as the claim that the color vision system estimates certain reflectance-types." (Byrne & Hilbert, 2003, p. 56)

---

[77] Note that Chirimuuta sees this point as meaning that the two researchers are dealing with distinct issues. As we saw in the last chapter, it is important that Zaidi's opinion not be that there may be reverse optics processes further down the line. Maloney could very well agree with the latter and then the two researchers would not be interpreting the same data in incompatible ways. Zaidi must be of the opinion that there is no reflectance recovery happening at any point in the process, be it at a visual or post-visual stage. Or Maloney needs to argue that it happens at the visual stage if anywhere meanwhile Zaidi disagrees on that very point.

[78] Note that it is unclear whether those two teams ought to interpret their respective data in such a strong way. More on this later.

When our visual system attributes color to objects, according to Hilbert's theory, it is trying to determine what kind of reflectance (or reflectance-type) the object has. This is exactly the output of the algorithmic process instantiated in color constant visual systems according to a reverse optics approach to vision. It therefore appears that a reverse optics approach contains a commitment to a form of reflectance physicalism. This is an undesirable result for remember that Hilbert's appeal to color constancy data was supposed to justify RP. But if such an appeal works only insofar as one already endorses RP, then it fails in the most circular way.

I discuss this second main thesis in the following chapter. For the time being, I will focus on dealing with the first one. More specifically, I will concentrate on one part of it, namely, that the data we do have speak for *Recovery* only if we already endorse a reverse optics conception of the nature of vision. There are a variety of ways in which one could start putting pressure on this thesis. One could for instance try to show that some color constancy theories whose success would speak in favor of *Recovery* are actually proposed by researchers who explicitly or implicitly reject any realist color ontology. This is a direct way of confronting Chirimuuta's point that interpreting the evidence in a way which speaks for *Recovery* requires endorsing a reverse optics approach to vision[79].

However, it is not always clear which ontology of color researchers in computational color constancy endorse, if any. It sometimes seems that they either do not really care about ontology or that they believe that it is not the time and place to address such issues[80]. A second, more indirect strategy would try and show that the empirical results of color constancy research are not open to the influence of preexisting conceptions in the way Chirimuuta's argument requires. That is to say, the evidence accrued in color constancy research is not ambiguous to the point that it can be interpreted in conflicting ways. More specifically, Chirimuuta appears to believe that whether an empirical result speaks in favor of *Recovery* is open to interpretation in such a way that there needs to be a previous endorsement of some conception of vision already in place to determine whether the result speaks for or against that hypothesis.

In this chapter, I will argue against that point of view. Although color constancy data certainly are open to interpretation (as is any scientific result), it is not so open as to validate

---

[79] Of course, one could also ask in what sense the two teams of researchers are interpreting "the same data" in different ways. After all, Zaidi and his colleagues are interpreting data coming from their experiments, whilst Maloney and colleagues are doing the same with their own data. So it needs to be explained in which sense those sets of data are "the same".

[80] In addition to the researchers who just straight up claim that colors are subjective properties in some sense (Land, 1983; Palmer, 1999; Zeki, 1983).

Chirimuuta's point. If it were the case, then there would be no clear implications of *Recovery*. However, the latter has empirical consequences which can be assessed and most researchers, including those who argue against *Recovery*, agree that the latter stands or falls with those consequences. Admittedly, this thesis is inconsistent with Chirimuuta's opinion, since if it were true, there would be no need to endorse reverse optics to see whether *Recovery* is true or not.

In the following, I will try and show that the literature on color constancy has already identified one such consequence and put it to the test repeatedly. Consequently, the situation is more nuanced than what Chirimuuta assumes. It is possible to directly falsify even a research program as she understands it or a particular conception of vision. As we shall see in the next chapter, this has implications for how we can and should study color ontology. In the next two sections, I discuss the notion of reverse optics. I will present how it is conceived to extract two conditions that need to be satisfied if a model is to count as involving reverse optics. I then test this proposal by seeing whether it correctly categorizes two prominent models in color constancy research. We will then see that one of the conditions is actually a testable hypothesis which has repeatedly been studied. This is argued to be a counterexample to Chirimuuta's opinion that it is not possible to provide evidence for or against a particular conception of the nature of vision.

In addition, note that Chirimuuta's picture entails that there can be no experiment able to unequivocally compare the two conceptions of vision at issue. That is, there cannot be a study comparing the predictions of the reverse optics approach to those of the heuristic approach without there being the need to endorse one or the other in order to interpret the data as favoring one or the other. If there were such a study, it would be possible to directly confront the two models of color constancy, something which should not be possible, or so Chirimuuta argues:

> "Note also that Zaidi and Maloney's groups are using different lightness constancy paradigms which, arguably, amounts to an exploration of different problems. […] Crucially, one group's model cannot be tested with the other group's data. With the current state of the evidence, the two models—and therefore, the two stances with respect to lightness recovery—cannot be compared directly against each other." (Chirimuuta, 2008, p. 572)

Although there may not have been data comparing the two models at that point in time (although this is by far not clear, as we shall see), we now have some studies which intend to compare color constancy models adhering to distinct conceptions of vision. In the final part of the chapter, I present those studies and argue that this counters Chirimuuta's picture further.

Color constancy theories with distinct conceptions of vision are not as experimentally isolated as she appears to claim. On the contrary, it is possible to have them go head-to-head and compare their predictions regarding the very same experiment. But, without further ado, let us for the time being dig into the notion of reverse optics and its empirical consequences.

The notion of reverse optics

The conception of reverse optics understands vision as the recovery of distal knowledge on the basis of the impoverished retinal stimulus. Here is a classic textbook description:

> "The early stages of visual perception can be viewed as trying to solve what is often called the inverse[81] problem: how to get from optical images of scenes back to knowledge of the objects that gave rise to them. From this perspective, the most obvious solution is for vision to try to invert the process of image formation by undoing the optical transformations that happen during image formation." (Palmer, 1999, p. 23)

Based on this quote, one can deduce that a reverse optics approach to vision comprises at least two distinct assumptions. First, the end goal of the visual process is knowledge of the environment it is performing in. More specifically, the visual system is portrayed as striving to acquire knowledge of the objects which are, in a certain sense, the cause of the visual processes themselves.

One could argue that the most clear-cut cause of visual processes is the bidimensional proximal array of light hitting the retina or a configuration of light in the observer's environment. However, information regarding this array of light is not very useful on its own since it conflates so many factors. It is usually thought that what the visual system is trying to do is acquire knowledge about the relevant objects in its environment. Those objects are relevant insofar as the observer can interact with them or if knowledge of their properties allows the observer to navigate her or his environment[82]. In the case of color constancy, for instance, the visual system will try and compute estimates of the SSRs of the surfaces present in the scene because it is these properties that are deemed relevant for the observer[83]. As Helmholtz famously argued, "colors have their greatest significance for us in so far as they

---

[81] "Inverse optics" and "reverse optics" are used interchangeably.

[82] This is why "objects" can be understood quite broadly here and cover the illuminant itself. However, remember that knowledge of the properties of the illuminant is not the same as knowledge of the color signal. The illuminant is only partly causally responsible for the color signal.

[83] Note that it is usually left out of the discussion exactly how those properties would be relevant to the observer. Helmholtz' quote does not help us much since it does not tell us in which way and to what extent colors aid in identification.

are properties of bodies and can be used as marks of identification of bodies" (Helmholtz, 1896, quoted in (Brainard & Maloney, 2011, p. 4)).

Note that this is already an assumption about the function of the visual system. One might argue that the final function of vision is, alongside other forms of perception, to contribute to behavior regulation. To argue that acquiring knowledge of the distal world causing the visual processes (or a subset of it) is the function of vision is already a proposal as to how it effects this final function of guiding behavior. Of course, it could be objected that any proposal as to how the visual system fulfils this final function will ultimately include some kind of knowledge acquisition related to the environment of the observer (but see Akins, 1996 for some nuance regarding how to understand this knowledge in at least certain sensory systems).

Alongside this assumption comes a second one regarding how the visual system acquires knowledge of the distal world. Vision consists in inverting "the process of image formation". To recover knowledge of, e.g., the 3D properties of the objects at the source of the visual process, the visual system uses any clues at his disposal to reverse the physical processes that compressed a 3D scene into a 2D retinal image. This is called an "inverse problem" and David Marr, a leading proponent and developer of the reverse optics approach, argued that vision was nothing but the solving of a series of inverse problems. Note here again that this is an assumption as to how the visual system recovers knowledge of the distal world. One might envision alternative ways of acquiring this knowledge.

However, if one assumes that the visual system performs an inversion of the retinal image, we can see why the problem is, in a mathematical sense, ill-posed. For a given configuration of the environment and one observer (that is, receptor such as a retina) situated in that environment, only one retinal image will result whereas a given retinal image could have been produced by a variety of configurations of the environment. That is to say, if we define a 3D environment and a visual system situated in this environment (i.e. a person looking at it from a particular vantage point), we can calculate the resulting retinal image exactly. Similarly, in the case of color, the combination of a particular SPD (of the illuminant) and a particular SSR (of a given surface) will yield one unique retinal image[84]. However, as we previously saw, the opposite is not true. Given one retinal image, there theoretically is an infinity of combinations of SPDs and SSRs which could have produced it. As Palmer puts it, "the inverse problem is *underspecified* (or *under-constrained* or *underdetermined*) by the

---

[84] Assuming we are in a flat-matte-diffuse environment. I shall come back to this point later.

sensory data in the image" (Palmer, 1999, p. 23) (original italics). There is no unique solution to the problem, and this is what makes it hard to solve for our visual system[85].

Approaches which conceptualize vision as an inverse problem (or *reverse optics approaches*) therefore have to propose a strategy to reverse the equation so that the visual system can compute an estimate of the properties of surfaces in the scene. They will propose that the visual system proceeds in two stages, which gives those strategies their name *two-stage algorithms*. First, the visual system will use an ensemble of visual cues at its disposal to recover an estimate of the SPD of the illuminant. It will then use this estimate in tandem with the color signal to compute an estimate of the SSRs of the surfaces viewed.

To summarize, we can see that an approach employing reverse optics strategies in color constancy research must fulfil at least the two following conditions: first, it must conceive of the role of the visual system as that of recovering information regarding the surfaces in the field of view. In all likelihood, this information will concern *inter alia* the surface spectral reflectances of those surfaces. Besides, the approach should presuppose that the visual system intends to acquire that information by reversing the process of image formation, notably by computing an estimate of the scene illuminant's SPD. Armed with this analysis of reverse optics, we are now in a position to test how the two assumptions categorize prominent positions in the field.

David Brainard, a leading researcher in color constancy research, has developed a particular approach to color constancy over the years (Brainard et al., 1993, 1997; Brainard & Maloney, 2011; Speigle & Brainard, 1996). According to the "Equivalent Illuminant Model" approach, the visual system proceeds using exactly two stages. First, and this is relevant to the second of our assumptions, it will compute an estimate of the illuminant, called *the equivalent illuminant*. This estimate will play the part of the actual illuminant in the visual process of reversing the operations that led to the formation of the retinal image in order to compute an estimate of SSRs in the field of view. If there are any mistakes made by the visual system in the computations it effects in the pursuit of color constancy, they will reside at the level of the equivalent illuminant. That is, if the visual system computes estimates of SSRs which diverge

---

[85] Note that Palmer actually believes that we know that our visual system performs this feat of strength on the basis of the fact that we can see, e.g., in 3D. That is, he presupposes that, since we are able to see in 3D, the visual system must be performing an inverse optics operation. It appears that he therefore argues from a given performance of the visual system (in that case, 3D vision) to a given strategy the visual system must employ to display such a performance. However, if it is possible to explain our ability to see in 3D without resorting to the notion of reverse optics, then the evidential relationship between the performance of the visual system and the strategy it employs to achieve this performance is also undetermined.

from the actual SSRs, it will be because the equivalent illuminant itself diverges from the actual illuminant (i.e., the actual illuminant is not as the equivalent illuminant portrays it).

Note that most if not all articles presuppose that, once there is a good illuminant estimate available (a good equivalent illuminant, in Brainard's terminology), the visual system should have no issue computing SSRs estimates. That is to say, once a visual system has successfully completed the first stage of the algorithm, the second stage is essentially a free meal. To the best of my knowledge, this assumption is taken for granted by the entirety of researchers[86]. If we restrict ourselves to flat-matte-diffuse scenes (in which there is no depth or illuminant orientation), this appears to make sense at a mathematical level. In flat-matte-diffuse conditions, for instance, using Mondrians as scenes, the color signal C(λ) is the product of the SSR of the surface being viewed (S(λ)) and the SPD of the illuminant hitting this surface (I(λ)):

$$C(\lambda) = I(\lambda)S(\lambda)$$

As the visual system knows the color signal, if it has an estimate of the illuminant's SPD, it is in a position to rearrange the equation to express the SSR as an estimate of the color signal over the illuminant estimate:

$$S(\lambda) = C(\lambda)/\hat{I}(\lambda)^{87}$$

Furthermore, regarding the first assumption we discussed, it is clear to the authors that the function of color constant processes is to acquire knowledge regarding the properties of the distal cause of the retinal image. As they put it:

> "As Helmholtz emphasized over a century ago, an important value of color appearance is to represent *the physical properties* of objects and to aid in identification […] For color appearance to be useful in this manner, the appearance of any given object must remain stable across changes in the scene in which it is viewed. To the extent that visual processing assigns the same surface color percept to a given physical surface, independent of the illuminant and the other surfaces in the scene, we say that the visual system is color constant." (Brainard & Maloney, 2011, p. 4) (my italics)

We can see that knowledge about the physical properties of objects is here deemed useful to the organism in question. More precisely, they are thought to be useful for object identification and re-identification. Therefore, the equivalent illuminant model satisfies the

---

[86] Even if some researchers do not agree that the visual system computes an estimate of the illuminant, they agree that, if it did compute one, then estimating reflectance would be a very straightforward task afterwards. Consider Foster's claim that "given the one estimate, the other is also implicitly available" (Foster, 2011, p. 690).

[87] The caret is meant to show that we are dealing with an estimate of the illuminant rather than the actual one.

first of our assumptions and Brainard's can be deemed a reverse optics approach to color constancy research.

In contrast, we can see and check whether our two assumptions correctly categorize a theory which is usually considered as eschewing reverse optics strategies, namely, the one of David Foster, Sergio Nascimento and colleagues (Craven & Foster, 1992; Foster et al., 1992, 1997; Foster & Nascimento, 1994; Linnell & Foster, 1996; Nascimento et al., 2002, 2004; Nascimento & Foster, 1997, 2000), who, as we saw in the last chapter, have introduced the notion of relational color constancy.

Remember that relational color constancy (RCC) is the constancy of the color relations between two surfaces different in reflectance. So, for instance, a surface standing in a two-to-one ratio of S-cones with another surface will be seen as conserving that ratio across an illuminant change by a visual system displaying RCC, according to Foster's theory of cone-excitation ratios. This is distinguished from a more traditional conception of constancy according to which the first surface will be judged as having the same color across the illumination change, irrespective of whether it still stands in the same color relations with the second one after the change.

First, note that RCC and the calculation of cone-excitation ratios (whether between two surfaces or between one surface and a chromatic average, whether between cone channels or between cone-opponent channels, etc.…) need involve only calculations over untransformed cone signals. Cone-excitation ratios (of any kind) cannot even compute an estimate of the illuminant. We can therefore conclude that theoretical approaches based on the notion of RCC and cone-excitation ratios do not satisfy the second of our assumptions, namely, they do not view illuminant estimation as necessary or even feasible in achieving color constancy.

Regarding the first of our assumptions, the situation is more unclear. Remember that a reverse optics approach is supposed to conceive of the goal of the color constancy process as the recovery of information regarding the surface properties of the objects in the field of view. Traditionally, this information is thought to refer to the SSR of those objects. In that sense, it is information that pertains to a non-relational property of visible objects[88]. However, there is no reason this has to be the only interpretation of the notion of surface properties. One could argue that cone-excitation ratios actually track differences in the surfaces' preferences for

---

[88] It could be countered that SSR is actually a relational property of surfaces as it is a property of a surface with respect to how this surface relates to the ambient light. However, it is non-relational if what counts as the other relata are other surfaces or an observer.

reflecting and absorbing lights of such-and-such wavelengths. In that sense, the visual system would acquire knowledge of relative and relational properties of visible objects[89]. Or one might say that the visual system tracks the differences in the scene's reflectivity of such-and-such wavelengths. In that scenario, the object one acquires knowledge about is the scene in front of one's eyes.

Independently of the way we resolve this issue, it is still the case that the relational constancy approach to color constancy still fails to satisfy the second of our assumptions. Subsequently, it cannot count as a reverse optics approach. This is a welcomed result as Foster has repeatedly eschewed this kind of approach to color constancy (see, e.g., (Foster, 2003, 2011)).

Reverse optics and the albedo hypothesis

We saw in the last section that reverse optics strategies require the visual system to compute an estimate of the illuminant as a prerequisite to reflectance estimation. The hypothesis that the visual system does indeed compute such an estimate has been called the illuminant estimation hypothesis or albedo hypothesis. One might thus conclude that the next step in our argumentation regarding the status of reverse optics strategies would be to check and see whether the data we have is overall in favor or against that hypothesis.

Unfortunately, things are more complicated than we might think. The albedo hypothesis has been conceived of in very different ways throughout its history and it is thus hard to argue that evidence is for or against *it*. I first present the various claims which have been proposed under the umbrella of "albedo hypothesis" and then evaluate the evidence according to which hypothesis it targets. In order to disentangle the complicated matter that is the study of illumination perception, it is useful to consider Alan Gilchrist's categorization of the main positions within the field of lightness constancy (Gilchrist, 2006)[90].

First, there are those, led by Katz or Helmholtz, who think that lightness estimation is derived from the estimation of the illumination[91]. According to Gilchrist, this view entails that, for a given luminance level, a change in perceived lightness should be accompanied by a compensatory change in perceived illumination. For instance, if the luminance of a target

---

[89] For a thorough discussion of this idea, see next chapter.

[90] Most of the data we have on the albedo hypothesis come from research on lightness constancy rather than color constancy.

[91] Remember that lightness is the achromatic counterpart of SSR in the sense that the lightness is equivalent to the average reflectance of a surface across all wavelengths.

stays the same but it is seen as lighter than before, the illumination should be perceived as darker.

Others, such as Koffka and Gelb, believe that the conditions associated with a successful assessment of lightness are the same conditions that lead to a successful estimation of the illumination but that there is no causal link between the two. In other words, an estimate of the illuminant is not causally required to estimate lightness because both are the results of parallel processes that benefit from the same conditions.

A third category of scientists have rejected the dependence of those two factors. Correct lightness estimates can arise in the absence of correct illumination estimation and vice-versa. Finally, it has also been argued that illumination perception actually precludes estimation of lightness and vice-versa. According to these views, one can see only the illuminant or the lightness of surfaces but not both[92].

The albedo hypothesis we are looking for is to be found either in the first or in the second categories of views just presented. One might argue that the albedo hypothesis is obviously the one that postulates a causal link between illumination estimation and reflectance estimation, that is, the first position presented. This is a fair point. Nevertheless, it may miss an important issue. If the first view (let us call it *the causal view*) entails the second (following Gilchrist, *the invariance hypothesis*), then an assessment of the latter is actually necessary in order to fully assess the first. Any evidence against the invariance hypothesis will be evidence against the causal view.

Therefore, we shall have to determine exactly the nature of the relationship between the invariance hypothesis and the causal view. Beforehand, however, we have to further clarify those views.

The causal view

Perhaps the most important point to clarify when it comes to the causal view is to determine whether an estimate of the illuminant is, in addition to luminance, the only element used to compute an estimate of reflectance or whether it is only one element amongst others which are all combined in order to compute this estimate of reflectance. We can call the first thesis *the exclusivity hypothesis* and the second *the partial cause thesis*.

---

[92] One may argue that the fact that we observe such diversity of opinion with regard to the link between lightness estimation and illuminant estimation is proof that reverse optics is already in trouble. But not that this is not the matter we are concerned with here. Rather, the thesis I will defend is the conditional one that, if endorsing reverse optics is the correct way to approach lightness or color constancy, then the albedo hypothesis is true.

If we come back to our understanding of reverse optics, the second of our two requirements might come in handy. It stated that, in order to acquire knowledge about the objects in one's environment, one's visual system has to reverse the optical process responsible for our retinal images. Notably, we saw that computing an estimate of the illuminant was a major element of this reversal. However, this particular requirement, the computing of an estimate of the illuminant, does not allow us to distinguish between our two theses since both of them hold that the visual system does perform such computation. But we can go back to the notion of inverting (reversing) the optical processes that yield the retinal image. If those processes include more than just the product of an illuminant and a reflectance, then these additional steps also need to be reversed, thus granting more weight to the partial cause view than to the exclusivity hypothesis. Here it pays to consider the properties of visual scenes.

A Mondrian scene, as we saw, is two-dimensional, is illuminated by a single illuminant and contains only perfectly Lambertian (i.e., matte) surfaces which do not cause any interreflections (since they are coplanar). In this kind of scene, the color signal (that is, the reflected light) is indeed the product of only the illuminant and the reflectance of the surface viewed and it is thus impossible to distinguish between our two theses, which would both predict the same outcomes. However, more complex scenes introduce additional variables whose effects need to be untangled from the color signal.

As an example, consider the case of achromatic color constancy (lightness constancy) in a three-dimensional setting. If directional lighting is introduced in a 3D scene, the amount of light which will be reflected by a surface will depend not only on the surface's reflectance but also on the angle at which the surface faces the source of light. An algorithm that took into account only the properties of the illuminant and those of the reflectance of a surface, ignoring its angle relative to the light, would systematically misestimate the lightness of the surface. Thus, a surface directly facing the directional light will be mistakenly estimated as lighter than the same surface facing away with an angle of 45° from the light[93]. An algorithm that did not take shading into account would also misestimate lightness in 3D settings. Imagine a pyramid casting a shadow onto itself because of directional lighting. If an algorithm subtracts only the influence of the illuminant (the combination of the diffuse and directional one), it will mistakenly compute a lightness that is darker than it really is for the surface of the pyramid in the shade.

---

[93] Provided the observer is occupying the same place as the light, i.e., facing the patch with the same 45° angle.

In those complex situations, an algorithm trying to compute an estimate of reflectance must reverse more than just the product of illuminant and reflectance. This is a consequence of the fact that, in complex, realistic scenes, the reflected signal is the result of a variety of variables. In addition to the angle of the surface in relation to directional lighting and shading, such variables include, amongst others, the presence of haze, transparent surfaces, and interreflections[94].

Consequently, our two theses, the exclusivity hypothesis and the partial cause thesis, can only be distinguished for scenes which approximate the complexity we find in our natural visual environment. In simpler environments, such as Mondrian scenes, the color signal is indeed a product of the illuminant and the reflectance of the surfaces in the scene and the partial cause thesis essentially collapses into the exclusivity hypothesis.

The invariance hypothesis

An apparently less stringent view of the relation between illuminant estimation and reflectance estimation can be found in what has been called the *lightness-illumination invariance hypothesis* (Gilchrist, 2006; Kozaki & Noguchi, 1976; Noguchi & Kozaki, 1985) or the *invariance hypothesis*, for short. This hypothesis holds that the sum of the lightness and the illumination estimate is invariant for a given luminance level[95]. In other words, given a certain luminance for a test patch, a change in the lightness judgment of this test patch entails a commensurate change in the illumination judgment and vice-versa.

For instance, if a test patch with a given luminance is judged as twice lighter than another test patch with the same luminance, then the illumination falling on the first should be deemed twice darker as the illumination falling on the second. This essentially means that lightness and illumination estimation are closely correlated with the actual reflectance and illumination of the test patch, respectively. At least under certain conditions, a test patch having a reflectance corresponding to half the reflectance of another patch but the same luminance[96] should be judged as being under an illumination that is twice as strong as the second patch's.

---

[94] "Interreflections" refer to what is also called "mutual illumination". For any object in the scene, the light reflected is not only due to the illuminant but also to light reflected off of neighboring objects.

[95] Note that, although the chromatic version of the causal view in achromatic color constancy has been discussed (Brainard & Maloney, 2011; Foster, 2011; Zaidi & Bostic, 2008)), I am not aware of any study investigating the chromatic counterpart of the invariance hypothesis.

[96] Remember that in Mondrian scenes luminance is achromatic reflectance multiplied by the amount of incident light. It is thus equivalent to the color signal in color constancy research and can be called the "light signal".

The relationship between the invariance hypothesis and the causal view

Depending on how we understand the invariance hypothesis, it can be compatible with the causal view. If we hold that it only consists in the claim that there is an invariance relationship between lightness and perceived illumination given a certain luminance, then it is entirely compatible with the causal view. The latter would merely add that the invariance relationship holds because illumination estimation is a causal prerequisite to reflectance estimation. However, if, to the invariance claim, we add that reflectance estimation and illumination estimation are distinct results of the same processes, and that neither of them underlies the other, then the causal view becomes indeed incompatible with the invariance hypothesis.

In addition, it is useful to consider the experimental results regarding the invariance hypothesis if we care about whether the causal view is true. Certain consequences of the invariance hypothesis are also consequences of the causal view and an empirical rejection of those consequences would therefore be a strike against both hypotheses, regardless of how we conceive of the invariance hypothesis. We thus have to be clear on the nature of the empirical consequences of the two theses.

First, both hypotheses entail that the conditions which favor illumination estimation also favor reflectance estimation. The invariance hypothesis entails it because both estimations are the results of the same processes. If those processes can be effected adequately (i.e. if the right scene conditions are in place and the visual processes can use the available cues), then the visual system will estimate both the reflectance and the illumination adequately. The invariance hypothesis states that the visual system can detect, for a given luminance, that a change in lightness ought to be accompanied by a change in perceived illumination. That is, in situations in which it can detect a change in lightness, it can also detect a change in illumination. Consequently, the conditions under which a change in lightness can be detected are also the conditions under which a change in illumination can be detected. It is also a consequence of the causal view because the latter states that a correct estimation of reflectance presupposes a correct estimation of the illumination and thus presupposes the conditions which favor the estimation of the illumination.

However, a little subtlety must be exercised in the case of the causal view since it can be interpreted in two different ways. The exclusivity hypothesis states that the only element needed by the visual system in addition to the color signal is an estimate of the illuminant. In that case, it is true that the conditions which favor illuminant estimation also favor reflectance estimation because there is nothing to estimate but the illuminant before the reflectance can be estimated.

If we consider the partial cause hypothesis, things are a little more complicated. In that case, illumination estimate is only one of the elements needed by the visual system to estimate reflectance. The conditions for illumination estimation might be present in a scene which lacks other features necessary for reflectance estimation. For instance, a particular scene might satisfy the conditions which favor illumination estimation but not those required for shading or haze estimation. Consequently, conditions which favor reflectance estimation will include, but will not be limited to, conditions which favor illumination estimation.

In those conditions, evidence that speaks against the invariance view might not speak against the partial cause hypothesis. For instance, it has been argued that there is more than one kind of lightness constancy, that is, more than lightness constancy with regard to illuminant changes:

> "When the same piece of gray paper is viewed successively against different backgrounds, the luminance ratio at the edge of the paper changes dramatically, yet the paper appears to change very little in lightness […] This constancy of lightness with respect to changing background has been labelled Type II constancy, to distinguish it from lightness constancy with respect to changing illumination, or Type I constancy. W. Ross and L. Pessoa […] have proposed the more memorable terms illumination-independent constancy (Type I) and background-independent constancy (Type II)."(Gilchrist et al., 1999)

A scene might thus satisfy conditions which allow constancy with regard to illuminant changes, that is, allow correct estimation of the illumination, but lack the features required for background-independence constancy. Our performance in such scenes would be evidence against the invariance hypothesis yet not against the partial cause hypothesis.

This is arguably what has been found by Rutherford and Brainard (Rutherford & Brainard, 2002). They hypothesized that, in a two-stage algorithm framework (in our case, a reverse optics framework), the relation between the estimate of the illuminant and the estimate of the lightness of surfaces is one-to-one. That is to say, for a given proximal stimulus (i.e. a given luminance), only one estimate of reflectance should correspond with a given estimate of the illuminant and vice-versa. Moreover, they assumed that the perceived illuminant (i.e. how we explicitly judged the illuminant to be) is determined by the very same estimate of the illuminant that is used to estimate lightness in a two-stage algorithm framework. Consequently, a change in the perceived illuminant signals a change in the implicit estimate of the illuminant and, in turn, a change in the estimate of lightness.

This is reminiscent of Gilchrist's comment that, according to the causal view, "if the luminance of a target is held constant, an increase in its perceived lightness should be accompanied by a decrease in its perceived illumination, and vice versa" (Gilchrist, 2006, pp. 217–218). This is the second of the consequences shared by both the causal view and the invariance hypothesis and this is the reason Rutherford and Brainard's article actually targets both hypotheses[97]. We shall later see how this conclusion needs to be refined. Following Rutherford and Brainard's assumptions, we can understand how they tested those hypotheses. If two scenes are judged as having the same illuminant (i.e. if the perceived illuminants for the two scenes are identical) and if two surfaces are estimated as having the same lightness, then the light reflected from those surfaces should be identical.

Subjects were seated in front of two experimental chambers which they could see in alternation. One of the chambers, the standard chamber, contained a variety of achromatic objects and its far wall was composed of two halves of different lightness (one darker than the other). On one of these halves, a target patch was set. The other chamber, the "match" chamber, contained the same variety of achromatic objects, walls arrangement and patch set up so that the two chambers were roughly mirror images of each other. However, whereas the reflectances in the standard chamber were of high or medium reflectance, the reflectances in the match chamber were of medium or low reflectance[98].

Participants were first asked to adjust the illumination in the match chamber so that it matched the illumination in the standard chamber. Once they judged the perceived illuminants to be equal across the chambers, they were invited to adjust the lightness of the test patch in the match chamber so that it matched the test patch in the standard chamber. If the albedo hypothesis holds, then after both the perceived illuminants and the lightness of the two target surfaces were equated, the luminance coming from such surfaces should be identical. This was contradicted by the results of the experiment. Even though the subjects judged that the illuminants and reflectances of the target patches were identical, the lights reflected from those patches were not.

A second experiment was designed to challenge the albedo hypothesis somewhat further. If the perceived illuminant is, in addition with luminance, the only element used to compute the estimate of lightness, then it must not be possible to manipulate the estimate of

---

[97] Note that Rutherford and Brainard actually give the name "albedo hypothesis" to what we called reverse optics frameworks.

[98] The difference in reflectance of the objects across scenes was meant to avoid the possibility that subjects would just match luminances from identical surfaces in the scenes. Including differences ensures that subjects could not do so.

lightness without also manipulating the perceived illuminant. In this experiment, the two different halves of the far wall were swapped for one another in the match chamber. This was not supposed to have any effect with regard to the matches if the albedo hypothesis indeed held true.

However, it did have an effect. The luminances of the target patches in the match chamber were set higher than the corresponding luminances in the first experiment. That is, even though the perceived illuminants and perceived lightness were matched and were thus identical with the first experiment's, the luminances from both experiments differed. This should not be possible if, as the exclusivity hypothesis holds, the perceived illuminant uniquely determines lightness given a proximal stimulus.

Rutherford and Brainard thus found an experimental situation in which the subjects' performance failed to be independent from the background's properties. Even though an estimate of the illuminant might have been involved in their judgments of the reflectance of the target patch, other features of the scene, most noticeably other features from the background, influenced those judgments. Their performance failed to be constant with respect to the background (i.e., it did not display background-independent constancy). This clearly represents evidence against the exclusivity hypothesis.

Nevertheless, this does not constitute evidence against the partial cause hypothesis. That the subjects' performance varied in relation with factors distinct from the illuminant does not speak against the estimate of the illuminant being one factor which is used in the computation of the reflectance estimate. Those results may have been expected. It has been known for a long time that there are simultaneous contrast effects at work in perception. How light (or red or blue or what have you) a surface is judged to be depends on the surrounding surfaces in a way that is inconsistent with the exclusivity hypothesis. The link between the invariance hypothesis and the causal view thus varies depending on how we understand the latter, be it as the exclusivity hypothesis or as the partial cause hypothesis.

The evidence regarding the albedo hypothesis

Now that we have a better grasp on the various claims which have been proposed under the label "albedo hypothesis", we are in a position to assess the evidence which has been stacked against or for them. We already saw that Rutherford and Brainard's results do not actually count against all hypotheses equally. Even if we admit that the perceived illuminant correctly represents the implicit illuminant estimate used in the computation of lightness, the performance of the subjects is only incompatible with the exclusivity and the invariance hypotheses[99].

To begin with, Kozaki and Noguchi (Kozaki & Noguchi, 1976) selected test patches whose luminances were approximately equal. That is, although the reflectance and illumination of the patches varied, the sum of a patch reflectance and the illumination falling on it was the same throughout the set. Participants had to judge the lightness and illumination of each patch on a scale of increments, going from "very black" or "very, very, very dim", respectively, to "very white" or "very, very, very bright", respectively.

The authors predicted that, if the invariance hypothesis holds, then the sum of the lightness and perceived illumination should be the same for a given luminance level. For instance, if a test patch is deemed pretty light, then the illumination on it should be judged pretty dim. The subjects' performance coincided with those predictions and thus supported the invariance hypothesis. However, Kozaki and Noguchi claim that their results are incompatible with the albedo hypothesis understood as the causal view. If, as the latter claims, the computation of lightness requires an estimate of the illuminant beforehand, then illuminant estimation should be better in the conditions in which lightness constancy is better. Though those conditions occurred in the study, illuminant estimation was not better in these compared to other conditions. If illumination estimation is the only variable used to compute lightness (in addition with luminance), then better lightness constancy should occur in tandem with better illumination estimation.

Note, though, that the challenge would have been fiercer if distinct findings had been observed. If illumination estimation had been better under some conditions under which lightness constancy was not improved, then this would have put more pressure on the albedo hypothesis, since it would have shown that correctly estimating the illuminant is not as important as one might think for correct lightness estimation. In the present case, however, there are a variety of rejoinders available to the defender of the albedo hypothesis.

One could argue, in the same vein as one could argue against Rutherford and Brainard above, that the perceived illuminant is not the same as the estimate of the illuminant. That is, one could argue that the estimate used to compute lightness is not identical with, or does not uniquely yield, the explicit judgment of the illuminant. If one could measure the implicit

---

[99] Obviously, it would be important to see a replication of this study. The lack of replication is a fundamental issue in science, especially in color and lightness constancy research. Another is the use of a very small number of participants, an important part of which is the authors of the study (see, e.g., Kelly et al., 2018; Logvinenko & Maloney, 2006; Morimoto et al., 2021; Morimoto & Smithson, 2018; Robilotto & Zaidi, 2004a; Zaidi & Bostic, 2008).

estimate, one could see that the conditions under which it would be easier to compute would be the same conditions under which lightness constancy would be better. However, this answer presupposes that the perceived illuminant is not the same as the estimate of the illuminant and it is unclear what this comes down to.

Perhaps the most obvious rejoinder available is that the results actually only undermine the exclusivity version of the causal view. If the illuminant estimate is the only variable used to compute lightness, then situations in which lightness constancy is better should indeed be the same ones in which illumination estimate is better. However, if one considers the partial cause hypothesis, other variables will be taken into account and we might thus find a discrepancy between those two kinds of situations. This is especially true in this study since Noguchi and Kozaki actually found pretty high correlations (above .977 for all three backgrounds) between the estimate of the illumination and the illumination itself. Perhaps the small differences in lightness constancy that were found across background albedos were not due to distinct illuminant estimates, since those were already correct with the backgrounds studied, but to other factors whose detection was better in the white background condition[100].

In another experiment, Noguchi and Kozaki tried to determine whether subjects could display lightness constancy with regard not to the test patch itself but to its background (Noguchi & Kozaki, 1985). Observers had to judge either the lightness of the background or the general illumination (in a given trial) on scales which were analogous to those of the above experiment. On the background were patches that were either white or black and the backgrounds themselves varied from dark gray to almost white. The illumination was varied in its intensity so as to give five different illumination levels. If subjects showed lightness constancy with respect to the background, then the lightness judgments of said background should not vary with the illuminance of the display.

Subjects were found to be more constant with the backgrounds with the white patches rather than with the black ones. However, independently of their lightness constancy score, what is noteworthy is that there was a discrepancy in the variables that affected lightness judgments and those affecting illumination judgments. Both kinds of judgments were independently fitted with least-square straight lines across the five levels of illumination. Those lines were then evaluated according to which factor influenced their slopes or y-intercepts.

---

[100] Or they were due to other factors which were more disruptive under certain conditions than others.

Lightness judgments lines were affected by the color of the patches with regard to their slopes (white patches yielded smaller slopes), whereas their y-intercept was dependent on the background albedo (higher reflectances (i.e. whiter backgrounds) yielded higher y-intercepts). In contrast, the slopes of illumination judgments lines were not influenced by either the patch color or the background albedo and were thus simply determined by the illuminance of the display. Their y-intercepts were dependent on the background albedo (higher albedo yielded judgments of higher illuminance) and on the interaction between patch color and background albedo.

The fact that there was a discrepancy between the factors influencing the lightness judgments and those influencing the illumination judgments is, according to the authors, incompatible with the albedo hypothesis. If the latter were true, lightness estimation would depend entirely on illumination estimation and, consequently, factors influencing lightness estimation should be identical with those influencing illumination estimation. But this was found not to be the case. The albedo hypothesis is thus undermined.

Again, note that those results only undermine the albedo hypothesis understood as the exclusivity hypothesis, while the partial cause hypothesis remains safe and sound. If lightness constancy depends on more than just the illuminant estimate, it is perfectly possible for there to be a discrepancy between the factors influencing one and the factors influencing the other. For instance, in the study, simultaneous contrast effects could have had an impact on the final lightness judgment, thereby explaining why they were influenced by the color of the patches, without having a similar impact on the illuminant judgments. Importantly, we need results of a study in which there is an increase in the quality of the illuminant estimate without a concurrent increase in lightness (or color) constancy[101]. Barring this, it is difficult for an experiment to be relevant to the partial cause hypothesis.

In the present section, we have seen how reverse optics approaches to vision have empirical consequences which can be tested for. The central aim was to argue that, although testing them is arduous, the very fact that it is possible to do so appears to undermine Chirimuuta's case that it is not possible to experimentally falsify or confirm research programs or abstract conceptions of vision or, more specifically, theses such as *Recovery* and the albedo hypothesis, which are important components of those programs. In the next section, I present data which take a different route. Rather than testing a specific empirical

---

[101] Lightness constancy or color constancy should not be close to perfect in that situation because it would make it harder to detect improvements. Neither should illumination estimation be too good since this could preclude meaningful improvements from happening.

consequence of an approach to vision, some studies directly compare color constancy models of distinct conceptual ancestry in a head-to-head match. This lends further weight to the thesis that such models can be experimentally compared without the need for us to endorse one or the other to read the data in one way or another.

Experimental comparison of color constancy models

Up until now, we have tried to delve into the notion of reverse optics to see whether we could assess Chirimuuta's opinion that the conceptual assumptions at play in color constancy research are not experimentally evaluable. We saw that adopting the notion of reverse optics usually entails a commitment to the albedo hypothesis which itself appears to make experimental predictions. Those predictions were put to the test repeatedly in numerous studies. We concluded that rejecting the albedo hypothesis would experimentally put pressure on reverse optics frameworks in general and *Recovery* in particular. This observation runs counter to Chirimuuta's thesis that we cannot rule out models in research on color constancy on the basis of experimental falsification without previously endorsing a particular conception of vision. The data in color science is not theory-laden in that strong a sense.

To advance our case further, there is another route available to put into question the latter thesis. Remember that Chirimuuta holds that, when it comes to Boyaci, Maloney and Hersh's article (Boyaci et al., 2003) and Robilotto and Zaidi's (Robilotto & Zaidi, 2004a):

> "Crucially, one group's model cannot be tested with the other group's data. With the current state of the evidence, the two models—and therefore, the two stances with respect to lightness recovery—cannot be compared directly against each other." (Chirimuuta, 2008, p. 572)

I believe that this amounts to the claim that a reverse optics model cannot be tested with Robilotto and Zaidi's data and that a heuristics model cannot be tested with Boyaci *et al.*'s data. This thesis is an important one in Chirimuuta's project to show that different lines of research in color constancy are experimentally isolated.

In this section, I will argue that certain experiments actually compare models of distinct conceptual ancestry. That is, rather than determining whether a consequence of some of the models (in this case, the reverse optics models and the albedo hypothesis) is borne out by the data, those experiments directly compare the predictions of those models without making use of a particular hypothesis such as the albedo hypothesis. Notwithstanding which model is declared the winner, if it is the case that we can directly compare those models, this constitutes a rejection of Chirimuuta's claim that we cannot compare conceptually distinct models directly in an experimental setup.

Robilotto and Zaidi

When we discussed Chirimuuta's position, we encountered the work of Qasim Zaidi's team (Robilotto & Zaidi, 2004a, 2006; Zaidi, 1998, 2000; Zaidi & Bostic, 2008). Zaidi and his colleagues used real scenes and a forced-choice paradigm to investigate the performance and strategies of observers regarding both lightness constancy and color constancy. Those observers sat in front of a display containing two sides, each side independently illuminated by a light different than the other side's. Each side contained two cups wrapped in paper. Three of the four cups were made of the same paper, the Standards, while the last one was of a different color or lightness, the Test. One side thus contained two Standards while the other one contained one Standard and the Test. Observers had to select the cup that was made of a different paper, that is, the Test. In order to do so, they first had to determine which side contained the Test by comparing cups under the same illuminant. When they had chosen the side, they had to compare both of the cups on this side to the cups in the other scene, that is to say, they had to perform a comparison across illuminants.

At least in their first experiment on the matter (Robilotto & Zaidi, 2004a), Zaidi modelled the results of observers according to three strategies. A first strategy, the "photometer-based strategy", has observers select the object that was most different in brightness (here, total amount of light reflected). Observers would first determine which side contained objects reflecting different amounts of light and then select the object most different in brightness. Observers using this strategy are only dealing with estimates of luminances and not of reflectances, that is, they are using brightness estimates. The second model of Zaidi's team adds adaptation to it so that the brightness of a stimulus is adjusted as a function of the mean intensity of the scene. The third model has observers estimate reflectance and illuminant and thus counts as a reverse optics model.

Although Robilotto and Zaidi claim that a photometer-based strategy with adaptation better accounts for their results, it is difficult to agree. **Figure 1** plots the pattern of responses of observers who would follow a photometer-based strategy with adaptation being set at a certain value. Three such values are displayed, based on the amount of adaptation taking place. The authors maintain that graph (b) actually mimics the actual answers of their participants quite well.

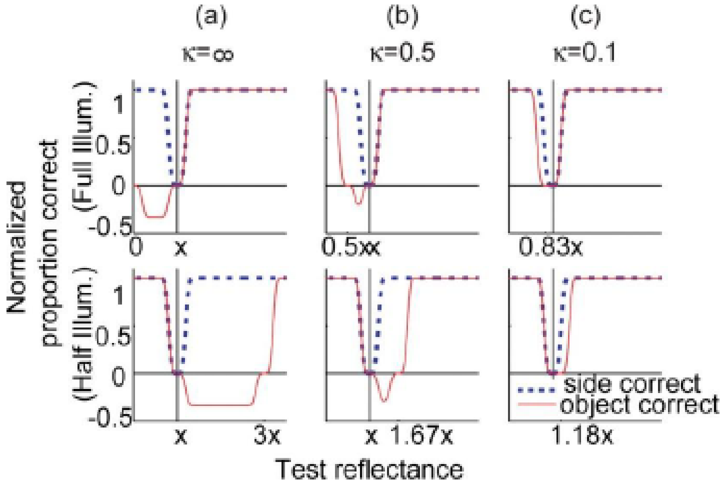

Figure 1. Normalized proportion of correct responses based on a model of brightness dissimilarity plus adaptation. (a), (b), and (c) represent models based on three levels of gain with decreasing adaptation κ values. Standard reflectances are denoted by x. Reproduced with permission.

Figure 2 represents the pattern of responses of observers who would follow an inverse optics strategy with a particular estimation of the between-backgrounds luminance ratio. Consider the graph which is second from the left. This strategy has observers slightly misestimating the ratio between the illuminants' intensities. Instead of finding out the correct ratio of 2 to 1 (so that the intensity in the full illumination scene is twice that of the half illumination scene), this strategy presupposes that subjects compute a ratio of 1.5 to 1 (i.e., the illumination in the full illumination scene is deemed only 1.5 times higher than the one in the half illumination scene).

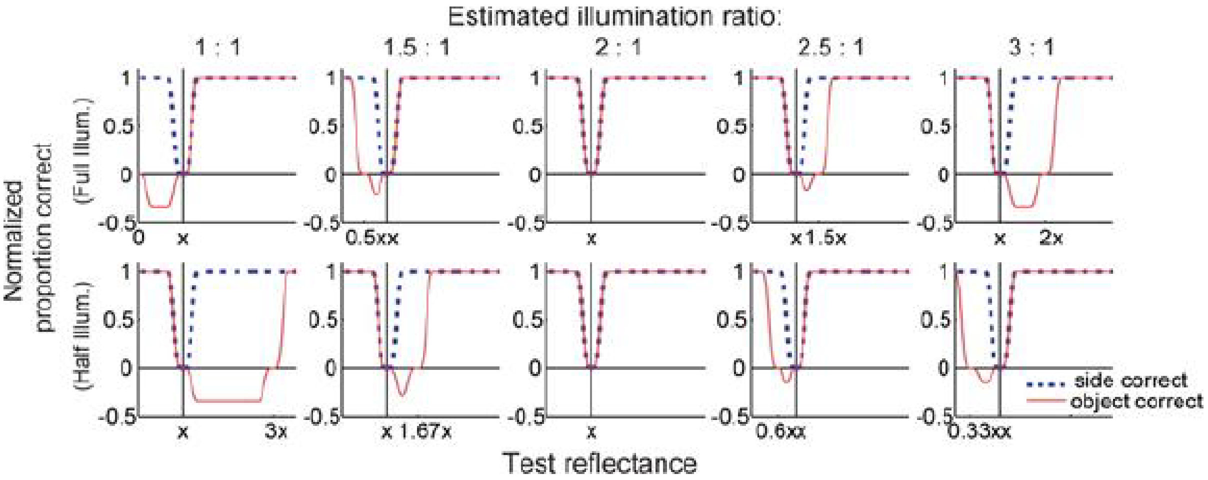

Figure 2. Normalized proportion of correct responses based on a model where the illumination ratio between the two compartments is estimated and used to factor out object reflectance. Reproduced with permission.

Notice that this pattern of responses is indistinguishable from graph (b) in **Figure 1**. In other words, observers would generate the same pattern of answers, regardless of which of the two strategies they followed.

If this is indeed an accurate portrayal of what is going on, then it becomes hard to agree with Robilotto and Zaidi's claim that their results are better predicted by a photometer-based strategy with adaptation than by an inverse optics strategy. Barring any evidence of a different kind, any evidence in favor of one of the models is evidence for the other. That is, the selection data alone underdetermine which strategy the perceptual system of the observers follows.

Independently of which of the three models the authors presented better accounts for the observers' performance, the central point is that theirs is a study comparing a model involving reflectance estimation as a central step to a model only using luminances, sitting more comfortably on the "heuristical" side of strategies. Chirimuuta concurs with that statement (Chirimuuta, 2008, p. 568). However, this is exactly what, she argued, cannot happen: two conceptually distinct models experimentally compared.

A rejoinder available to her is that it is unclear that the inverse optics strategy the authors model is indeed a strategy involving inverse optics. Remember that we understood an inverse (or reverse) optics strategy as a strategy involving the recovery of a physical trait which was at the source of the sensory signal involved. Given the luminance of a cup as the sensory signal, the recovery of the physical trait at the source of the signal would here consist in the recovery of the reflectance of the cup, i.e., an estimation of its lightness. If, e.g., the cup reflects .7 of the light it receives, then the visual system should try to compute an estimate of this value in order to be said to be applying an inverse optics strategy.

However, this is absolutely not what the visual system is thought to do according to Robilotto and Zaidi's proposed inverse optics strategy:

> "Many models of lightness and color constancy assume that the visual system estimates the scene illuminant and uses this estimate to determine material reflectance […]. In terms of this illumination-estimation hypothesis, the questions arising about this study are, what information is available in the display, and what is the best any visual system could estimate? Backgrounds in the two boxes have similar statistics. Theoretically, a system could calculate the mean luminance of both backgrounds, then measure the mean luminance of each of the four cups, and take the ratios."

According to this strategy, the visual system computes ratios of luminance between backgrounds and cups, respectively, and then tries to determine which of the ratio is not identical with the other three. But the issue here is that, by doing so, the visual system does not estimate the reflectance of any surface in the scene. It only deals with luminances and luminance ratios. It therefore cannot be said to apply an inverse optics strategy. Comparing ratios of luminance rather reminds us of Foster's strategy of comparing color relations between different surfaces across illuminant change (e.g. Craven & Foster, 1992; Foster, 2011; Foster & Nascimento, 1994; Linnell & Foster, 1996), that is, of relational color constancy[102].

In addition, we understood reverse optics models as necessarily involving a recovery of an estimate of the illuminant. Here, however, the visual system is thought to try and determine merely the ratio of the illuminants' respective intensity. But this ratio can be compatible with a number of pairs of illuminant intensities. That is, the visual system computing such ratios does not single out any illuminant in particular; it just puts restrictions on what the intensity of one illuminant could be, given the other one.

Finally, a reverse optics model's output is that the reflectance of an object in the field of view has such-and-such reflectance. Robilotto and Zaidi's depiction of reverse optics postulates that the model's output is to "equate the reflectances of the standards across compartments" (Robilotto & Zaidi, 2004a, p. 784). That is, rather than reaching a conclusion regarding which wavelengths a particular reflectance preferably reflects and which ones it does not, the visual system is thought to reach a conclusion regarding which reflectance in the scene is identical with which one in the other scene. This is a consequence of the intermediate conclusion that a given reflectance in one scene stands in some relation with another reflectance in the other scene that is identical to the relation of the backgrounds of the two scenes with one another. More precisely, if one background is found out to reflect, e.g., twice the amount of light the other background reflects, this intermediate conclusion will be that such reflectance in one scene also reflects twice the amount of light as this other reflectance in the second scene. The conclusion of the visual system is thus not conceived as being that a given surface or object in the field of view has such-and-such reflectance.

It could thus be argued that Robilotto and Zaidi's experiment does not actually compare a reverse optics model to a heuristics model. However, this conclusion does not lead us as far as Chirimuuta would like to. Even though it can be argued that the model Robilotto and Zaidi maintain involves reverse optics does not actually involve reverse optics, it still

---

[102] Remember that Foster has explicitly rejected inverse optics strategies (Foster, 2003, 2011).

represents a model that is different from the one they favor. The question which remains is therefore whether this counts as the right kind of difference.

If one restricts oneself to comparing models with different levels of adaptation, it might be argued that they do not count as different models. This is because those two models would conceive of perception in the exact same way. For two models to count as distinct models in our case, the models must make distinct assumptions regarding the nature of perception. We saw that reverse optics conceive of perception as the solution to inverse problems and thus as essentially involving the recovery of properties of objects in the scene.

However, we must be careful regarding how we read "properties of objects" here. In Zaidi's team's favored model, the visual system is thought to recover the brightness of objects after adaptation to the illuminant. Though it would be a transient one, the brightness of objects can be conceived of as a property of the objects of the scene. Both this model and a reverse optics one would then see vision as involving the recovery of properties of the objects of the scene and thus not count as distinct models, although they are clearly thought as so.

Maybe the key lies in the fact that the reflectance is, to speak roughly for the time being, a more robust property of objects than their brightness[103]. The latter changes with the illumination, conditions of viewing, adaptation of the visual system, and so on. In contrast, an object will have a stable reflectance in most conditions[104]. Consequently, tracking objects by using their brightness is much more suboptimal than tracking them by using their reflectance. Of course, this worry could be assuaged by considering the relative brightness of objects. That is to say, although the brightness of a particular object will vary across viewing conditions, illumination level, etc., the ratio between its brightness and a neighboring one will stay the same. A model could aim at recovering those ratios instead of the brightness (even after adaptation) of objects. This is reminiscent of Foster's relational constancy (see, e.g., Craven & Foster, 1992; Foster & Nascimento, 1994; Linnell & Foster, 1996; Nascimento & Foster, 2000). However, this might be an unwelcome change to the theory. Computing relative brightness may be too similar to the computation of ratios involved in the "reverse optics" model proposed by Zaidi.

It is not clear that this is the right way of thinking of "distinct models" according to Chirimuuta. According to her, the important difference between the two models at issue is that the reverse optics model construes perception as the recovery of the physical properties at

---

[103] I here use the term "robust" in order to avoid the debate around intrinsic and non-intrinsic properties.

[104] Prolonged exposure to sunlight will change the reflectance of many objects but, over a short time span, this does not impact it meaningfully.

the source of the color signal, whereas Zaidi's model sees perception as a way to "get by" in the world, a way to use the color signal in the most efficient way to deal with one's environment. As I do not have convincing answers to the present question, I shall leave the matter aside, at least for the time being.

Robilotto and Zaidi actually replicated their experiment with achromatic patterned cups (Robilotto & Zaidi, 2006) and Zaidi performed a similar experiment with plain chromatic cups (Zaidi & Bostic, 2008). In the first study, Robilotto and Zaidi argued that their results again favored models based on perceptual heuristics (such as the one based on perceived brightness with adaptation in their former study) because they were not well accounted for by a reverse optics model. At first glance, it appears that we have here another instance of a study comparing two distinct models, which is what Chirimuuta argued could not happen.

However, it is worth looking more deeply into Robilotto and Zaidi's comparison. Note that, again, the reverse optics model they describe does not seem to be a reverse optics model as traditionally understood. It is conceived in the same way as it was in their former study (Robilotto & Zaidi, 2004a). Contrarily to this study, however, Robilotto and Zaidi did not model the predicted results for this "reverse optics" strategy. This in turn makes it harder to say with confidence that their results were better predicted by a perceptual heuristics model than by the "reverse optics" one[105].

One of the reasons they assert such a conclusion is that their "results show that despite naturalistic binocular viewing of information-rich situations, under many conditions, observers cannot identify objects of the same reflectance across different illumination levels" (Robilotto & Zaidi, 2006, p. 33). In order for this criticism to go through, it must be the case that the use of reverse optics strategies entails a perfect performance in tasks similar to theirs. However, it is something we do not have any reason to accept. Even if the visual system does use a reverse optics strategy, it does not entail that its performance in color constancy tasks should be perfect across the board. This is something Robilotto and Zaidi seemed to agree with, albeit somewhat reluctantly, in their former study when they discuss reverse optics and the possibility that their subjects employed perceptual strategies involving them in their study:

> "In terms of this illumination-estimation hypothesis, the questions arising about this study are, what information is available in the display, and what is the best any visual system could estimate? Backgrounds in the two boxes

---

[105] For the same reason, Maloney's study was in a poor position to say that his model was *better* than other models at predicting their results.

> have similar statistics. Theoretically, a system could calculate the mean luminance of both backgrounds, then measure the mean luminance of each of the four cups, and take the ratios. According to this design, there is sufficient information to perform the lightness identification task perfectly […], but it is clear that observers are unable to do this. Observers can tell which side has the brighter illuminant, but *either they are unable to calculate the ratio exactly* or else they cannot put it to use." (Robilotto & Zaidi, 2004, p. 792) (my italics)

Note that they admit that observers could be unable to calculate the ratio exactly[106]. But the very fact that there probably are no conditions in which color constancy could be perfect is something that is well agreed upon (see e.g. Foster, 2011; Smithson, 2005). More specifically, it is agreed upon that any mistake in the estimation of the illuminant in reverse optics models would lead to a distorted estimate of the reflectances.

It thus appears that this second study is somewhat less complete than Robilotto and Zaidi's first, especially because of the lack of modelling of the reverse optics strategy's results. Zaidi and Bostic's study (Zaidi & Bostic, 2008) apparently suffers from the same flaw and essentially replicated the 2004 study but used chromatic cups instead of achromatic ones. The perceptual strategy they propose to account for those is that the visual system uses relative colors of the cups to determine which one is the odd one.

Participants first completed the find-the-odd-one-out task which was analogous to the task in the earlier studies, except that cups and illuminants varied chromatically. They were then seated in front of only one scene containing two distinct cups and asked to rate whether one of them was predominantly greener, redder, yellower or bluer than the other one. It was then analyzed whether the relative colors of the cups correlated with performance in the find-the-odd-one-out task. For tests under the blue-green light, the proposed strategy predicts that:

> "Tests perceived as bluer or greener than the Standard under the same light, will be picked correctly as the odd object, but Tests perceived as yellower or redder than the Standard under the same light, will not be picked as the odd object, unless they are perceived to be at least as yellow or red as the Standards under the yellow-red light in the other box. For Tests under the yellow-red light […] Tests perceived as yellower or redder than the Standard under the same light, will be picked correctly as the odd object in the identification experiment, but Tests perceived as bluer or greener than the Standard under the same light, will not be picked as the odd object unless they are perceived to be at least as blue or green as the Standards under the blue-green light in the other box." (Zaidi & Bostic, 2008, p. 2680)

---

[106] The fact that one may argue that this is not properly speaking an illumination-estimation hypothesis that is at issue here (such as the causal view we discussed) is not relevant. What matters is that Zaidi and colleagues believe that it is.

The response pattern of subjects who would follow this strategy predicted 88% of correct identifications and 93% of misidentifications actually found in the experiment.

Whereas the relative colors strategy well predicted the pattern of results of the subjects in the find-the-odd-one-out task, the authors argue that a strategy using reverse optics could not have done as well because of one particular result. In a cone opponent space, the correctly identified and misidentified tests fell on opposite sides of the standard patch's location and switched side across illuminants (see **figure 3**). That is, the tests which were correctly identified under one of the illuminants were misidentified under the other illuminant. Whereas their favorite strategy, based on relative colors, could account for this qualitative result, there is no reason to expect it according to reverse optics strategies.

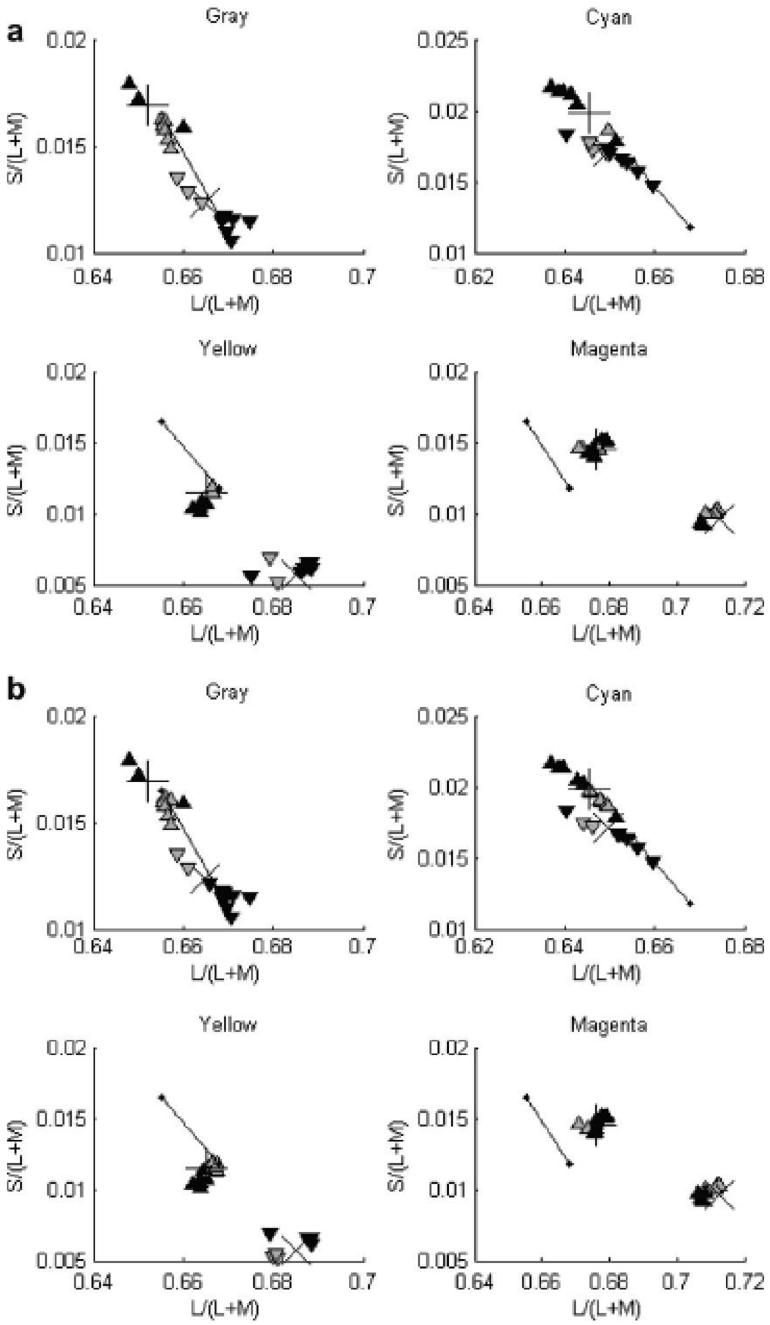

Figure 3. Each panel represents identification results for each kind of standard cup under each of the two lights for tests that could be reliably discriminated from the standard cup for observer a and observer b. The color name at the top of each panel specifies the color of the standard cup in that condition. The symbols "+" and "X" represent the chromaticity of the standard cup in that condition under the bluish-green and yellowish-red lights, respectively. Triangles pointing up are the chromaticities of the tests under the bluish-green light, whereas those pointing down are the chromaticities of the tests under the yellowish-red lights. Dark triangles represent tests that were correctly identified, whilst gray triangles represent those that were misidentified. Reproduced with permission.

Unfortunately, it is not clear that this 'result' of theirs counts in fact as such. Without any statistical test to determine unambiguously whether correctly identified and misidentified tests switch sides across illuminants, it is hard to agree that one can just see that this is so. If we consider **figure 3**, it is clear that some of the graphs do present such an asymmetry in the results[107]. However, probably as many seem to present results which do not appear to be easily coerced into that interpretation. The problem is that so much of their argumentation against most theories of color constancy in the field revolves around this supposedly obvious result, which means it would be damaging for their case were it to turn out spurious.

It still remains that, not unlike Robilotto and Zaidi's second study, this study fails to model any other strategy than their favorite.[108] This makes it harder to unequivocally conclude that theirs is the only strategy which could well account for their results. Of the three studies we just analyzed, only the first one modelled predictions of the subjects' results according to two different perceptual strategies. Even though it might yield a piece of the puzzle, we can and ought to look somewhere else.

Olkkonen & Allred

More recently, there have been studies which try to determine the relationship between the processes responsible for color memory and those responsible for color constancy. Most of the time, it is assumed that color memory and color constancy are independent processes. However, this independence assumption is rarely questioned. Olkkonen and Allred (Olkkonen & Allred, 2014) performed an experiment designed to test whether it holds.

Subjects were seated in front of a rendered display which contained two squarish stimuli, ranging from yellowish-green to bluish-green, each on their own background. The stimuli were either presented with a 2-seconds delay or simultaneously and on different or identical backgrounds. There were thus 4 distinct experimental conditions. The delay was meant to analyze the effect of memory, whereas the change in background was supposed to mimic, and thus be interpreted as, a change in illuminant and represented the color constancy conditions. The subjects executed a forced-choice task. They had to decide which of the stimuli they had just seen was the bluer one. For instance, in the joint condition (that is, combining memory and color constancy), the subjects would first see a stimulus under an achromatic illuminant with the other empty scene under the blue illuminant next to it. The stimulus would disappear and subjects would see the empty scenes for 2 seconds. The test

---

[107] The reader is thus invited to focus on whether, for each panel, the triangles representing the correctly identified tests under the bluish-green light end up misidentified under the yellowish-red light and vice versa. I wager that this will be easier said than done.

[108] Note, again, that this is not a trend that is restricted to studies favoring heuristic strategies. For studies which favor reverse optics and fail to model anything else, see, e.g., Boyaci et al., 2003.

stimulus would then appear in the scene under the blue illuminant and the subjects would have to determine which of the stimuli was bluer than the other one.

After the subjects completed the task, it could be determined what stimulus were chosen as bluer in a given condition. For instance, in order to be seen as bluer than the reference on the grey background (i.e., under the achromatic light), a stimulus on the blue background (i.e., under the blue light) had to send back much more blue light than a stimulus that would also have been on a neutral background. That is to say, the response of a color constant subject in a given trial would involve a compensation for the bluer 'illuminant'. In that case, the stimulus that would be chosen as equally blue as the one on the neutral background (i.e., the one that would be chosen 50% of the time as bluer and the rest of the time as less blue[109]) would send significantly more blue light than the latter[110]. In that way, it could be determined what amount of bias was applied to the reference stimulus in the scene with the different background (or delayed, in the memory condition).

Olkkonen and Allred's line of thinking was the following. If there is no interaction between color memory and color constancy, then the bias in the experimental condition consisting of both the delay and the change in illuminant (by a change of the background), the "joint" condition, should amount to no more and no less than the sum of the biases in the memory condition and the results of the constancy condition. However, this is not what was found. The bias in the joint condition was much smaller than the predicted bias and was actually not very far from the baseline value. That is to say, subjects were less constant in the joint condition as compared to the constancy condition. The addition of a 2-sec delay worsened color constancy. Therefore, it seems that there is an interaction between the processes respectively responsible for color constancy and memory.

In addition, and most importantly for our purposes, Olkkonen and Allred maintain that it is unclear how a model based on contrast coding, albeit able to account for simultaneous conditions, would explain memory effects[111]. They argue that a model based on a Bayesian approach and essentially involving reverse optics could qualitatively account for all of their results. However, as they do not model the predictions of those two strategies, it is unclear

---

[109] That is, if forced to choose, a subject would select randomly, giving a percentage of 50% choice for each stimulus.

[110] Of course, subjects could have overcompensated for the simulated illuminant change but this was not observed in this study.

[111] The authors are not very forthcoming in what constitutes a model based on contrast-coding. However, if we consider the source the authors give (Shevell & Wei, 1998), then contrast coding refers to the coding of chromatic contrasts at the edges of a target with its background.

whether this counts as a win for reverse optics strategies. Fortunately, a later study remedies that disadvantage.

A few years later, Olkkonen, Saarela and Allred performed an experiment whose main goal was not to determine the relationship between memory and perception but to directly compare two distinct models regarding how well they could account for human data (Olkkonen et al., 2016). The experimental task essentially replicated their early article's with the minor difference that stimuli and backgrounds were achromatic rather than blue-green. Participants thus had to choose which stimulus appeared lighter in each condition.

The authors modelled two distinct strategies. The first one involved contrast-coding and thus consisted in computing local contrasts, in that case, the ratios of the center and surround luminance. That is to say, to which extent an object is seen as light or dark is determined by the ratios of center-to-surround luminance. The higher the ratio (i.e., the more light the center reflects in comparison to the surround), the lighter the object is seen. The second strategy represented reverse optics models and involved computing an estimate of the reflectance of both center and surround. In order to account for memory effects, both models were formulated probabilistically. That is, both of them used prior knowledge in combination with sensory evidence to reach a verdict concerning the trait it is supposed to compute, be it contrast ratios or reflectance estimates. This prior knowledge consists in information regarding the likelihood of encountering certain stimulus values in the environment and, combined with sensory evidence, yields a posterior probability of a given stimulus value. For instance, if a stimulus has a very high occurrence in the environment but is not very likely given the sensory measurement at hand, there is only an average likelihood of it being at the source of the sensory measurement. A second stimulus that is somewhat rare but has an extremely high probability of giving rise to the measurement might have a higher posterior probability (i.e. a higher likelihood of it being at the source of the measurement) than the first stimulus.

There were two variants of the contrast model, a simple and a complex one. The simple variant just learnt the value of the mean contrast across all experimental conditions and used it as a prior probability (or "prior") indiscriminately, that is, regardless of which condition it is in. The complex model adapted its prior in relation to the condition it was in. The reflectance model had two priors, one for the mean reflectance estimate of the background and one for the mean reflectance estimate of the center, but, as the simple contrast model, those did not change across conditions.

All three models were able to predict relatively well the performance of human subjects. However, there were a few interesting differences between those predictions. All models reproduced the simultaneous lightness contrast effect seen in the human data[112]. However, only the reflectance model was able to reproduce the magnitude of the effect. In addition, the authors investigated whether the independence condition for perception and memory held for their human data. The bias of the joint condition (context change and memory load together) was smaller than the bias predicted by the sum of the context change bias and the memory bias. That is, adding a memory component to the context change condition reduced the compensation effected by the visual system to offset the illumination change. Adding a memory component therefore reduced color constancy.

The simple contrast model did not display a reduction in bias for the joint condition. Its bias in that condition amounted to the sum of the bias in the memory condition and the one in the context change condition. In other words, the simple contrast model did not display any 'subadditivity' (i.e., the fact that the bias in the joint condition was less than the sum of the biases in the memory and context change conditions). The complex contrast model did display some subadditivity but the magnitude of the bias varied with the reference luminance in a way distinct from the human data's bias. Again, only the reflectance model qualitatively and quantitatively reproduced the pattern found in the human data. Overall, the authors concluded that the reflectance model better accounted for the human data.

It seems that we have here found an experiment which can compare distinct color constancy models to see which one can better explain our performance in color constancy tasks. Although exploratory in nature, this study is a direct counterexample to Chirimuuta's doubt that we can experimentally compare models with distinct conceptual roots. It appears that we do not need to endorse particular conceptions of vision to interpret at least some results of color constancy research as speaking against or for a given model.

Summary

In this chapter, we have first examined Chirimuuta's view of the relationship between ontological assumptions regarding the nature of vision on the one hand and color constancy data and experiments on the other. On that view, it is in particular possible to subscribe to *Recovery* in light of color constancy data only if one already adheres to a reverse optics

---

[112] Note that this is relevant to the interpretation of Rutherford and Brainard's study we touched on before. If reverse optics models can account for simultaneous contrast phenomena, it is not a strike against them if one can show that those phenomena occur by changing the background or anything else in the scene.

conception of vision. In addition, it appears that such adhesion automatically entails a commitment to some form of color realism. If this is true, Hilbert's argumentative strategy is threatened by a vicious circle.

Nevertheless, this picture entails that there can be no clear empirical consequences of reverse optics in general and of *Recovery* in particular which could be assessed experimentally. I argued in the following section of the chapter that color constancy research has already identified and experimentally studied one such consequence. The albedo hypothesis is viewed by foes and friends of reverse optics as an entailment of reverse optics models and any empirical results contradicting it would *modus tollens* stand in contradiction with this kind of models.

Finally, we have seen that another consequence of Chirimuuta's picture is that there can be no pitting against each other of the distinct conceptions of vision. At least in the recent years, and probably even in Zaidi's own work, there have however been such comparisons and this observation puts further pressure on Chirimuuta's thesis that those distinct conceptions cannot be experimentally compared.

In the next chapter, I will address Chirimuuta's point that reverse optics appears to be committed to a form of color realism. In my view, this may be true but it is not an objection to Hilbert's strategy. Theories of color constancy (and color appearance in general) can be seen as including operationalizations of various color ontologies[113] or, alternatively, color constancy theories can be seen as motivating a particular account of color ontology or at the very least certain ontological claims about color. I will try and show that we ought to revise our methodological approach to color ontology in light of this finding.

---

[113] Although talk of 'operationalization' may induce the impression that first there was the color ontology and then and only then there was the theory of constancy. In practice, it may be true sometimes but, in certain cases, the opposite may well be the case.

Chapter 5: On the methodology of color ontology

Introduction

In the last chapter, we saw how empirical evidence can infirm or confirm a particular research program in color perception, such as reverse optics. Despite Chirimuuta's worry, it is possible to test the empirical consequences of reverse optics, here exemplified by the albedo hypothesis, in a somewhat direct manner, thereby testing the validity of the research program itself. We also saw that, pace Chirimuuta, it is possible to compare two distinct research programs in a direct experimental way, thus showing that they are not experimentally isolated from one another.

In the present chapter, I would like to take a step back and consider the role color constancy theories can play in color ontology. I will argue that not only is there a strong link between the two but that an acceptable methodology for research on color ontology can start with unearthing the ontological claims motivated by color constancy theories. At certain points, I will draw on material of the two previous chapters to support the argument. This chapter can be seen as an extension of chapter 3 and an argument to the effect that we ought to pay attention to color constancy research if we want to do color ontology[114]. I will thus propose that color ontology could and perhaps ought to start with the study of color constancy theories since the latter embodies constraints on what the correct ontology of color could ultimately look like.

In the following, I defend the central thesis of the chapter, namely, that there is a high likelihood that by making progress on our understanding of color constancy, we are thereby making progress in color ontology. I thus propose that there is a much stronger evidential link between color constancy and color ontology than what may have been presupposed before. More precisely, I argue that color constancy theories usually motivate, if not a definite theory of color ontology, then at least a particular view on color ontology. In other words, a given theory of color constancy sets constraints on what the correct color ontology could look like,

---

[114] In appendix B, I provide another reason why philosophers thinking about color ontology ought to pay attention to color constancy research, a reason that to the best of my knowledge has not been pointed out before. More specifically, I argue that we make assumptions regarding the form which perceptual processing takes in the case of color vision in the course of making ontological arguments about color.

especially when it comes to whether the latter will turn out to be a realist, eliminativist, relationalist theory, or something else entirely[115].

In the following section, I consider three distinct color constancy theories, each of which embodies distinct constraints on what our ontology of color should look like. Note that I will not be focused on whether the constraints would yield an ontology that is phenomenologically plausible. I follow up by discussing how this novel approach to color ontology can be applied. I end the chapter with some objections that could be made to this approach and attempts to answer them.

The methodology in practice

It is now time to present what this methodology could come down to in practice. As the current work is rather exploratory, the examples I shall give are presented in broad brush strokes. The point of the following sections is more about giving an idea to the reader rather than convincing her that a particular ontology is definitely to be paired with a particular color constancy theory.

I will thus present three color constancy theories and try and argue that those theories, were they to be respectively correct, would constrain the form of the color ontology we may endorse. That is, the truth of any one of those three theories would have repercussions for the kind of color ontology which could be the correct one. Although the ontology might not be defined all the way down, it would still be limited in important ways, such as whether it is a form of eliminativism, relationalism or realism.

We already encountered two of those theories in the third chapter when we discussed Foster's and Zaidi's respective work. Although the focus was there put on whether those theories could accommodate some form of reflectance physicalism, we shall here pay attention to which ontologies of color they most readily embody. The third theory we shall discuss is Maloney and Wandell's. Although Hilbert has famously argued that it essentially justifies only his ontological theory, we will try and see whether the constraints it puts on color ontology research may not be broader than what has been assumed. I shall start with this theory.

---

[115] If this is true, then a proper understanding of constancy will also constrain which arguments are receivable or not. If the account of constancy we land on disqualifies, e.g., color realism, then any argument which apparently supports color realism will automatically look suspicious. That is, if the present reasoning is sound, then finding out the nature of color constancy will not only constrain which kinds of color ontology are acceptable but also shape the remaining debate in an important way.

Later on, we shall point out some general features of the approach and answer some objections. Note that the present chapter is thus essentially a continuation of chapter 3. Although there we were trying to determine which color constancy theory could represent evidence for reflectance physicalism, and thus starting with the color ontology, we are in both cases attempting to see the links between theories of color constancy and theories of color ontology. In the present chapter, we are taking a different route by starting with the color constancy theory and seeing where it leads us regarding color ontology.

The theory of linear models

Of course, there is already a famous appeal to color constancy theories in the literature on color ontology. Starting in 1987, Hilbert notoriously argued that the truth of Maloney and Wandell's theory of linear models justifies the claim that colors are SSRs and this was and arguably still is a central piece of his whole argument for reflectance physicalism (Byrne & Hilbert, 2003, 2020b; Hilbert, 1987). We will start by discussing the import of this theory of color constancy as far as color ontology is concerned. Remember that the theory of linear models is a proud example of the reverse optics approach put to work. That is, it conceives of vision (and of perception more generally) as the visual system's attempt at solving a series of inverse problems. In the case of color constancy, the visual system is trying to determine the causes of a given color signal.

More specifically, Maloney and Wandell's theory aims at the recovery of the SSR and SPD which jointly yielded a given color signal. According to the theory, the visual system decomposes the color signal into two weighted sets of basis functions. The first set corresponds to an estimate of SSR, whereas the second corresponds to the SPD estimate[116]. Each set is constituted by weighted basis functions which do not vary across scenes. In the case of SSR, basis functions represent basis reflectances. The curb describing a basis reflectance stipulates the amount of light reflected per wavelength for the whole visible spectrum. Thus, each basis reflectance represents a dimension along which particular SSRs can vary. Each SSR thus scores a particular number on each basis reflectance and, as the latter is fixed, knowledge of the weights pertaining to a given SSR represents complete knowledge of the SSR in question. The same analysis is performed in the case of illuminants' SPDs.

---

[116] Of course, in more complex scenes, such as realistic ones, the visual system will also have to determine what angle the objects sustain with regard to the viewer, whether haze is present, etc.… But the important point here is that, in all scenes, the algorithm will have to estimate the SPD of the illuminant and the SSR of the surfaces. For more on the notion of linear models, see appendix A.

When it comes to color ontology, Hilbert has appealed to Maloney and Wandell's theory in order to defend reflectance physicalism (or RP) (Byrne & Hilbert, 2003, 2020b; Hilbert, 1987, 2005, 2008). According to RP, colors are sets of reflectances. That is, each (determinable or determinate[117]) color we encounter in our color experiences, such as blue, red and green, is actually a set of SSRs. Pairing RP and Maloney and Wandell's theory thus makes sense. The latter explains how the visual system can come to know about the surfaces' SSRs it is presented with and RP is the further claim that those SSRs are the very colors we encounter in our color experiences.

According to that reconstruction, can we really say that Maloney and Wandell's theory *entails* RP? We can say quite confidently that RP is a natural addition to their theory but this by no means signifies that it is entailed by it. Chirimuuta nevertheless appears to say that it is:

> "Interestingly, the theoretical commitments of a scientist such as Maloney map smoothly onto a worked-out philosophical doctrine such as reflectance physicalism. That is, it can be said that the Marrian framework *implies* colour realism, by positing that, since vision *is* the recovery and representation of external world properties, there must *be* a physical property that we see as colour. And this is entirely consistent with Maloney's explicit statements about the objectivity of colour." (Chirimuuta, 2008, p. 579) (original italics)

Here she seems to clearly state that the Marrian framework, more specifically, Maloney's theoretical commitment, entails if not RP itself then at least a form of color realism. In order for this reading to go through, however, the representation of external world properties mentioned must be understood as a conscious representation, that is, as part of our color phenomenology[118].

Barring a solution to this issue, it is hard to understand how Maloney and Wandell's theory could be made to entail RP. From the fact that vision is "the recovery and representation of external world properties", it does not follow that such representations are part of our color experiences. This recovery and this representation could be entirely subpersonal. Note that Hilbert and Byrne especially take pains to emphasize that their theory does not operate at a subpersonal level:

> "The representational content of a subject's experience specifies the way the world appears *to the subject*. So the content of an experience is content at

---

[117] Determinable colors are general families of colors such as red, green, yellow and blue, whereas determinate colors are particular shades such as crimson, lime, canary yellow and azure.

[118] Here it seems that I am saying that phenomenology involves representation. However, note that the use of the concept of representation in psychology is not identical to the one involved in philosophy. The former is in all likelihood a much less loaded concept.

> the personal level – it is not subpersonal content. If the proposition that there are such-and-such edges, blobs, and bars is part of the content of an early stage of visual processing, it does not follow that that proposition is part of the content of the subject's visual experience." (Byrne & Hilbert, 2003, p. 5) (original italics)

Therefore, if we hold that Maloney and Wandell's algorithm works at the level of "edges, blobs and bars", then it cannot be concluded that its output will also be the content of our color experiences. Yet the authors give us no reason to conclude that their algorithm does not work only at that subpersonal level.

Indeed, Chirimuuta points out elsewhere that Maloney has actually distanced himself from RP:

> "Maloney (personal communication) has emphasized that an important difference between his own understanding of inverse problems and Hilbert's reflectance physicalism is that Maloney sees no requirement for explicit representation of recovered properties (reflectances), or of any of the estimates of illuminant chromaticities used to recover them." (Chirimuuta, 2008, p. 574)

Here it seems that we ought to read "explicit" as "conscious" and that Maloney is thus saying that the outputs of his and Wandell's algorithm may not be the colors we encounter in our color experiences. The algorithm may work only subpersonally, that is, without ever yielding a conscious percept[119]. But if Maloney holds that his theory does not necessarily entail the "explicit representation of recovered properties", then it cannot entail RP either if the colors are thought of as figuring in such representations.

The outputs of the algorithm may indeed be quite different than what is assumed by Hilbert:

> "Maloney (personal communication) notes that the goal of his work, strictly speaking, has not been to devise algorithms to reconstruct reflectance, but rather to assign invariant colour descriptors to surfaces, and these descriptors are correlated with SSR." (Chirimuuta, 2008, p. 566)

Colors are here construed as descriptors of surfaces but Maloney does not appear to endorse the claim that they are identical with SSRs, only that they are correlated with them. If there is theoretical space for the claim that something can be a descriptor of a mind-independent property or object without being that property or object, then indeed it seems that we have identified a reason why Maloney and Wandell's theory does not strictly speaking entail RP or

---

[119] This is not to say the outputs of the process are not used further down the line to determine our color phenomenology (whatever that may mean).

some other form of color realism. It is because more than one color ontology can be paired with RP and the latter cannot therefore entail only one.

If there is indeed theoretical space for that claim, then it seems we have identified another color ontology which is compatible with Maloney and Wandell's theory. According to this ontology, colors would be descriptors of surfaces without being mind-independent properties of those surfaces (such as reflectances). To grasp this point, consider road signs. The road sign "bear left" points towards some action that is required of the driver. However, the sign itself is distinct from the instruction it points towards. On the theory of color we are considering, colors would be identical with the road signs, not the action it points towards or the instruction associated with it. Or, to motivate the point further, consider labels. Colors could be like labels attached to objects (or surfaces) that would covary with some of the properties of those objects but also with certain properties of the scene (such as the presence of mutual illumination). The point being that labels are so to speak useful tools designed with some practical purpose in mind but not necessarily coextensive with mind-independent properties.

However, we need to be careful at this point. There are two ways to understand this notion of descriptor. One may conceive of descriptors as part of the outputs of subpersonal algorithms which would consist in the ascription of a color descriptor to a surface. Those outputs are not identical with the contents of our color phenomenology. Or descriptors can be understood as figuring in the contents of our color experiences. In that sense, color descriptors would *be* the blues, greens and reds of our color experiences. If some perceptual process yields a descriptor in the first sense, then it does not thereby mean that this descriptor will be a descriptor in the second sense or that it will entail the presence of a descriptor in the second sense. In other words, it is not necessary that a perceptual process yielding a descriptor in the first sense will be at least partly involved in explaining color phenomenology, although it may.

To sum up, one could thus devise an ontology in which colors are descriptors of surfaces which are correlated with SSRs but also with other things. This would allow for colors to change with factors unrelated with SSRs. For instance, a situation in which simultaneous contrast occurs is one in which the colors objects are judged to have vary with the background or neighboring objects. If colors are descriptors of surfaces which are only partly correlated with SSR but also with other variables of the scene, then simultaneous contrast does not constitute a counterexample to that claim.

One version of that view would have colors consist in mental, mind-dependent properties[120]. This theory would distinguish between a SSR and the color descriptor, the latter being part of the output of the perceptual process and not the SSR itself. Those descriptors could be construed either as parts of outputs of mere algorithms, that is, as subpersonal, or as parts of our color phenomenology. Depending on which interpretation we opt for and depending on the link we postulate between those subpersonal processes and our color phenomenology, this theory of colors as descriptors may or may not explain this phenomenology.

This mentalist ontology would probably count as a form of eliminativism if the latter consists in the claim that ordinary objects like raspberries, balloons and leaves are not colored. In that case, the visual system would really create a fiction when it attributes color descriptors to objects. It may be a useful fiction (to categorize objects and remember them more easily later on) but a fiction nonetheless. Alternatively, the ontology may also be construed as a form of relationalism where colors are properties of the outputs of perceptual processes which are determined by a variety of factors such as the SSR of surfaces, the presence of simultaneous contrast, illumination, etc.[121]. Here the visual system would detect regularities in the environment and attribute color descriptors to them. Those regularities would consist in covarying properties. For instance, green would be attributed to particular objects (e.g., leaves) under a particular illuminant (e.g., dappled sunlight) with other variables being present in varying amounts (e.g., a particular amount of interreflections from the other leaves).

However, there is a feature of the theory of linear models which seems to speak against this relational reading of the descriptor theory. Note that, if all goes well, the estimate of SSR is supposed to be independent of the estimate of the illuminant's SPD. In situations approximating natural environments, it will also be independent from estimates of other confounding variables, such as the detection of haze, mutual illumination, etc.. But then the estimate of SSR does not really appear to be a relational property. It may be a mental one but it is not determined by anything else than the SSR of the surfaces at issue.

However, the very fact that the color descriptor is determined by other variables of the scene beyond SSRs was one of the reasons we postulated that the descriptor theory was a viable alternative to reflectance physicalism. If we then back-pedal and state that, actually, the

---

[120] Another version we will not concern ourselves with has colors consist in relational properties of objects. We will see a similar theory when we consider Foster's theory of constancy.

[121] Note that this is reminiscent of Chirimuuta's adverbialism, according to which colors are properties of perceptual processes constituted by objects, observers and viewing conditions (Chirimuuta, 2015).

descriptor is correlated only with the SSR of a surface, then this puts pressure on the idea that we have identified a theory which is distinct from RP. If colors are descriptors correlated uniquely with SSRs, then it is not clear that the theory of descriptors is distinct from RP. The only difference between the two would appear to be that the former does not identify the color with the SSR. The descriptor theory has the colors identified with estimates of SSRs but does not follow up with an identification of the colors with the SSRs. It is not clear that this is a viable alternative to RP in the present context, even though it is possible.

Nevertheless, we are now in a better position to answer Chirimuuta's remaining worry to the effect that reverse optics, and more specifically Maloney and Wandell's theory, presuppose some form of reflectance physicalism. First, it is not at all clear that it does. We may have identified a distinct theory which is very distant from any sort of color realism. It seemingly cannot be true that we need to endorse RP in order to believe that the evidence speaks in favor of Maloney and Wandell's theory. Secondly, even if it is true that their theory only represents justification for RP, and if it is therefore wrong that the descriptor theory we identified is a viable option, then it is not clear that this is an issue.

This is where the resources developed in chapter 4 come in handy. We saw there that color constancy theories, such as Maloney and Wandell's or, more generally, frameworks such as reverse optics, have empirical consequences which can be tested. The fact that they have certain consequences is accepted by all theorists alike, whether or not they endorse reverse optics or any other framework. It is therefore possible to empirically falsify a theory of color constancy or the framework which it is an operationalization of without first endorsing any color ontology. This means that there is no risk that we will enter a methodologically vicious circle.

The original theory of cone-excitation ratios

We can now turn to a second theory of color constancy, namely, Foster's theory of cone-excitation ratios (or CER). In chapter 3, we saw how Foster's theory can be interpreted as requiring what Gilchrist calls an *anchoring rule*. Supplemented with the latter, I tried and argued that the empirical success of this modified theory would constitute evidence for reflectance physicalism. As pointed out above, the aim of the present chapter is to examine which ontological consequences could be motivated by distinct color constancy theories, not whether reflectance physicalism itself can be shown to be motivated by certain (modified) theories. In order to do so, we will now consider Foster's unadulterated theory.

Remember that Foster argues that asymmetric matching experiments actually measure a different ability than color constancy as traditionally understood. Rather than relying on the latter, asymmetric matching relies on our ability to perceive constant chromatic relations across scene variation (including, but not limited to, illumination variation). Color constant behavior is supposed to be achieved not through the ability to recover estimates of SSR to objects independently of factors such as the illumination but through the fact that our perception preserves the color relations of the surfaces in the field of view across scene changes such as illuminant changes.

In order to more precisely capture the difference between traditional approaches to color constancy (such as Maloney and Wandell's) and Foster's, let us consider color spaces. Color spaces are color order systems. That is, they order colors according to certain variables. For instance, one of the most famous color order systems still in use is the Munsell Color System. According to the latter, colors are ordered on the basis of three variables: their hue, their value (similar to lightness) and their chroma (related to the notion of saturation). Each color will score a particular number on each of the three variables, situating it in the space. The positions of the colors are usually represented on an oval shape (similar to a rugby ball), with the center axis representing lightness, horizontal distance from the axis representing chroma (saturation) and each direction from the axis representing a different hue. All achromatic hues, running from white to black, are on the center axis, whereas, e.g., a particularly dark purple stands on the periphery of the color solid, on the particular direction representing that hue of purple.

More importantly for our purposes, colors are ranked on the Munsell system (and most if not all other color systems) according to how they rank on each of the three variables. That is, their ranking reflects their absolute position on the three variables and not their relations with regard to the other colors[122]. This is arguably how we conceive of traditional color constancy. Although the variables encoding the colors in our visual system are without a doubt of a different nature than those of the Munsell Color System, we conceive of this encoding as a particular value in a multidimensional space. That value ought not to depend on the particular illuminant at issue, nor on any other variable which changes across scenes (such

---

[122] Of course, ranking in a particular way on the three variables will entail particular relations with regard to the other colors. But this is not how colors are ranked on the order system.

as the presence of haze). Crucially, a particular color being encoded in a particular way does not depend on its spatial relations to other colors[123].

This is where Foster departs from the traditional understanding of color constancy. He proposes that the encoding necessary to perform (at least) asymmetric matching be realized in a space in which surfaces are situated according to their relations with other surfaces. The space in question is supposed to be a space defined by the excitations of within-cone-channel excitations ratios[124]. That is to say, rather than there being an absolute scale on each of the axis, those axes encode ratios of excitations. For instance, being on a certain point on the axis representing the excitations ratios corresponding to the L cones channel would not mean that the surface excites the L cones to a certain degree (i.e., on an absolute scale), but would indicate that it excites the L cones to a greater or lesser extent than another surface. The two surfaces could be positioned on the L dimension in such a way that the first surface is encoded as exciting the L cones to an extent which is twice the amount of L cones excitations corresponding to the second surface. If all the surfaces of the scene viewed were classified on the space, we would obtain a map of the differences of excitations pertaining to each pair of surfaces for any cone channel. But this map would not indicate whether a given surface scores high or low on any of the dimension, only whether it scores higher or lower than any other surface by a given proportion.

Note that the point of this exposition is not to say that there is a conflict between the Munsell space and a space based on cone-excitation ratios. Nor is it to say that the Munsell color space is deficient in some regard or other or not fit for some purpose. It is rather to point out the way colors could be conceived of if one endorsed Foster's theory of constancy. The difference between the Munsell space and one based on cone-excitation ratios which we thus need to focus on is that in the latter case but not in the former the position of one point in the Munsell space is determined by its relations with other points, not by its absolute positions on each of the axes[125]. With that understanding in mind, we can turn to the consequences of the theory of cone-excitation ratios for color ontology.

---

[123] Note that some may argue that this encoding is actually relative encoding in disguise. That is, it may not be possible to specify, e.g., the hue, saturation and lightness of a surface independently of the hue, saturation and lightness of other surfaces. However, as far as lightness is concerned, this does not appear to be true. Encoding a surface according to which portion of the incident light it reflects does not appear to imply a comparison with other surfaces.

[124] Or any space which depends in one way or another on those excitations. A space of cone-opponent excitations would fit the bill. Note, in addition, that the values in such a space would not depend on factors such as illumination or the presence of haze, similarly to the Munsell color space.

[125] Of course, there are many other similarities and differences between the two spaces but this is the one we need to focus on for our purposes.

Following our discussion of Foster's theory in the third chapter, we may first note that this relative encoding of within-cone-excitation ratios does not necessarily imply that Foster's theory cannot be paired with some form of color objectivism like reflectance physicalism. As we saw, it is not clear that the theory as presented by Foster can account for all our color constancy-related abilities. Neither is it clear that it can account for our color phenomenology, more specifically, for our phenomenology of color constancy[126]. Nonetheless, those issues can be remedied once an anchoring rule is provided. Once the latter is done, the map of ratios we have presented can be fixated on a non-relative map of cone excitations which is going to strongly correlate with the map of SSRs of the surfaces viewed[127]. SSRs can be measured with Foster's proposed algorithm once an anchoring rule is added to it.

This is, however, only one way ontological questions about color could be informed by the theory of CER. If we reject the thesis that an anchoring rule must be added to the theory, there is another way the latter can inform us about ontology[128]. Consider again this quote from Foster we touched on in the third chapter:

> "For example, in Fig. 3, the scene is illuminated by different daylights, with correlated color temperatures (a) 17,000 K, (b) 4000 K, (c) 6500 K, and (d) 4000 K. The color of the light reflected from the sphere in the bottom left corner in a–c is clearly different. Nevertheless, […] it can be seen that the sphere has the same or similar surface color in each image by comparing it with the nearby foliage and by looking over each image as a whole." (Foster, 2011, p. 680)

In this passage, Foster does not appear to deny that objects have colors. In particular, he talks of objects having surface colors[129]. He explicitly talks about the surface color of a sphere and of it being "the same or similar" across scenes with distinct illuminants. Although it is not directly addressed by Foster in his writings, the theory of CER may motivate a form of color objectivism, at least under certain conditions.

---

[126] Note that an advocate of RCC may argue that the phenomenology of color constancy is not part of the explanandum of RCC. In that case, not being able to explain it would not represent a counter-argument to the theory. Alternatively, the advocate of RCC could claim that phenomenology is actually neutral with regard to whether colors are presented as relational or not in our color phenomenology (for a similar strategy, see Chirimuuta, 2015).

[127] This non-relative map of cone excitations could involve, for instance, cone excitations from objects under an achromatic illuminant. Those cone excitations would strictly correlate with a map of metameric sets of SSRs. For the impact the choice of this non-relative map would have on ontology, see the discussion of Zaidi's theory below.

[128] This would be required by Occam's Razor if we believed that Foster's theory in its unadulterated state was able to account for all relevant color phenomena.

[129] For more on the notion of surface color and its distinction with perceptually unasserted color, please see the next section.

Note that what is lacking from the above explanation of relational color constancy if it is to motivate a form of objectivism is that a visual system employing the corresponding algorithm would only generate a map of excitations *ratios*. This is where, as we saw, the anchoring rule came in handy by helping fixate this map of ratios on a non-relative map. But there is another way the theory could explain the phenomenological fact that we do not appear to see mere color relations[130] and, at the same time, provide grounds for another kind of color ontology.

For ease of exposition, let us for the time being restrict our attention to pairs of scenes which share the same sum of color relations. In such scenes, for any color relation such as e.g., an object emitting twice the amount of longer wavelengths than another object, if that relation obtains in one of the scenes, it obtains in the other. The paragon of such scenes is obviously the same scene under distinct ambient illuminations. If it used Foster's proposed strategy, a visual system exposed to such scenes would be able to determine the sum of chromatic relations a surface enters into under some illuminant and use that information to re-identify that very surface under another illuminant. That is to say, across those scenes, the sum of chromatic relations of an object is constant. If this is the case, then it becomes quite easy to identify this sum of color relations with the color of the object.

According to this line of thinking, an object being a certain color is equivalent to its entering into certain chromatic relations with other surfaces. That is to say, being a certain color is identified with a relational property of the object. However, the latter does not consist in a relation obtaining between an object and a subject but between an object and other objects. It is thus a form of relationalism of color which is quite distinct from the traditional relationalism found in the literature on color ontology (see e.g. Averill, 1992; Cohen, 2004, 2009; Thompson, 1995), where one of the factors determining color is the subject herself. According to the relationalism we are considering here, colors are relational properties which do not require a subject. Let us call that theory *objectual color relationalism* (or OCR, for short). Perhaps more interestingly, OCR can also count as a form of color realism.

Consider that, although this has been disputed (see, e.g., Chirimuuta, 2015), one of the hallmarks of color realism is the claim that colors are deemed to be mind-independent properties. A second hallmark, of course, is the thesis that colors are non-relational properties. Nevertheless, it may be important to point out two things. First, it is entirely possible that the non-relationality clause of traditional color realism is at least in part a way to avoid making

---

[130] If we take for granted such a fact, at least for the time being.

reference to minds, perception or experience in the definition of color. That is, it is possible that it is another, additional way to ensure mind-independence. Second, although it may be considered fair to count a theory as a form of color realism only if that theory fulfills the non-relationality condition, note that it is also question-begging against theories which have the theoretical resources to count as a form of color realism but imply that colors are indeed relational[131].

However, some may complain that OCR secretly refers to observers in its definition of color. Since chromatic relations will be encoded by cone excitation ratios and those ratios obviously imply reference to an observer, those relations actually involve reference to an observer and are thus not mind-independent relations. OCR, according to the argument, could not count as a form of color realism. We already tackled this issue in chapter 3 when we discussed objections to construing Foster's theory as a potential justification for reflectance physicalism. In order to report facts of the outside world, a perceptual system has to encode information in one way or another. Arguing that such encoding represents a bar to that perceptual system representing facts is indeed asking for perceptual systems to perform an impossible feat.

Nonetheless, this complaint prompts another, more interesting question: what is, indeed, the information encoded by cone excitation ratios? In other words, it encourages us to determine what sets of relational properties the colors are supposed to consist in according to OCR. As we just saw, the answer will not be that the information encoded are facts concerning the cones excitations. Those excitation ratios will track differences between the preferences of objects for reflecting certain wavelengths over others. For instance, if an object O reflects ten times the set of wavelengths captured by the activity of the L cones than another object P does, then part of the information encoded by the visual system regarding O will be that O possesses the property of *reflecting ten times the amount of longer wavelengths reflected by P*. Obviously, it will also encode the corresponding differences for the shorter and middle wavelengths. O may be encoded as *reflecting half the amount of shorter wavelengths reflected by P* and *reflecting twice the amount of middle wavelengths reflected by P*. In addition, the visual system will encode O's preferences relative to the other objects in the scene.

Another possibility is that a visual system displaying relational color constancy will, according to Foster's theory, encode O's preferences for certain wavelengths relative to the

---

[131] Although the reasons for her saying so are slightly distinct from the reasons presented here, Chirimuuta has also claimed that relationalism is "realism enough"(Chirimuuta, 2015, p. 118).

scene chromatic average (see, e.g., Amano & Foster, 2004). That is, the visual system will first determine the chromatic average of the scene by calculating the mean amount of light reflected according to its wavelength. It will then situate O in relation to that average. O might then be judged as *reflecting 0.7 times the amount of shorter wavelengths the chromatic average would reflect*, *reflecting 1.4 times the amount of middle wavelengths the chromatic average would reflect* and *reflecting 2.2 times the amount of longer wavelengths the chromatic average would reflect*[132]. In a scene with a different chromatic average, the visual system will identify an object as similar or identical in color with O if and only if this object is similarly related to the new scene chromatic average.

We are now in a position to consider how the information encoded by the visual system differs depending on whether objectual color relationalism or some more common form of color realism, such as reflectance physicalism, is considered. According to reflectance physicalism, judging an object's color boils down to judging to which class of SSRs the object belongs to. According to OCR, it comes down to determining which types of wavelengths it reflects relative to the other objects in the scene. That is, reflectance physicalism has color consist in a sum of proportions of light reflected relative to the incident light, whereas OCR has it consist in a sum of proportions of light reflected relative to the same kind of light reflected by neighboring objects. So reflectance physicalism has color consist in a ratio of reflected light to incident light, while OCR has it consist in a ratio of reflected light to reflected light (of other objects). Here we can see that OCR really qualifies as a form of color realism. Even though the properties which are identified with the colors are relational properties, the latter are properties which refer to other objects beyond the object itself. There is no reason to believe that those properties could not be instantiated in a world devoid of observers.

If we consider the version of OCR which states that the attribution of colors entails reference to the chromatic average of the scene instead, the above conclusion still follows. Rather than referring to other objects' propensity to reflect the same kind of light, this version of OCR would state that the color of an object is its preferences for reflecting certain types of wavelengths relative to the scene chromatic average's preferences for reflecting those same

---

[132] Note that it could be argued that it is possible to design an experiment to distinguish between attributions referring to other objects' color and attributions referring to the chromatic average. In an asymmetric matching paradigm, participants could be set the task of matching a center patch in a Mondrian scene constituted by different patches but with the same chromatic average as the standard test. If the participants' performance is not affected by the change in neighboring patches, then it could be argued that the attributions involved in Foster's theory refer not to particular colored patches but to the chromatic average of the scene. See Brown & MacLeod, 1997 for a similar experiment which seems to motivate the claim that the distribution of colors around the average is also a factor to take into consideration.

types of wavelengths[133]. Here the relational property which OCR identifies with the color of the object is a relational property involving not a pair of objects but an object and a statistical or mathematical construct.

However, note that we have been restricting our discussion to pairs of scenes which share the same set of color relations for ease of exposition. There may be important differences to our conclusions if we now consider scenes which do not. For instance, let us focus on pairs of scenes where the difference between the two consists in a change of the surfaces' SSRs. As the SSRs change, the color relations which can be seen to obtain in the scene will change. An object that has in the first scene the relational property of *reflecting ten times the amount of longer wavelengths reflected by P*, where P is another object with a particular SSR, will not be seen as possessing that property in a scene which does not include P.

Let us call the sum of the color relations a particular object O is seen as entering into in the first scene as the sum S and the sum of the color relations O is seen as entering into in the second scene as the sum T. S and T are themselves subsets of the whole set of color relations O can enter into when we consider all possible scenes. A question here arises: if S and T are insufficiently similar that the visual system cannot come to the conclusion that they are subsets of the same set, how will it attribute colors to O in the two scenes? If the visual system's attribution of a color to O just is its attribution of the sum of color relations S to O in the first scene and its attribution of T in the second scene, then one cannot help but conclude that the colors we see objects as having change as scene change in important ways.

One way to construe the resulting ontology would be to say that, when scene changes are so dramatic that a completely different set of color relations is attributed to an object following such a change in the scene, the resulting color attribution is actually a mistake on the part of the visual system. This is equivalent to holding that the visual system tells us about the true colors of objects only under certain conditions. When those conditions do not obtain, it is unable to do so and we should see the colors attributed as systematically wrong. An important task of this theory would be to tell us under which conditions are true colors experienced and to do so in a way which seems principled enough. As the literature on standard conditions and standard perceivers has told us (see e.g. Averill, 1992, 2005; Cohen, 2004, 2009; Hardin, 1988, 2014), this may be easier said than done.

---

[133] Of course, the scene chromatic average does not reflect wavelengths per se, as it is an abstraction. This is just a simpler way of saying that the relational property is between an object and another abstract object reflecting the same color signal as the scene chromatic average consists of.

Perhaps a way out of this conundrum is to consider what would happen if the visual system actually had one of the relata of the color relations be the chromatic average of the scene rather than a given object. If one construes color properties as relational properties between an object's preferences for reflecting certain wavelengths compared to the scene chromatic average, then objects may be conceived as possessing only one color. As long as the object can stand in the same color relations with regard to the chromatic average as it can in any other scene, any surface will always occupy the same location with respect to the average. In other words, it will only have one color (the property consisting in its color relations with the average) and the latter will not change[134]. In that case, the visual system will be seen as trying to recover that color by using the color relations it has access to. If those are sufficiently representative of the relational property the object has with regard to the global chromatic average[135], then the visual system will manage to estimate the latter.

To sum up, any variety of OCR will have the following characteristics:

a) Colors are sums of relational properties.
b) The relations which determine those relational properties are relations between an object and other objects or between an object and a chromatic average.
c) Those relations consist in differences of preferences for reflecting certain (types of) wavelengths.

One might want to add that OCR entails that colors are mind-independent properties. A good case can be made for this claim. If objects have preferences for reflecting certain types of wavelengths (which we know they do), then this entails that there are differences between objects regarding those preferences. But then each object or surface will automatically inherit a set of relational properties regarding those differences. According to OCR, this set of properties is the color of the object. Nowhere in that argument is an appeal to observer, minds or perception required[136]. We can thus add the following feature of OCR:

d) Colors are mind-independent properties.

However, it could be argued that neither a), b), c) or d) nor a combination of some of these would seem to imply that objects can have at most one color. OCR is compatible with the claim that objects have numerous colors. But this may be too hasty. If the "other objects" in b) include all possible objects, then objects can have at most one color. The visual system would

---

[134] In scenes in which the color relations the surface has with respect to the average do not reflect accurately the relations it would have in the average scene, it could be said that the visual system is here unable to correctly estimate the color of the object.

[135] The global chromatic average need not be the chromatic average of all possible scenes. It may refer to the chromatic average of the natural scenes found in our actual environment.

[136] Note that the same argument can be run in the case of the color involving a relation to the chromatic average.

then be seen as trying to determine the overarching set of relational properties of an object (that is, its color) by relying on, hopefully, a representative subset of this overarching set. We can thus add a final feature to our list:

e) For each object O, if O has an SSR, O has one and only one color.

According to OCR, then, the red of a tomato is actually not a non-relational property of the tomato, as traditional color realists would have it. On the contrary, this red is equivalent to the tomato's preferences for reflecting certain wavelengths as compared with the preferences of other objects (or with the chromatic average of the scene) for reflecting those same wavelengths. That is, it is identical with a relational property of the tomato[137].

The sensed and concluded color of Zaidi

We can now turn to a last color constancy theory, namely, Zaidi's. Remember that the core of Zaidi's theory consists in a set of algorithms which would, were they to be implemented, allow the visual system to identify similar objects across illuminants. More precisely, those algorithms allow the visual system to conclude that a particular object seen under an illuminant is of the same material or has at least the same surface as another object seen under another illuminant. But, most importantly, those algorithms are silent on the nature of the surface properties of the objects in question. That is, although the photoreceptors excitations corresponding to the color signal from the objects are used in the algorithms, a visual system implementing them would not go on to separate those excitations into the aspects corresponding to the SSR of the objects and those corresponding to the SPD of the illuminant. This is consistent with Zaidi's rejection of reverse optics models of color constancy.

In other words, Zaidi's core theory merely explains our ability to identify objects across illuminants, not our color experiences or the apparent stability of the colors we think we see when we have those experiences. That is, his theory explains what we may call *corresponding color constancy*. In that sense, the possession of corresponding color constancy does not strictly entail the possession of the ability to attribute colors to objects, at least not in a non-relative sense (such as "this is blue")[138]. Neither does it involve the ability we discussed

---

[137] I here avoid saying "the set of relational properties" since "relational property" can be understood as already involving a set of relations.

[138] It may involve the ability to draw conclusions regarding whether objects have the same or distinct color without telling you anything about those colors. Conclusions such as "those objects have the same color" without a further specification of which color it is would fit the bill.

when we touched on Foster's theory, namely, the ability to attribute color relations to pairs of objects (such as "this is x times bluer than this"). As Zaidi puts it himself:

> "The present paper, however, suggests that the problem for the visual system to solve is not to bring about stable color appearance under different illuminants by discounting them but to recognize that objects are indeed being viewed under different illuminants and to discover what the illuminant properties are." (Zaidi, 1998, p. 1767)

Here we can see that Zaidi's approach is not concerned with a supposed phenomenological constancy of color but with a performance-oriented view of the phenomenon, understood in a particular sense. If there are any to be found, we have to look for ontological consequences elsewhere. Since I believe that the rest of Zaidi's theory concerning essential colors is not fully developed yet, the following is more of an exploration of what it could be rather than what it definitely consists in.

Remember, first, that Zaidi appears to side with Lichtenberg in holding that there needs to be a distinction made between the sensed color and the concluded color. One can refer to the color inconstancy of a tomato placed under skylight and then under a lightbulb by pointing out that the tomato appears to have a different color in both situations. When one uses the term "color" in that sense, one is actually referring to what Zaidi calls *the sensed color* of the tomato. As the illumination falling on the tomato changes in SPD, so does the sensed color.

But one can talk of the tomato as having the same color under both illuminants. In that sense, the term "color" refers to what Zaidi calls *the concluded color*[139]. The concluded color of the tomato remains constant across illumination changes. Having a particular concluded color is equivalent to displaying a certain set of sensed colors under the appropriate viewing conditions. Two objects therefore have the same concluded color if and only if, for any illuminant, they display the same sensed color under that illuminant[140]. According to Zaidi, the concluded color is not part of our phenomenology. It is a cognitive achievement based on

---

[139] This is consistent with Chirimuuta's emphasis on the fact that Zaidi and other similarly minded theorists, contrarily to Maloney and his team, actually embrace the distinction between sensation and perception, where sensation is an immediate response to, here, the color signal, whereas perception is the result of subpersonal inferential processes (Chirimuuta, 2008, p. 577). According to this picture, finding out the sensed color of an object is a matter of having a sensory response to the objects, whereas determining its concluded color is done by inferential non-sensory processes.

[140] It is interesting to point out that using this criterion to determine whether two objects have the same (concluded) color implies that metamerism is not a sufficient condition to determine whether they do. Metameric objects are indistinguishable under a given illuminant but they may differ in sensed color, and thus be distinguishable, under another. According to the present criterion, those objects would not have the same (concluded) color. Note furthermore that this criterion is not in conflict with the thesis that colors are visibilia, i.e., the thesis that we can determine the color of an object only by looking at it. If we use that criterion, we just have to look a lot more than we thought we did.

a processing of the sensed color[141]. In other words, we do not have a visual phenomenology of color constancy, but rather of color inconstancy[142] [143]. Our phenomenology is inconstant because only the sensed colors are part of it. As the sensed colors change with viewing conditions, our color phenomenology is inconstant[144].

Note that this distinction between sensed and concluded color is reminiscent of Cohen's way of understanding color constancy (Cohen, 2004, p. 459). Considering a red cup half in shadow, Cohen argues that the two distinctly illuminated regions of the cups have distinct colors. At the same time, both regions also have a color that they are not manifesting right now, namely, the color of the other region. For instance, the illuminated region of the cup will have the property of *being light red under the current illuminant*. But it will also have another property, that is, the property of *being dark red under the illuminant I*, where the illuminant I is the illuminant found in the shaded region of the cup. According to Cohen, color constancy consists in the ability to determine which color a particular surface would have under another illuminant. One difference between Cohen's and Zaidi's conceptions is that the former reserves the name of color to what Zaidi calls the sensed color[145]. However, both authors would probably agree that the concluded color in the case of Zaidi or, alternatively, the judgments regarding which colors would be instantiated under other illuminants in the case of Cohen are firmly located on the post-perceptual side of things.

Going back, Zaidi links the notion of the concluded color with that of the essential color. The latter refers to the cone excitations corresponding to an object under an equal-energy illuminant. One may understand the essential color as a criterion for determining *which* concluded color an object has whereas the possession of the same concluded color is a

---

[141] Note that the concept of concluded color is explained by the concept of sensed color.

[142] The distinction between sensed and concluded color is reminiscent of another distinction we touched on in chapter 3, namely, the distinction between perceptually unasserted color and surface color (Arend et al., 1991; Arend & Reeves, 1986). The latter is somehow equivalent to the notion of concluded color as it is supposed to be invariant under illuminant changes, whereas the perceptually unasserted color (or "unattributed") roughly corresponds to the sensed color in the sense that it is supposed to change with illuminant changes. However, note that perceptually unasserted color and surface color are both supposed to be part of the content of perceptual judgments. Although the concluded color is also conceived of as the result of a perceptual judgment (more specifically, an inferential process), the sensed color is deemed a sensory response in that it is thought an immediate, non-inferential response to stimuli.

[143] Importantly, since Zaidi denies that there is color constancy on a (visual) phenomenological level, his theory does not have to explain it. Of course, he may have to provide in due time a reason as to why certain people believe that there is such a phenomenology of constancy. One such reason may for instance be that people mistakenly view the constancy of the concluded color as an element that is present at a phenomenological level, whereas it is actually only something they judge.

[144] Note that we are talking about visual color phenomenology. If there is such a thing as a cognitive phenomenology of color, color constancy may well be a part of it.

[145] Of course, there are other differences at play here. For instance, as we shall see below, sensed colors are conceived by Zaidi as a creation of the visual system, a fiction that has no counterpart in the mind-independent world. Cohen believes that colors are real albeit relational properties.

criterion for determining *that* they have the same essential color. In other words, to determine which color an object has (in the sense of a stable, unchanging color), we have to put that object under an equal-energy illuminant and see how it appears. That is to say, we have to determine the sensed color of that object under this illuminant. To determine whether another object has the same color as the first object, we have to see whether they display the same color (in the sense of a changing, unstable color), i.e., whether they appear the same colorwise, under all illuminants[146]. Under that construal, an equal-energy illuminant reveals which concluded color an object really has. A surface may appear for example blue under a certain illuminant, that is, it may display a blue sensed color under that illuminant but this surface's concluded color is blue only if the illuminant in question is equal energy[147].

An equal-energy illuminant is here construed as a special illuminant. That is, it is thought to be authoritative when it comes to determining the concluded color of an object. This is obviously reminiscent of the much discussed notion of standard conditions (Averill, 1992, 2005; Boghossian & Velleman, 1991; Hardin, 1988; Johnston, 1992; Levin, 2000). In the research on color ontology, it is usually pointed out that an object may appear of various colors depending on the type of illuminant it falls under, the neighboring objects' colors, the observer doing the viewing, etc.[148]. If one wants to preserve the thesis that objects have at most one color, then it needs to be determined under which viewing conditions is the object supposed to appear with its 'true' color.

One common solution is to argue that certain viewing conditions are special in that they reveal this true color. Those conditions are dubbed 'standard'[149]. Although it has been pointed out that there are numerous issues with this notion, it is here mentioned only to draw a parallel between it and Zaidi's approach to color constancy and color ontology. One issue repeatedly mentioned is that there can be no non-arbitrary criterion to identify standard conditions. For instance, conditions in laboratory settings are chosen based on the kinds of tasks subjects are asked to do. The right conditions of viewing for makeups may just be the conditions under which the makeup at issue is more likely to be seen. It is thus interesting to

---

[146] What if two objects appear to have the same sensed color under an equal-energy illuminant but not under other illuminants? According to the notion of essential color, they would have the same essential color but the fact that they have distinct sensed colors under at least another illuminant seems to point in the direction that they do not. There thus appears to be a tension if we analyze the notion of concluded color in terms of essential color. To the best of my knowledge, Zaidi and his team have not explicitly addressed this issue in their writings.

[147] Note that "blue" and other color terms can refer either to the sensed or the concluded color under that construal. An object may appear blue under a certain illuminant (sensed color) but not be blue, i.e. appear blue under an equal-energy illuminant (concluded color). Color terms are according to this understanding always ambiguous.

[148] For more on the notion of standard conditions, see appendix B.

[149] And the observers which are supposed to be authoritative are called "standard" or "normal" observers.

note that Zaidi does have an answer to this question. Under an equal-energy illuminant, the cone quantum catches from an object for an observer are identical to the scalar products of the function describing the SSR of the object and the spectral sensitivity functions of that observer's photoreceptors. In other words, an equal-energy illuminant does not distract from the SSR of an object[150].

When we consider color ontology within the context of Zaidi's approach to color constancy, we have to keep in mind the distinction between sensed and concluded color. Our ontological account of sensed color may be quite different from the corresponding account of concluded color. Sensed colors are supposed to be transient properties whereas concluded colors are stable properties.

Let us then first consider the ontology of sensed color[151]. As we saw, sensed colors depend on a number of variables. Once the illuminant's SPD is changed, so will the sensed colors associated with the scene. Interreflections added to the scene will automatically impact sensed colors. As a final example, consider that color contrast cannot be changed without influencing the sensed color associated with the target surface. However, it is unclear whether we ought to consider those factors as merely causally relevant to sensed colors or whether they are constitutive of sensed colors. For instance, are sensed colors merely causally determined by the illuminant or are they constituted in part by the illuminant? According to the latter, a sensed color subjectively indistinguishable from another one would not be identical with it if they are constituted by illuminants with distinct SPDs. On the former view, sensed colors are non-relational mental entities or properties and it may be possible (or even necessary) for two indistinguishable sensed colors to be identical.

In addition, although it may or may not include the claim that sensed colors are relational properties, this ontological account will not include dispositionalist elements. Sensed colors are occurrent properties. If I see an object as having a red sensed color, I am seeing this object as appearing red right now. That is, the sensed color red is a property of the object only as long as it stays under the relevant illuminant. If I were to say that this object does not appear red right now but has the potential to appear red under other circumstances, I would merely be talking about concluded colors.

---

[150] Note that this strategy does not deal with interpersonal variation such as variation in the cones' sensitivity. Furthermore, accepting this construal of standard conditions would essentially entail that colors almost never appear as they really are (i.e., of the sensed color which gives them their name, namely, their essential color), since equal-energy illuminant is essentially only found under laboratory conditions.

[151] That is, at least as far as we can go with the schematics we have drawn of Zaidi's theory. None of this is supposed to be a definitive analysis of the ontology accompanying his theory.

Note that we have not yet addressed the question of whether sensed colors are properties of objects or mere projected properties. According to the way we construe them here, projected properties are subjective properties which are not seen as such. In fact, experiences represent mind-independent objects as having those properties. They are "projected" onto material objects and this projection is thought to be a mistake[152]. Zaidi clearly seems to want to endorse some sort of projectivist account when he says that "by selectively absorbing wavelength bands of electro-magnetic radiation in specialized photoreceptors and performing neural operations on the resulting signals, the human brain *creates* the percept we call color" (Zaidi, 2000, p. 192)[153](my italics). If one decides that a projectivist ontology is the one best fitted to Zaidi's (completed) theory of sensed colors[154], then they would have to argue that sensed colors are, in some sense or other, properties figuring in the content of color experiences. For instance, they could argue that they are properties mistakenly represented by color experiences as belonging to material objects. This theory would then count as a type of error theory.

When it comes to our account of concluded color, the first important point we have to mention is that it will require an understanding of sensed colors. Concluded colors are indeed defined by sensed colors. As Lichtenberg puts it, and Zaidi seems to concur (Zaidi, 1998, 2000):

> "People certainly know what color it is that they call white, but how many have ever been confronted with a pure, white colour? In ordinary life we call white, not what looks white, but what would look white if it was set out in pure sunlight, or in a light whose quality did not differ much from sunlight. It is more the potential to be white and become white, in all its gradations, that we call white in some object, rather that the pure white colour itself." (Joost et al., 2002, p. 302)[155]

In other words, if we want to talk about a stable property when we talk about the dark blue of the blueberry, we have to hold that being dark blue is actually the dispositional property of the

---

[152] For a projectivist theory defended by philosophers see e.g. Boghossian & Velleman, 1989.

[153] However, note that it is a risky project to just take such quotes at face value. For one, color scientists sometimes write them as a preamble (such as in the case of Zaidi), without really putting much stock into them. Secondly, it is not clear that they express their real opinion about ontology with those claims.

[154] Alternatively, one might say that an eliminativist theory fits Zaidi's picture better. Of course, this issue will involve determining whether projected properties need to be properties which *are* instantiated by other things (such as the visual field) but are then wrongly attributed to objects. For a theory which does require this instantiation, see Boghossian & Velleman, 1989. If they need not be instantiated anywhere, then some eliminativist theories will count as projectivist as long as they also endorse the claim that our color experiences represent objects as colored. That is to say, they will count as projectivist if they endorse an error theory of color experience.

[155] Note that one issue with this quote is that sunlight is not strictly speaking an equal-energy illuminant.

blueberry to appear dark blue under certain conditions. Being dark blue, as a concluded color, entails looking dark blue (i.e., having a dark blue sensed color) under those conditions[156].

A consequence of this is that the blueberry does not need to look dark blue in any other conditions. In the proper circumstances, it might look black, more reddish or more greenish. That is, the sensed color associated with it may be dramatically different than dark blue if one takes pains to change the viewing conditions enough. If we take Lichtenberg's quote at face value, we might even say that, even if we can conclude, based on our perceptual experiences, that blueberries are dark blue, it is entirely possible that we have never seen it as dark blue, that is to say, as displaying a dark blue sensed color[157].

Our account of concluded colors will thus be a dispositional one and concluded colors will be analyzed as the dispositions of objects to cause particular sensed colors under certain viewing conditions. The ontology of concluded colors would thus have to be paired with an ontology of sensed colors in order for the former to be complete. For instance, if sensed colors are thought of as projected properties, then concluded colors would be identified with dispositions of objects to be (mistakenly) represented by color experiences as having particular sensed colors under certain viewing conditions[158].

Note that this discussion of Zaidi does not fit the structure of our other two expositions. In the latter case, we tried to determine what could be said at an ontological level on the basis of an approach to color constancy. However, it is not clear that this is what we have been doing in Zaidi's case. Here it seems that we are delving into Zaidi's direct opinion on ontological matters. If this is indeed what we have been doing, then it is not clear that this represents an example of the proposed methodology in action. However, note that the similarity of Zaidi's opinion with Cohen's opinion on color constancy may actually indicate that we are indeed talking about Zaidi's approach to constancy. The discussion would be at a very abstract level but it would still be a discussion about color constancy and not about color ontology. If this is true, then it has to be admitted that Zaidi owes us an explanation regarding both how the visual system comes to know about sensed colors and how the rest of the neurological machinery comes to know about concluded colors.

---

[156] Note that this is reminiscent of Cohen's understanding of color constancy (Cohen, 2004). According to the latter, color constancy is achieved through our ability to tell what color an object would have under another illuminant, that is, by a counterfactual judgment.

[157] This would entail that we have the ability to correctly imagine what dark blue would look like under those conditions in which it would look dark blue.

[158] Note that many will deem this view a pretty cumbersome one.

The methodology and its objections

The relationship between the color ontology and the theory of constancy

We now have a few examples of what it would mean to tease out the ontological consequences of a particular color constancy theory. This process is argued to be the first step of the methodology for color ontology we are here discussing. Before we move on to address some obvious and less obvious objections, I would like to take some time to explain the methodology further and, most importantly, discuss the nature of the evidential relation there may be between the ontological consequences of the color constancy theory and the theory itself. I start with the latter task. As this chapter is exploratory, the following should not be seen as the last word on the matter. On the contrary, it is to be seen as the start of the reflection.

If color constancy theories have ontological consequences, one may rightly ask what the nature of the relationship between the theory itself and its consequences is supposed to consist in. That is to say, we can ask whether the consequences are logically entailed by the claims made by the theory of constancy or whether the confirmation of the latter would only represent empirical evidence for the ontological claims. In the second case, color constancy theories would motivate certain views of color ontology, respectively, but they would not logically entail them.

I believe it can be argued that the first of these options is a non-starter, as long as we limit ourselves to current scientific knowledge. The thesis that color constancy theories on their own logically entail claims about color ontology seems impossible to defend. In order to show that it does not hold, we have to show that a particular theory of constancy does not logically entail the ontological view of colors we have associated with it. Of course, demonstrating this point is not equivalent to demonstrating that ontological claims about color never logically follow from a color constancy theory. But I would argue that the latter would be more the exception than the rule.

In order to demonstrate this, I therefore suggest that we consider one of the pairs of color constancy theory and color ontology we have examined in the current chapter, namely, Maloney and Wandell's theory and its associated color ontologies. Of course, it may be noted at the outset that the very fact that two distinct ontologies have been identified as justified by Maloney and Wandell's theory is already proof that the claims a color constancy theory makes do not logically entail a given color ontology. However, the fact that we have identified two theories of color ontology and not one is mostly due to taking into consideration Maloney's comments on his theory and not strictly speaking the theory itself.

To be fair, we may want to check whether the claims made by the theory itself logically entail claims about color ontology.

The core of the theory can be summed up as follows. First, it openly embraces a reverse optics view of vision (and thus also of color vision). Remember that, according to reverse optics, the final phase of the visual process will consist in visual judgments regarding the causes of the retinal image. More specifically, in the case of color constancy, the visual system is thought to reach conclusions regarding *inter alia* the nature of the SSR and the SPD causing the color signal in question[159]. Maloney and Wandell's theory is a particular implementation of the framework of reverse optics. It proposes that the visual system uses an algorithm which decomposes the color signal into two weighted sets of basis functions representing the visual system's knowledge of the relevant SSR and SPD of the scene, respectively.

As we already discussed above, from the outset, it appears difficult to understand how we could derive RP from this theory. It is opaque how we are meant to identify the colors we supposedly see, the greens, blues, reds and so on we are familiar with, with the SSRs as recovered through the proposed algorithm. It seems that we would need some sort of claim to the effect that, when we see a particular surface as, e.g., yellow, the yellow we see is actually the visual system's representation, through a particular weighing of the fixed basis functions, of the SSR of the surface. For instance, we would need to argue that the color content of the yellow experience is actually that a particular surface has a certain weighted set of basis functions. The output of Maloney and Wandell's algorithm would here be the content of the yellow experience which corresponds to its yellow aspect. Nevertheless, it is not clear how we are supposed to motivate such a claim on the sole basis of Maloney and Wandell's theory.

If it is indeed the case that we need an additional hypothesis so as to identify the colors we see and the representation of SSRs effected by the visual system according to Maloney and Wandell's theory, then it seems to be undeniable that RP does not logically follow from this theory on its own. At the very best, RP would follow from a combination of the color constancy theory combined with a particular assumption linking the colors of our color experiences (the 'colors-as-experienced') and the representations of SSR. For instance, a causal claim to the effect that the colors-as-experienced are a direct effect of our subpersonal representation of SSR may be enough. Or it is possible that our understanding of the

---

[159] The visual system should also reach conclusions regarding the angle of viewing of particular objects, the presence or absence of particular atmosphere (such as haze or fog), etc.…

relationship between our phenomenology and its neurological implementation is wanting. In that case, discovering the links between the two would bridge the gap.

But this is just to say that more scientific advancements would help in this situation. What if Hilbert and Byrne want to defend their theory with Maloney and Wandell's model of color constancy *right now*[160]? Are there any philosophical claims they can make to make RP follow from the theory of linear models? It is difficult to imagine such a claim, given that the link we are missing at this point lies between a particular set of weights, that is, an output of the algorithmic process described by Maloney and Wandell, and the particular colors we supposedly see when we are having color experiences. Once this link is provided, it seems to follow logically that those colors are indeed SSRs.

Regarding the particular claim needed, one could argue that, in having a color experience, we are becoming aware of the outputs of some perceptual processes[161]. This could represent the roots of a claim that we need to bridge the gap between reflectance physicalism and Maloney and Wandell's theory. We could specify this claim further and hold that the outputs we are becoming aware of are the outputs of an algorithm which is correctly described by Maloney and Wandell's theory. Having an experience that an object is red would then be equivalent to being aware of the output that this object has such-and-such reflectance. By having an experience of redness, I would thus be having an experience of a particular SSR. It would then seem to follow that the red is the reflectance. And so on for the other colors.

However, it does not look as if this is a philosophical claim. One could marshal distinct reasons to show that it is not. First, it is unclear how we could establish it using the traditional methodological tools of philosophy. Neither phenomenological reflection nor appeals to intuition could seem to establish a thesis which involves reference to estimates of surface spectral reflectance or subpersonal perceptual processes. The fact that it involves reference to such scientific objects of study is another reason why one may hold that it is not a philosophical claim[162].

Regardless, note that the important point is that, even if we had reasons for a claim Hilbert and Byrne could add to the conjunction of RP and Maloney and Wandell's theory which would make the first a logical conclusion of the second, this still entails that the color constancy theory itself, on its own, does not imply a particular color ontology.

---

[160] For starters, it may be noted that, as Chirimuuta rightly emphasizes in her work, Maloney and Wandell's theory has not been shown as clearly superior to other theories in explaining our color constancy abilities.

[161] For more on this particular claim, see appendix B.

[162] However, note that arguing that it is not a philosophical claim may force us to also argue that Hilbert's ontological theory is not a philosophical theory either.

Considering Hilbert and Byrne's original argument from color constancy for RP does not help us either. According to its most recent form (Byrne & Hilbert, 2020a), colors are experienced as illumination-independent properties of objects. Moreover, we have reason to believe our color experiences are generally veridical. Therefore, colors are illumination-independent properties of color. More precisely, they are ways of altering the light, that is, surface spectral reflectances. Although it would need to be unpacked quite extensively, note that the argument in this form does not license the conclusion that RP is true. Since the immediate conclusion is that colors are illumination-independent properties of color, many relevant properties of objects could fit the bill. If we put aside the relational properties we discussed in the case of Foster's theory, note that, for instance, the categorical bases of SSR could be the colors according to this argument. Those bases are after all illumination-independent and are obviously causally related to our color experiences since they realize the SSR at issue. Therefore, Hilbert and Byrne's argument from constancy does not logically lead us to RP[163].

We are effectively moving further and further away from a simple logical relationship between Maloney and Wandell's theory on the one hand and a form of color realism such as RP on the other. The same conclusion could have been drawn had we looked at the other color constancy theories discussed in this chapter. Whether it be Zaidi's corresponding color constancy theory or Foster's theory of CER, both of them only provide defeasible justification for the respective color ontology[164]. We therefore have to conclude that color constancy theories provide empirical justification for ontological claims about color but do not logically entail them. A given ontological claim will be empirically motivated but not logically mandated by a color constancy theory[165].

Objections and replies

Now that we have a better grasp of the proposed methodology of color ontology, we can try and answer some objections regarding its viability. We shall first address the objection that the present approach seems to have too narrow a focus. It can be asked why we restrict ourselves to color constancy when there are so many other parts of color science which may

---

[163] However, that very fact actually constitutes a reason to endorse the thesis that constancy theories constrain but do not logically entail ontological claims about color.

[164] Subsequently, when it is said that some color ontology or some ontological claim about color is a consequence of a color constancy theory, this is not to be read as a logical consequence.

[165] Note that this is in concordance with Hilbert and Byrne's opinion. Nowhere in their work is it claimed that their theory of colors follows logically from Maloney and Wandell's theory. They have always strived to show that recovering reflectance is a possibility according to color constancy research and that, if we then identify colors with reflectances, it cannot be said that colors are perceptually inaccessible.

be relevant to color ontology. Secondly, we shall touch on what I call *Akins and Hahn's challenge*, according to which it has to be explained why in the case of color we consider evidence coming from research on the *perception* of color when we would obviously not do that in the case of other properties such as shape or size. Finally, we shall determine whether the present approach can be reconciled with the thesis known as *Revelation*, according to which the nature of colors is revealed in perceptual experience. Answering those objections will allow not only to alleviate any doubt one may have regarding the methodological point at issue but also to be even more precise in our specification of its traits.

Why color constancy ?

One astute reader may wonder about the reason the methodology proposed in the current chapter was restricted to the use of color constancy research when there are so many other fields in color science. For instance, research on color afterimages aims to determine why color afterimages occur, how they are neurologically implemented and whether they serve a function or are a mere byproduct of "such a complex and superb machinery known as the human visual system" (Gori, 2015, p. 12). Other researchers have tried for decades to determine why there is such a thing as simultaneous color contrast, where surfaces with identical SSR appear to have distinct colors depending on the SSR of the background on which they are placed[166] (see, e.g., Soranzo, 2016).

The astute reader may thus ask "why is it that it is theories of *color constancy* and not theories in some other area of color science which can empirically motivate ontological claims about color?". There may indeed be theories in other fields of color science which could also be seen as motivating such claims. So our methodology should either embrace those other areas or explain why it is only color constancy that should be the focus of color ontology.

It may be answered that color constancy theories have a form which make them suitable to represent justification for certain ontological claims about color. In the phenomenally-characterized conception of constancy, the role of the visual system is to "stabilize" the color percept attributed to a surface in the face of varying unrelated factors. This process can be effected through an inversion of the optical processes responsible for the retinal image, as is found in theories embracing reverse optics. Or it can be realized by other kinds of strategy, such as described by Foster's theory. The important point is that the stable

---

[166] For a striking example of simultaneous contrast, see appendix B.

percepts which are attributed to objects in the course of this process have the highest claim to the title of color since we tend to think of colors as stable across scene changes.

Of course, as we have seen, there are exceptions to this. Zaidi appears to reject the idea that colors are stable at least if we are talking about sensed color and restricting our focus to our color phenomenology[167]. Some other researchers have embraced this idea and rejected, in one way or another, the notion of phenomenally-characterized constancy (Logvinenko, 2013; Logvinenko et al., 2015; Logvinenko & Demidenko, 2016; Valberg & Lange-Malecki, 1990). In other words, Zaidi and those authors hold that there is no color constancy at a phenomenological level. If this is true, then color constancy research cannot be about finding out an algorithm which can 'stabilize' the color percept.

Moreover, our behavior and performance in asymmetric matching, color naming or achromatic adjustments may be explained by strategies which do not involve the attribution of stable percepts to objects by our visual system, contrarily to the theories which, as we saw in chapter 3, involve a *measurement* of SSR. Furthermore, there are other phenomena which push against the hypothesis that there is such a thing as phenomenally-characterized color constancy. For instance, the existence of simultaneous color contrast (and assimilation, which is the opposite phenomenon) appears to sustain the claim that colors are not stable across scene changes[168]. For those reasons, it cannot just be claimed that there is a feature shared by all color constancy theories but not by theories outside of the field which explain why the former and not the latter are suited to support ontological claims about color.

An alternative answer would be to draw on Davies' defense of color constancy as the necessary and sufficient component of color vision (Davies, 2018). According to Davies, an organism possesses color vision if and only if that organism possesses color constancy. Therefore, if we believe that the ontological questions about color can be informed by an answer to the question regarding what a visual system with color vision is doing, then it is only towards the phenomenon of color constancy that we ought to turn to. It would thus be justified to say that it is only color constancy research that can provide justification for various ontological claims about color.

---

[167] When it comes to concluded colors, Zaidi may be on board with the idea that post-perceptual judgments regarding them involve an element of constancy.

[168] Of course, one may counter the relevance of these observations by arguing for a more nuanced view of constancy, according to which color constancy is more or less present in certain conditions. Following that line of thinking, one could argue that the theories of constancy all point towards the existence of color constancy as a phenomenological phenomenon. However, this does not answer for the numerous researchers which fruitfully approach constancy from a non-phenomenological perspective.

However, Davies' argumentative strategy is quite unclear. Through various expositions that span ethological studies, psychophysics and other areas of color science, Davies aims at showing that color constancy is constitutive of color vision in the sense that it is both a necessary and a sufficient feature of color vision. But to show that the possession of feature F is necessary and sufficient for the possession of feature G, it should be possible in principle to determine which objects possess G independently of whether they possess F. For instance, if we want to determine whether all birds can fly, we ought to have a criterion to determine what counts as a bird besides having the ability to fly. In the case at hand, one ought to be able to identify organisms possessing color vision through a criterion that does not refer to color constancy. But it is unclear that Davies can do this. The term "color vision" is used in various and conflicting ways throughout the literature and, barring a definitional project, it is thus bizarre to say that there is a set of organisms which can be said to have color vision and can be independently picked out. It would be akin to saying that every piece of gold is shiny by picking out pieces of gold only if they shine.

A more ecumenical answer to the objection could actually be to just bite the bullet and accept its conclusion. There are many other theories in the field of color science and theories which have ontological import regarding color need not necessarily concern color constancy. In that sense, the present work uses color constancy research as an example in which (at least some) theories have such import. But this does not mean that there are no other examples to be found. There are few downsides to biting the bullet here. The only one we may think of is that the importance of research on color constancy is downplayed to some extent if we allow other fields of color science to play an evidential role in color ontology. I believe we can live with that.

Akins and Hahn's challenge

Byrne and Hilbert, in one of their most important papers on the subject (Byrne & Hilbert, 2003), argue that "the problem of color realism is primarily a problem in the theory of perception"[169]. They therefore view the question at hand as a question concerning our perception of the external world. The methodology proposed in this chapter appears to endorse a similar view of the issue. By encouraging us to first determine which ontological claims are justified by color constancy theories, it stands alongside Byrne and Hilbert in claiming that one needs to consider color perception in order to answer ontological questions

---

[169] More specifically, they hold that the problem lies with the representational content of color experiences but, regarding the present objection, we do not need to discuss that further precision.

about color, even if this is not by far the whole story. That is, addressing ontological questions about color requires an understanding of our perception of color[170].

Philosophers who have expressed concern regarding an approach to color ontology which necessarily involves such extensive consideration of color perception and color experiences include Akins and Hahn (Akins & Hahn, 2000, 2014). They hold that the contemporary framing of the problem of color has been bequeathed to us by Hardin in his seminal book (Hardin, 1988). According to their reading of Hardin, the problem of color realism boils down to the issue of explaining why the relationships between color experiences on the one hand and their physical causes on the other are so unruly[171].

If one can do so while showing that our color space is actually matched by the space of their physical causes, then color realism is saved[172]. Otherwise, one can endorse eliminativism but then has to explain why color perception is such an important part of our perceptual apparatus even though it does not track any real property of the outside world. Independently of the answer we favor, the road to be followed is clear. On the one hand, we should have a firm grasp on the nature of SSRs and the part of physics which concerns them. On the other, we should have a clear picture of our color space and its properties, that is, of the relations which hold between the various colors we see.

Akins and Hahn state that this way of conceptualizing the tasks of color ontology is a peculiar one:

> "The framing of the problem ought to give us pause. There are goat experiences and size and shape experiences too. But it would be very odd to claim that whether or not goats or shapes or sizes are real features of the world is to be determined by whether we can find anything in the world which is related systemically enough to the content of our goat, shape, or size experiences. Odder still, to claim that what goats, or specific sizes and shapes are, is constituted by whatever appears to be goats or specific shapes or sizes in standard conditions. […] And should there be a section [in this book] on shape or size, one would expect it to be about shape and size perception and not expect to find anything there about whether shapes or sizes are real properties of the world." (Akins & Hahn, 2000, p. 17)

---

[170] However, the methodology argued for in this chapter distances itself from Hilbert and Byrne's by acknowledging that the questions we may have regarding color perception can and probably definitely will include questions regarding the subpersonal processes involved in color perception. Again, given the present objection, we can just ignore this difference in outlook.

[171] Those relationships are unruly especially because of the phenomenon of metamerism, the existence of perceptual variation, the particular properties of our color space (which include the existence of similarities and differences between the colors), and so on.

[172] Or, at least, it is saved for Hilbert and Byrne, who endorse representationalism about our perceptual experience. According to this theory, if an object appears blue to me, then this is analyzed as this object being represented as blue. Representationalism ensures that the blue of the object, when the content of my experience is true, is an actual property of the object.

Akins and Hahn here compare the usual understanding we have of color ontology with our attitude regarding other properties such as shapes and sizes (and, quite comically, goats) and notes that the two differ quite starkly. While we do not look to perception, its mechanisms or the resulting experiences for guidance on matters pertaining to the ontology of size or shape, we do precisely so in the case of color.

Akins and Hahn's challenge is thus the following. Why is it that we adopt such an attitude with regard to color when we would promptly reject it in the case of other properties such as shape or size? We are starting our investigation by asking about color perception when the investigation itself is about color, not the perception of it. The question we ask ourselves is *What is color?*, not *How do we perceive color?*. This calls out for an explanation, according to Akins and Hahn. I believe that this is a serious objection to the methodology of color ontology I present in this chapter. If we do not find a suitable answer to this objection, then the whole project will appear very dubious. Fortunately, I also believe that there are ways to deflect the objection and here I lay out two of them. I start with the more optimistic one and end with the one which concedes the most to the challenge.

The first answer we can have to this issue holds that the proposed methodology is a viable way to do color ontology, that is, one which is worth pursuing. To begin with, consider this quote from Hardin:

> "When someone tells me that she has a theory about colors, I expect it to be a theory of yellow and green and the like, and if I get a story about spectral luminance or reflectance profiles, or whatever, I want to know how all of that relates to those qualities that I know and love. If a pusher of chromatic theory can't spell out these relationships particularly well, I am disappointed, but if she tells me that colors as she understands them don't include the hues, I feel cheated. No matter how brilliant her discourse, she has changed the subject." (Hardin, 1988, p. xx)

Hardin is here pointing out a very common desideratum of an ontology of colors, one which is accepted by many if not all philosophers working on the subject (see, e.g., Byrne & Hilbert, 2020; Cohen, 2004; Hilbert, 1987; Johnston, 1992; Wright, 2019). A theory trying to explain the nature of the colors has to explain why we see the hues which we do see, what those hues are and, depending on the nature of the theory involved, what the relationship between those hues and spectral reflectances consists in. That is to say, the theory has to tell us about green, blue, red, yellow, and so on[173].

---

[173] Even though Hardin only mentions the supposed four unique hues, he does include all the remaining hues, as is made clear in the paragraph following the quote.

We can understand Hardin's point in the following way. Whereas we may have theories of what physical magnitudes could explain the brightness and saturation of colors (respectively explained by the absolute intensity of the light and the predominant wavelengths of that light), it is not so in the case of hue. Hues seem not to be correctly explained by any physical magnitude. Therefore, we have to look elsewhere and where would we turn ourselves to if not to the science on color perception?

Another way to understand this requirement is to see it as demanding that the theory of colors put forward should make theoretical contact with the hues (or colors) as visual qualities. To say that hues are visual qualities is not to say that they are mental entities. It only means that we encounter or come to know about hues through vision[174]. An investigation of what it is our visual system is doing when we see a given hue thus seems a promising start in order to establish the explanandum. If we understand what our visual system is trying to do when it processes the retinal information in order to generate our perceptual experiences of hues (or colors), we may be able to determine whether it is trying to track a particular property when it does so or whether it is creating an arbitrary classification to ease later recognition and identification of the object.

Other options exist. For instance, if Chirimuuta's recent opinion is correct (Chirimuuta, 2015), what our visual system is doing when we have perceptual experiences of color is to present us objects or some of their properties in a particular way. According to this line of thinking, which Chirimuuta argues is motivated by the recent color science, seeing an object in color is equivalent to seeing that object (or some of its properties such as its shape and texture) in a certain way. Colors are ways objects (and their properties) appear to us. The important point here is that, in investigating color perception at the outset, we are effectively trying to determine why we came to have it and color experiences in the first place and this will motivate an answer to the ontological question[175]. Accordingly, the question about ontology is subsumed under the question about perception[176].

A more pessimistic answer to the challenge concedes to Akins and Hahn that the proposed methodology is not the ideal one. However, it also holds that there is no better

---

[174] The issue with the latter phrase is of course that it may imply that hues are real, mind-independent properties. The expression "encounter" may here be more neutral.

[175] Alternatively, a very likely possibility is that, in studying color perception, we may come to realize that the ontological question as we have formulated it is a bad question and the only way to put it to rest is not to solve it but to dissolve it.

[176] Another potential issue with Akins and Hahn's challenge is that it may beg the question for non-realist theories. If we study color directly rather than the perception of it (whatever that means), we are effectively saying that there are colors and eliminativism and theories which hold that colors are perception-dependent properties are automatically discredited. Starting with perception avoids that issue.

option available. If we are to do color ontology, we have to start here, with color vision, and especially with color constancy. To see why this answer is actually a live one, we have to consider the alternative and show that it does not seem more fruitful but quite the opposite. To see this point, consider the way this alternative is phrased by Akins and Hahn themselves:

> "Contrary to the facts of our phenomenal space, we could find out that purple *qua* a colour of the distal world is not equidistant between blue and red, that brown and orange are extremely closely related to one another, that red is closer to heat than it is to violet (no matter how nonsensical that sounds to our ears). There are facts about colours, and then there are facts about how we perceive colours. And there is no a priori relation between them, although there are many empirical ones." (Akins & Hahn, 2014, p. 19) (original italics)

That is to say, we ought not to study color experiences if what we want an answer to are ontological questions about color. We ought to study color itself, not the way we perceive them, regardless of whether this last expression is supposed to refer to color experiences or the perceptual processes responsible for color vision.

However, it is methodologically unclear how this proposal is supposed to be implemented. How exactly would we find out the things we could find out according to Akins and Hahn? For instance, consider the question as to whether purple is equidistant between blue and red. If studying color experiences is out of the question, what are the experiments we should perform, the results of which would motivate one answer over the other? Barring any identification of purple, blue and red with certain SSRs or other physical properties, it is a mystery how we ought to proceed. Even if we were to perform such an identification, we would also need to know what it means for an SSR (or a set of SSRs) to be equidistant between two other SSRs (or sets of SSRs).

Or consider the question as to whether red is closer to heat than it is to violet. We could see whether surfaces with SSRs biased towards the longer end of the visible spectrum are more similar (whatever that means) to heat than they are to SSRs on the lower end of the spectrum. If they were, would we conclude that we have shown that red is closer to heat than it is to violet? Again, barring any identification of the longer-end SSRs with red, it seems impossible to conclude such a thing. It thus appears difficult to follow the methodology proposed by Akins as long as we do not have any clarification on how it is supposed to be implemented.

It is my hope that one of these answers, either the optimistic or the pessimistic one, is enough to answer this challenge, which is, I reckon, an important counterargument to the methodology proposed in this chapter.

Revelation

Another objection to the methodology of color ontology presented in the current chapter rests on the much-discussed epistemological thesis of *Revelation*. This thesis appears in its formal expression in Johnston's work:

> "Revelation. The intrinsic nature of canary yellow is fully revealed by a standard visual experience as of a canary yellow thing." (Johnston, 1992, p. 223)

And so on for the other colors. If *Revelation* is true, then color experiences reveal to us what colors are[177] [178].

Of course, there is a number of precisions which can be offered to clarify Johnston's idea. *Revelation*, as stated, seems to imply that it is sufficient to have a single experience of a color to then know its nature. However, it is more likely that a somewhat repeated exposure would be necessary if any amount is. In addition, just having the experiences of color will usually not be enough to then know the nature of the color at issue. Rather, the experiences put one in a position to know (Allen, 2011). So a young child, who might not be able to carefully reflect on their experiences, might not come to know about the nature of the relevant colors. It is usually thought that careful reflection on the experiences is an implicit element of *Revelation*.

Moreover, it is not clear exactly what Johnston meant by the 'intrinsic nature' of color. For our purposes, however, it is enough to understand this phrase as referring to the set of properties which are necessary for the color at issue to have in order to be that color. In other words, if a property P is such that it is necessary for a color C to have P, then having P is part of the nature of C. For instance, it may be in the nature of yellow to be more similar to red than to blue. A color that would be more similar to blue than to red could simply not be a

---

[177] This talk of 'revealing' could be fleshed out in distinct ways. For instance, one can say that an experience reveals the property P if and only if P is part of the content of the experience. Of course, the thesis is to be understood in a much stronger way. More on this below.

[178] If color experiences are identical with certain brain functions or states, someone might argue that *Revelation* is ambiguous in that careful reflection on the corresponding brain functions or states could reveal to us what colors are. But note that it is having color experiences from the first person point of view that is supposed to reveal to us what colors are.

shade of yellow. In contrast, it is not in the nature of blue to be someone's favorite color. Blue could fail to be anyone's favorite color but still remain blue.

Another important distinction is the one brought by Byrne and Hilbert between two versions of *Revelation* (Byrne & Hilbert, 2007). According to the weak version of *Revelation*, color experiences reveal the nature of colors but the latter is not revealed as such in the experience. For instance, if a property P is part of the nature of what it is to be blue, then it is revealed in an experience of blue that blue has P but not that having P is essential to what it is to be blue. On the contrary, a strong reading of *Revelation* entails that P is revealed in the experience of blue as an essential property of what it is to be blue. In other words, P is revealed *as* part of the nature of what it is to be blue.

As it is the strong reading that is usually assumed in the literature, Byrne and Hilbert analyze it as the conjunction of the two following theses:

> "SELF-INTIMATION If it is in the nature of the colors that p, then after careful reflection on color experience it seems to be in the nature of the colors that p.
>
> INFALLIBILITY If after careful reflection on color experience it seems to be in the nature of the colors that p, then it is in the nature of the colors that p." (Byrne & Hilbert, 2007, p. 77)

As can be seen quite easily, *Self-intimation* raises an issue for many theories of color. For instance, if colors are reflectances (or reflectance sets), then by *Self-intimation* this should be revealed in color experiences. But it clearly is not. So reflectance physicalism is false.

Crucially for our purposes, *Revelation* is also in conflict with the methodology proposed in the current chapter and this for at least two reasons. First, if we manage to find out how color constancy is achieved, that is, if we find out the correct theory of color constancy, this theory could motivate claims about the nature of color which are not apparent after reflecting on our color experiences. This would contradict *Self-intimation*. Secondly, *Revelation* entails that scientific investigation, including the whole of color science, will be at the very best superfluous and in all likelihood useless when it comes to discovering the nature of colors. This worry is best put by McLaughlin[179]:

> "There is much that we've learned about light and how it interacts with matter. We've investigated how the chemical properties of surfaces and volumes made up of various materials dispose them to interact with light;

---

[179] We also owe to McLaughlin the distinction between *Revelation* regarding the nature of colors and *Revelation* regarding what it is like to see colors. According to McLaughlin, the former appears so intuitively true because the latter is so intuitively appealing (even though McLaughlin will end up rejecting both of them).

> how rods and cones respond to light; how electrochemical impulses are propagated along the optic nerve to the visual centres of the brain, and resulting patterns of neural activity therein. From an information-theoretic perspective, we've investigated how the visual system might process information about the scenes before our eyes […] While there is much that we still don't know, we've learned an enormous amount; and research in vision science is moving apace. It is a consequence of the doctrine of Revelation, however, that all we've learned and, indeed, all we can ever hope to learn by scientific investigation will contribute not one whit to our knowledge of the nature of colours themselves. For Revelation entails that there is nothing more that we can learn about the nature of colours than what visual experience teaches us. […] As concerns knowledge of [colors], there is no substitute for experience." (McLaughlin, 2003, p. 477)

The last sentence actually expresses a component of *Revelation* that is usually assumed but not explicitly acknowledged. Note that *Self-intimation* and *Infallibility* do not together entail that experience is the only source of knowledge we may have regarding the nature of colors[180]. However, this is exactly how *Revelation* is supposed to be read. As McLaughlin explains it, our scientific investigation into, e.g., how color constancy is achieved or the kinds of properties which determine how objects will react to light shined on them will not in any way advance our understanding of colors themselves.

But this last assumption is at the core of the present methodological advice for color ontology. According to the latter, there may be empirical discoveries which can lead to a better understanding of the nature of colors, discoveries which could not be made through an examination of our color experiences. In our particular case, it was argued that the correct theory of color constancy could motivate a particular view regarding the nature of colors. This is exactly what is denied by the thesis of *Revelation*. Subsequently, in order for the methodology at issue to be a viable path to tread, we have to reject *Revelation*.

Of course, a very easy way to reject it would be to say that it does not fit with the project of the present work. The methodology proposed here hopes to be part of a naturalistic approach to color ontology. But note that the conception of perceptual experiences that is brought with *Revelation* is just not one which can fit within this kind of approach. The latter will probably endorse a view of perception and perceptual experiences where those are seen as tools biological organisms use to navigate their environment. Perception is seen as a mean to an end and this end does not require us knowing about the nature of the properties we find

---

[180] It is possible, although very unlikely, that we would reach the same conclusion regarding the nature of color, independently of whether we reached it through a careful examination of our color experiences or through empirical study of color vision. *Self-intimation* and *infallibility* do not strictly speaking entail that this could not happen.

in color experiences. So there is no reason to believe that this nature is revealed in such experiences.

However, we can probably do better than this blunt response to the proponent of *Revelation*. I believe that the best way to reject this thesis is to be found in Allen's defense of naïve realism (Allen, 2011). According to the latter, which counts as a form of primitivism, colors are *sui generis* mind-independent properties. In the course of defending this theory, Allen develops an argument to show that *Revelation* is actually not tenable. This argument is in fact independent from Allen's naïve realism and may be accepted by theoreticians from any board.

In a nutshell, Allen's reasoning is the following. If *Revelation* is true, then by definition *Self-intimation* and *Infallibility* are true. But if the latter are true, then for any proposition p that is such that it appears after careful reflection on color experiences to be in the nature of the colors that p, then it should be relatively obvious to everyone carefully reflecting on their color experiences that p is in the nature of the colors. However, this is exactly the opposite of what we find. Instead of widespread agreement or superficial disagreement, we find that philosophers, scientists and artists who have considered their color experiences very carefully disagree in a pretty profound way with regard to what is essential or accidental to colors. Therefore, by *modus tollens*, *Revelation* is false.

To motivate his argument, Allen focuses on the questions of whether colors are causally efficacious and which higher-order properties are essential properties of colors. For the sake of brevity, I shall only develop the latter question, noting that the reasoning regarding the causal efficacy of colors is structurally similar to the one concerning the higher-order properties of color.

As instances of higher-order properties, Allen focuses on the elemental-compound distinction. Elemental colors are thought to be phenomenally uncomposed, whereas compound colors are phenomenally composed. Phenomenal composition is distinct from physical composition. For instance, mixed pigments can give rise to a color distinct from the color of the original pigments. The resulting color can be said to be physically composed but it does not entail that it is phenomenally composed from the two original pigments' color. In contrast, a color that is phenomenally composed somehow looks like the colors which compose it.

Take orange as an example. Orange is usually thought to be a compound color, that is to say, a color that is phenomenally composed. In this particular case, the composing colors are red and yellow. Every shade of orange is supposed to look both reddish and yellowish to varying degrees. But yellow appears to admit of a shade that neither looks reddish nor greenish. That is, it admits of a shade that is unique, i.e., phenomenally uncomposed. The same appears to be the case for green, red and blue.

That red, green, blue and yellow are the only colors which admit of unique shades is supported by the fact that, when people are asked to describe various shades using only those color terms, they can do so pretty well and without felt loss in the description (Sternheim & Boynton, 1966). This is not so when they are restricted for instance to red, yellow and blue or when the color terms include one or two elemental colors (say, red and blue) and phenomenally composed colors (like orange and teal).

As Allen argues, the elemental-compound distinction will be part of the nature of the colors if it applies at all to them. So it should be in the nature of, e.g., red that it admits of a unique shade, whereas the opposite is true of the nature of orange. If there are only unique hues of red, green, blue and yellow, then it is in the nature of those shades that they are unique. Issues arise when we consider that the extension of the distinction has been much debated. Even worse, the very existence of the distinction has been put into question.

Even though it is not pointed out by Allen, note first that even naïve subjects can behave in a way which indicates that they may see some shades of orange as unique. One experiment found that there were twice as many subjects that were willing to describe some shades as orange without also describing those colors as reddish or yellowish than those who were not so willing (Sternheim & Boynton, 1966)[181]. However, if orange was a compound color (i.e., if all shades of orange were a phenomenal composition of red and yellow), then there ought not to be any shade of orange which does not look both yellowish and reddish.

Nevertheless, it may be argued that naïve subjects are not the best source of evidence regarding the extension of the elemental-compound distinction. It could be said that they are not experts at carefully reflecting on their color experiences to determine whether the shade they are having an experience of is an elemental or compound one. This is the reason Allen actually focuses on whether there were deep disagreement regarding the extension of the distinction at the level of researchers who can perhaps be said to be experts in this particular domain.

Allen thus points out to a tradition of thinkers who rejected the idea that green was an elemental color. Those included Berkeley, Goethe and Brentano. In addition, many thought

---

[181] Unfortunately, Sternheim and Boynton do not really comment on this finding of theirs and still end up concluding that "the hues associated with the long wavelength part of the spectrum can be adequately described without the orange category" (Sternheim & Boynton, 1966, p. 776).

that brown suffered from the opposite issue. It did not seem evident to many that brown was a compound color that could be entirely described without using the word "brown" (Fuld et al., 1983). It was seen as necessary to design an experiment that showed it was, in fact, a compound color (Quinn et al., 1988)[182]. Furthermore, contrarily to the dominant opinion, some researchers have argued that orange and purple must also be considered as elemental colors (Beare & Siegel, 1967; Padgham & Saunders, 1975)[183]. This matter was also apparently put to rest with the designing of new experiments showing that those colors do not behave as elemental ones (see, e.g., Fuld et al., 1981).

It is not only the extension of the elemental-compound distinction that has been fiercely debated. At the end of the nineteenth century, the existence of the distinction was also put into question. Whereas Brentano was maintaining that green was phenomenally composed, "other participants, including Külpe, Titchener and Ebbinghaus, denied that any colours are phenomenally composed, even orange or purple" (Allen, 2011, p. 165). Allen quotes Boring which "describes an attempt to settle the question of whether there is an elemental-compound distinction at a meeting of the American Psychological Association:

> Ladd-Franklin [a proponent of the view that some colours are phenomenally composed] appealed to the consensus of expert opinion by exhibiting [differently coloured] disks and asking for the judgements of psychologists. The problem was never settled. It simply disappeared." (Boring, 1942, p. 131 in Allen, 2011, p. 166)

It does not even appear clearly to everyone involved that there is such a thing as a compound color.

Those deep disagreements represent an issue for the proponent of *Revelation*. Remember that *Self-intimation* states that if it is in the nature of the colors that p, then after careful reflection on color experience it seems to be in the nature of the colors that p. Take the issue of whether there is a shade of green that is elemental. Let us say that there is. In that case, after careful reflection on color experience, it should seem that there is such a shade. But Brentano, Goethe and others have defended quite strongly the opposite thesis. Now let us imagine that it is false and that all shades of green are compound colors. It should then seem

---

[182] This already puts pressure on *Revelation*. If the latter were true, we should not need to design experiments to prove that some property is or is not an essential property of a color.

[183] Note that an issue with those studies, as well as with the one of Sternheim and Boynton (Sternheim & Boynton, 1966), is that they all use a method which involves applying color terms to *lights*. Yet, if there are qualitative differences between our perception of colored lights and our perception of colored surfaces, it is not clear that those results will generalize. For instance, it is unclear what could be said in the case of contrast colors (brown, black, olive, …) which can only be seen with surfaces set on a visible background and not with mere lights projected directly on people's eyes.

after careful reflection that it is so. As we saw, however, there are many psychologists who rejected the whole notion of compound color, arguing that all colors, even orange and violet, are elemental. Either way, it seems like *Self-intimation* is falsified by the fact that there are deep disagreements running.

Infallibility is also in trouble. Remember that it claims that if after careful reflection on color experience it seems to be in the nature of the colors that p, then it is in the nature of the colors that p. Consider now the question of whether there even is an elemental-compound distinction when it comes to the nature of the colors. We know that it has seemed to many that it is obvious that there is such a distinction. Therefore, by *Infallibility*, it is in the nature of the colors that an elemental-compound distinction applies to them. But to others it has seemed that the distinction does not apply to colors. By *Infallibility*, then, it is not in the nature of the colors that the distinction applies to them. Thus, if *Infallibility* is true, then it both is and is not in the nature of the colors that such a distinction applies to them, which is an unwelcome result.

Both components of *Revelation* therefore appear to be falsified by the presence of strong disagreement with regard to *inter alia* the elemental-compound distinction and how or whether it applies to the colors. Another way to make the same point is to consider what ought to have happened if *Revelation* were true. If it were true, then we ought to have discovered the nature of the colors by merely carefully reflecting on our color experiences. By the use of the term "merely", I want to point out that *Revelation* entails that scientific discoveries and technological advancements should be of no help whatsoever in finding out the nature of the colors. In turn, this should entail that we could have found out about their nature at any point in history since we have begun theorizing about our experiences. Arguably, it means that we should have found this out at least since the time of philosophers from Ancient Greece. It can indeed be argued that they were amongst the first to theorize on the basis of their experiences.

But if this is true, then it is curious that we still find important disagreements regarding the nature of the colors more than two thousand years later. If *Revelation* were true, it would be very bizarre to argue that it should take us more than two millennia to come to an agreement regarding the nature of the colors[184]. Our inability to agree on central matters regarding this subject appears to be evidence that *Revelation* cannot be true. However, if *Revelation* is not true, then it is quite normal and even expected to find such disagreements.

---

[184] A lot more, since the current philosophy of color is contemporary evidence that we still have not come up with a shared conclusion on the subject.

Even worse for the proponent of *Revelation*, it appears that some of those disagreements may have been settled by the use of experiments. As Allen puts it:

> "To the extent that these issues have been resolved […] it was not by colour experience. These issues were resolved by the discovery of neural mechanisms that implement the opponent-process pathways which predict two sets of opponently organised hues: red/green and yellow/blue […]. Far from being *revealed* by colour experience, the elemental-compound distinction was a *discovery*." (Allen, 2011, p. 167) (original italics)

If those disagreements were resolved, it should have been through the use of reflection on color experiences, not by experimental study of the neurological bases responsible for color vision[185].

Of course, there are ways for the proponent of *Revelation* to counter the argument. However, in any case, it seems that she will be forced to maintain that the mentioned disagreement does not constitute a falsification of her thesis. Admittedly, the only way to do so is to hold that at least some of the proponents on either side of each debate have not carefully reflected on their color experience and that this is the reason we observe such a disagreement. The disagreement is merely superficial and, were those thinkers to be more careful when considering their experience, would evaporate quite quickly.

Consider the case of whether green admits of an elemental shade. The counterargument of the proponent of *Revelation* would here consist in the claim that Brentano, Goethe and Berkeley have failed to carefully reflect on their color experience or in the claim that their opponents on the issue have. And she would have to use that strategy wherever any disagreement regarding the nature of the colors arises and say that all of the thinkers on one side of the disagreement have failed to carefully reflect on their color experiences. As both components of *Revelation* require such a careful reflection, the observed disagreement would thus not count as evidence falsifying those components.

The issue with this answer is that it seems quite hard to maintain, e.g., that Brentano, Goethe and Berkeley and all similarly-minded thinkers failed to reflect on their color experiences in the manner which is mentioned in *Self-intimation* and *Infallibility*. Why is it that those thinkers, well-versed in both philosophy and psychology (or their dissenting comrades) have failed to do so? Not only should the proponent of *Revelation* have an answer

---

[185] Note that it cannot be said that issues regarding the elemental-compound distinction and the more general structures of hues (such as color opponency) have been resolved. Indeed, a few years after it had been proposed that color opponency was explained by cone opponency, we realized that the perceptions of the unique colors that were supposed to be opponent (red versus green and yellow versus blue) did not align with the cone-opponent excitations when only one of the systems is balanced. For more information on this subject, see Wuerger & Xiao, 2016.

to this question but she must also explain which procedure is the one which constitutes the right way to carefully reflect on one's color experiences[186]. Barring this information, her argumentative strategy appears as an *ad hoc* maneuver destined only to save *Revelation*.

Another way to answer what we may call for lack of a better name Allen's challenge is to dilute down *Revelation*. Remember that we saw that Byrne and Hilbert distinguished between a weak and a strong version of *Revelation*. According to the latter version, *Revelation* entails that the nature of colors is revealed *as such* in color experience. The weak version, on the contrary, only consists in the claim that their nature is revealed but not *as such*. Thus, for any property that is an essential property of a color (such as *being elemental*, *being more similar to yellow than to blue*, etc.), that property will be revealed in color experiences. However, careful reflection on the experience will not reveal that it is an essential property of the color at issue.

Allen's challenge only applies to the strong version of *Revelation* and thus endorsing the weak version sidesteps it. Indeed, weak *Revelation* entails neither *Self-intimation* nor *Infallibility*. However, the issue now becomes that reflecting on our experience is not enough to find out about the nature of colors. To see this, note that under weak *Revelation* our experience will attribute many properties to colors. Some of those will be essential to them, some won't. But the experience itself does not tell us which is which. So the proponent of weak *Revelation* has to come up with a criterion to help us distinguish between the essential and non-essential properties of color. Without this criterion, it is unclear how weak *Revelation* can help us come to a conclusion with regard to the nature of colors[187] [188].

However, weak *Revelation* is still incompatible with our methodology. In all likelihood, the ontological claims motivated by a particular color constancy theory will not be part of the content of our color experiences. That is, for any proposition p that is such that it is in the nature of the colors that p according to the color constancy theory, p will not be part of the content of any of our color experiences. But if weak *Revelation* is true, then those propositions should be part of the content of our color experiences. So weak *Revelation* is still a claim we ought to reject.

---

[186] To the best of my knowledge, there has been no work on answering those questions. At the point of writing, it does not seem that anyone has risen up to Allen's challenge.

[187] Note that denying *Revelation*, either in its strong or weak form, is entirely compatible with holding that the particular knowledge we acquire by having color experiences is knowledge by acquaintance. In that case, by having color experiences, I would know the colors themselves but I would not thereby know anything about their nature (which would amount to knowledge by description). For more on this issue, see Allen, 2011, p. 173.

[188] However, note that the criterion in question cannot come from phenomenological reflection. Indeed, the proponent of weak *Revelation* admits that the answer to the question "is this property that I am seeing this object as having an essential property or an accidental one of that object?" is not to be found in the experience itself.

For instance, if Foster's theory is correct and we in turn accept that colors are differences of preferences for reflecting certain wavelengths over others, then this will not be part of the content of our experience. It is arguably not the case that our color experience is such that it reveals that such a theory is true. This phenomenological claim is in all likelihood false. Perhaps the canonical example is the one of Byrne and Hilbert. If it is indeed true that colors are sets of SSRs, then this is something that is not revealed by a careful consideration of our color experience.

To the best of my knowledge, there has not been any discussion of weak *Revelation*. This could be explained by the fact that this thesis somehow cripples our efforts at finding out about the nature of colors. Although the strong reading of *Revelation* does give us a way of finding this out (by carefully reflecting on our color experiences), this is not so in the case of weak *Revelation*. If the latter is true, then although the essential features of colors are present in experience, there is no way for us to determine which ones of all the features so presented they are. I take it that this is an unwelcome result of the thesis and potentially the reason it has not been more discussed or embraced in the literature[189].

Summary

In this chapter, I have defended a methodological point about the study of color ontology, namely, that the latter could start with an exploration of what our best theory of color constancy tells us about the nature of colors. I first presented a few examples of this point in action, focusing on the cases of Maloney and Wandell's, Foster's and Zaidi's respective theories.

I tried and showed that Maloney and Wandell's theory may motivate an ontology of color which is distinct from Hilbert's. This ontology would have colors consist in potentially mental descriptors attributed to objects by the visual system. Importantly, those descriptors would thus not be equivalent to physical properties of objects. However, it was not clear that the theory is viable as it seems to require Maloney and Wandell's algorithm to function poorly.

I argued that Foster's theory could motivate what I called *objectual color relationalism* (or OCR for short). According to the latter, colors are relational properties of objects constituted by their preferences for reflecting certain wavelengths compared to the

---

[189] Note that what we may call the argument from perceptual function, presented above, also shows weak *Revelation* to be untenable. If the function of perception is to be an effective guide to action in our natural environment, there is no reason to expect that the nature of what we see need to be revealed in experience (even if not *as* such nature). In that sense, weak *Revelation* is not part of naturalistic research.

preferences of other objects in the scene or its chromatic average. Interestingly, OCR can be seen as a form of realism as it does not need to involve references to minds in the specification of the nature of colors.

Thirdly, we touched on Zaidi's theory and the kind of ontological claims it could motivate. Most importantly, the claims at issue would probably involve a distinction between a transient property (the sensed color) and a stable one (the concluded color), where the latter notion is explained by the former one. The account which would here be motivated is one according to which colors are fictional or mental properties, created by the visual system to bring order to a messy visual world.

We then reflected on the link between on the one hand the color ontology and on the other the theory of color constancy at issue. I argued that the link between the two had to be conceived as one of empirical motivation from the theory of constancy to the color ontology (or for some ontological claims), rather than as a logical entailment. As an example, we considered Maloney and Wandell's theory. It was shown that their algorithm could not, on its own, logically entail RP. If successful, the theory of linear models would represent empirical and thus defeasible justification for RP.

Finally, I addressed a few objections which could be made regarding the general approach. First, I touched on the fear that our approach might be too narrow in focus by avoiding other fields in color science. It was concluded that this fear is appropriate and that we ought to accept that color science theories in general may have evidential import for color ontology. We then considered Akins and Hahn's challenge, according to which it is unclear why we ought to consider color perception research when our interest lies with color ontology. I argued that, at the very least, it is unclear how any other methodological practice would get off the ground. To end up, we discussed the thesis known as *Revelation*, which states that the nature of colors is revealed in color experiences. It was shown that this thesis is incompatible with the proposed methodology. As Allen shows us, it appears to entail that we ought to widely agree on matters of color ontology. However, since we find exactly the opposite to be true, we can by *modus tollens* reject it.

In the next chapter, I say a few words to conclude the present project. I start by summarizing each chapter and then move on to some points which have not been discussed yet.

Chapter 6: Conclusion

Summary of the present work

Now that we have come to the end of the present work, I believe it would be fruitful to summarize the main points of the critical chapters. In the second chapter, we dissected a particular debate around the evidence for Hilbert and Byrne's favorite theory of colors, reflectance physicalism (or RP). The argumentative strategy of Hilbert and Byrne was described in the following way. As commonsense would have it, colors are mind-independent properties of objects such as tomatoes, leaves and flowers. However, Hilbert and Byrne argue that we can find motivation for such a view or, more specifically, for reflectance physicalism, in color constancy research. The experimental success of Maloney and Wandell's theory represents a powerful motivation for the identification of colors with reflectances. In that way, the authors provide an empirical justification for a philosophical theory.

However, we saw Chirimuuta has actually expressed a pretty dissenting opinion regarding this issue. Her case rests on two distinct points. First, she argues that the consensus around the theory of Maloney and Wandell which we would need for Hilbert and Byrne's argument to go through is nowhere to be found. Secondly, this absence of consensus is explained by the presence of diverging conceptions of vision which influence the interpretation of the data in conflicting ways.

Chirimuuta first focuses on the question of whether there is a consensus to the effect that Maloney and Wandell's theory correctly describes the perceptual processes responsible for color constancy. She motivates her point by considering the question of whether reflectance recovery is held to occur by all authors in the field. For various reasons, she concludes that it is not the case. First, it is not even clear that our visual system could perform reflectance recovery in theory. Remember that Maloney and Wandell's algorithm has the color signal decomposed into two weighted sets of basis functions, one corresponding to the SSR of surfaces, the other to the SPD of the illuminant. One condition which must be met for the algorithm to function is that this number not be greater than two (that is, the number of cones minus one). Nevertheless, it seems that to correctly render the complexity of visual scenes more than two basis functions would be needed. According to the requirements of the algorithm, it seems that our visual system could never recover reflectance (i.e., *Recovery* would have to be false).

In addition, Chirimuuta shows that many researchers in the field of color constancy research have opted to study distinct visual strategies, in particular, ones which do not portray reflectance recovery as a viable or even desirable step in visual processing. As an example, the work of Zaidi and his team is presented. According to this point of view, the goal of the visual system is to use the color signal in such a way that the organism at issue can 'get by' in the world. The means to this end are argued to consist in heuristical strategies which involve simple treatments of the color signal. Most notably, those strategies eschew reflectance recovery.

Furthermore, Chirimuuta argues that the lack of consensus we can observe is due to researchers adopting distinct research programs. The latter in turn naturally bring along distinct conceptions of the nature of vision. Those conceptions motivate different interpretations of the data in such a way that it cannot even be claimed that the latter speaks unequivocally for or against *Recovery*. As an example, we saw that researchers adhering to the framework of reverse optics construe vision as the recovery of distal information through a reversal of the color signal. Ideally, that reversal would yield a reconstruction of the color signal into the distinct components which generated it, notably, the reflectance of the surface and the SPD of the illuminant. For researchers rejecting that conception of vision, evidence which may have spoken in favor of it is allegedly interpreted in a way which does not.

Chapter 3 consisted in a discussion of the first of Chirimuuta's claim. Even though Chirimuuta is right regarding the absence of consensus concerning whether *Recovery* is true, it is argued that the question of whether there is such a consensus is not relevant in assessing the viability of RP. A theory of color constancy which has the resources to justify RP is not necessarily one which requires the truth of *Recovery*. The pool of theories which could justify RP includes but is not limited to theories which require *Recovery*. The relevant criterion is rather that the strategy described include a measurement of SSR, where the latter is a broader notion than the one of recovery.

To motivate this thesis, I suggested that we consider a theory of color constancy based on the work of Foster. The theory in question has the visual system use the color relations between the distinct surfaces in the scene, expressed as cone-excitation ratios, to achieve color constancy[190]. I argued that, with the addition of an anchoring rule, this theory could represent justification for RP. It thus appears that there are counterexamples to Chirimuuta's claim that

---

[190] Or, rather, something very close, namely, relational color constancy.

there must be a consensus regarding *Recovery* in order for color constancy research to justify RP.

I then addressed a few objections. Most notably, I discussed the question of whether we have a notion of measurement of reflectance which is clear enough to categorize theories of color constancy according to whether they can justify RP. To answer it, I used the list of criteria Akins described in her characterization of veridical sensory systems. According to the latter, a sensory system which can be rightly said to track, measure or detect a property of the mind-independent environment is one that exhibits constant correlation between the properties tracked and the signals corresponding to those properties, preserves the structure of the tracked properties' domain and is servile by only reporting on the properties tracked and not interjecting its own idiosyncrasies in its reports. I then showed that both Maloney and Wandell's and Foster's theories describe visual processes in a way which would fulfill those three criteria.

I ended the chapter by describing Zaidi's theory. Contrarily to Maloney and Wandell but in concordance with Foster, Zaidi eschews reverse optics. According to the author, we ought to adopt a performance-oriented view of color constancy. The latter encourages us to view color constant behavior, and not color constancy as understood phenomenologically, as the goal our visual system strives to achieve. In that sense, the aim of the visual system is to deal efficiently with its environment, even if this entails that there is no constancy at a phenomenological level. Remember that the theory's central element is the strong correlation within cone-opponent channels' excitations across illuminant change. With the help of this correlation and simple operations such as rank ordering, the visual system would be able to achieve identification of similar materials across scene changes.

I tried and showed that, since the strategy at issue does not in any way involve a measurement of reflectance, it cannot justify RP. Indeed, since this strategy does not involve the determination of some of the properties of the visible surfaces but only whether there are pairs of identical surfaces which are present, there can be no measurement or tracking of some of the surfaces' properties. But then it is a wonder how the theory could justify the identification of colors with SSRs.

If the material presented in this chapter is reasonably viable, then this represents a counterargument to Chirimuuta's claim that it is central that there be a consensus regarding *Recovery* in order for color constancy research to support RP. On the contrary, I tried and showed that the pool of theories which can be appealed to in support of RP include but are not limited to theories according to which *Recovery* is true. This opens up the door to new possibilities. For instance, if our final contenders for explaining color constancy are theories which all include a measurement of reflectance in the sense discussed, then this would count as strong motivation for RP.

The aim of chapter 4 was to discuss Chirimuuta's assertion that there are distinct conceptions of the nature of vision at play in color constancy research and that those differences explain why researchers interpret common data in incompatible ways. I argued that this claim entails that there can be no unequivocal consequence of the framework of reverse optics in general and of *Recovery* in particular. However, it seems that some of such consequences have been identified in the field. In addition, some experiments have directly compared color constancy models endorsing distinct conceptions of the nature of vision, which should be impossible according to Chirimuuta's claim.

I started by discussing the notion of reverse optics and what it takes for a theory of color constancy to be counted as endorsing this framework. According to the analysis, a theory endorsing reverse optics ought to (a) consider that the goal of visual processing is acquiring knowledge of the environment it is currently functioning in and (b) achieve this goal by computationally *inverting* the processes responsible for the formation of the retinal image. To validate the analysis, I tried to determine whether it correctly classified a theory which is thought to endorse reverse optics and one which explicitly rejects them. The first was Brainard's "equivalent illuminant model" and it was shown that it does fulfill both conditions and thus counts as endorsing reverse optics. In contrast, the theory of Foster and colleagues, relying on cone-excitation ratios, fails to fulfill criterion (b). This theory does not involve an inverting of the color signal and thus does not count as a reverse optics model of color constancy.

It was then shown that criterion (b) requires that the illuminant be estimated in order to then use that estimate in conjunction with the color signal to estimate the reflectance of surfaces. That the visual system performs such an estimation of the illuminant has come to be known as the albedo hypothesis. We combed through the various meanings which have been ascribed to this hypothesis. For example, some authors saw it as entailing that there was a causal relation between the illuminant estimate and lightness estimate so that the former was causally required for the latter. We dubbed this position *the causal view*. Others merely viewed the albedo hypothesis as specifying an invariant relationship between both estimates. For instance, for a surface with a given luminance, estimating that the illuminant is now lighter should lead to an estimation of the surface as darker than it was before. This invariant relationship was not necessarily seen as compatible with the causal view.

After reviewing the various construals of the albedo hypothesis, I suggested that we focus on the causal view. The latter can be cashed out in various ways, depending on whether the illuminant estimate is considered the sole factor necessary to compute the reflectance estimate (the *exclusivity hypothesis*) or whether the visual system has to know about other variables to do so (the *partial cause hypothesis*).

Various studies testing those hypotheses were presented and discussed. For instance, Kozaki and Noguchi performed two experiments on the subject. In all conditions, the luminance was kept levelled (i.e., the amount of light emitted from a test patch or background wall held constant) as the reflectance (of the patch or the background) and illuminant varied in compensating ways (so as to preserve luminance). Participants had to judge how light and bright the figure (or background) and illuminant were, respectively. The results were consistent with the existence of an invariant relationship between the estimate of the reflectance and the estimate of the illuminant. The authors argue that they were inconsistent with the albedo hypothesis understood as the causal view. But it is not clear that the results were relevant to the partial cause hypothesis, as opposed to being relevant only to the exclusivity hypothesis.

Although they were no definite conclusion, the very fact that this hypothesis is seen by foes and friends of reverse optics as a prerequisite of the viability of the framework and the fact that it has been tested quite extensively in numerous experiments put pressure on Chirimuuta's claim that there can be no unequivocal testable consequence of research programs in color constancy research.

In addition, Chirimuuta appears to claim that there can be no experimental comparison of, e.g., theories endorsing reverse optics and theories rejecting the framework, I showed that such experiments have been conducted. In one of Zaidi's experiments, four cups wrapped in paper were presented to participants, one of them having a differently achromatic color of paper. Two of the four cups were presented under a light illuminant, the other two under a darker one. Participants had to determine which cup was wrapped in the different paper. The authors modelled possible responses according to three models, one of which supposedly being a reverse optics model. They argued that the latter did not account for their participants' results as well as a model based on heuristics. However, we saw that their data actually underdetermine which model is correct. This issue, in addition with conceptual ones, prompted us to look elsewhere.

I thus presented the work of Olkkonen, Saarela and Allred which directly put to the test various models of color constancy. In their experiment, two displays, simulating either a lighter or darker illuminant, each contained a square of a particular lightness on a simple background. The displays were either presented simultaneously or with a delay. The latter case, the "joint" condition, allowed the determination of the effect of memory on color constancy. Participants had to choose the lightest square. The authors modelled the possible results according to three distinct models, two of them based on contrast coding and the last one being an example of reverse optics strategies. Their results were only accounted qualitatively and quantitatively by the reverse optics model. This work seems to present a counterexample to Chirimuuta's claim that models adhering to distinct research frameworks cannot be directly compared.

The last chapter was the most exploratory part of the present project. I there laid out what I hope can be a foundation for a methodological stance regarding color ontology. It is argued that research on color ontology could and perhaps ought to start with an examination of our best color constancy theory (or theories). Were the latter to be empirically successful, it could indeed be shown to represent justification for certain ontological claims about color.

The chapter started with a few examples of that methodology in action. As it is perhaps the most famous example, Maloney and Wandell's theory of linear models was the first theory considered. Remember that, according to this theory, the visual system decomposes the color signal into two distinct weighted sets of basis functions corresponding to an estimate of the illuminant and an estimate of the reflectance of the object, respectively[191]. Although it is usually assumed that the truth of Maloney and Wandell's theory would only represent justification for RP, it is not clear that this is the case. The theory may also represent justification for an ontology of color according to which colors are descriptors correlated with the objects' reflectance but also with other variables of the scene. If this is true, colors would be akin to mental markers, probably used by the visual system for bookkeeping purposes and not designed to track a physical property of surfaces (such as their SSR). Nevertheless, it was not clear in the end that this theory was a viable alternative as it seems to presuppose that Maloney and Wandell's algorithm is not correctly implemented by the visual system.

Foster's theory of relational color constancy was next considered. Remember that, according to this theory, the visual system tries to recover the chromatic relations standing between the objects of the scene, expressed as cone-excitation ratios within cone-channels,

---

[191] That is the case in flat-matte-diffuse conditions as we saw. In more complex scenes (such as natural conditions), the visual system will also have to determine whether haze, interreflections, shading, etc. are present.

rather than the SSR of surfaces. Whereas in the third chapter we tried to determine whether the theory could be modified so as to represent justification for RP, we here focused on the unadulterated theory. It was argued that the latter could motivate what I called *objectual color relationalism* (or OCR, for short). According to OCR, colors are relational properties of objects determined by the chromatic relations between objects in the scene or between objects and the chromatic average of the scene. For instance, an apple being red would consist in the apple reflecting more longer wavelengths than the other objects in the scene but much less short and middle wavelengths.

We then touched on Zaidi's approach to color constancy and his distinction between sensed and concluded color. Remember that the latter refers to two distinct ways of understanding the notion of color. When one speaks of the color of the maple leaf as changing when it is taken inside and put under a lightbulb, one is referring to the sensed color of the leaf. If I insist that the leaf has not changed color when the illuminant has changed, I am referring to the concluded color of the leaf. Finding out about sensed colors is argued to be a part of perception, whereas the processes responsible for determining concluded colors are thought to lie beyond perception. In that sense, Zaidi argues that color constancy is not part of our visual phenomenology but is a cognitive achievement. On the contrary, our phenomenology is one of inconstancy.

In addition, the notion of concluded color was linked with that of essential color. The essential color of an object corresponds to the cone quantum catches of the object under an equal-energy illuminant. Under such an illuminant, the color signal emitted from the object is proportional to its reflectance. Zaidi therefore argues that seeing an object in this condition counts as special in the sense that it defines the concluded color the object really has. For instance, if a leaf appears dark green under an equal-energy illuminant (that is, displays a dark green sensed color), then the leaf is dark green, notwithstanding its appearing differently (i.e., not dark green) under other illuminants.

When it comes to the ontological import of this account, it was concluded that sensed colors were mental properties which wrongly appear as relational properties of objects. Although objects appear to have sensed colors, the latter are a mere construction of our perceptual system and, depending on how we construe projectivism, can be seen as properties projected onto objects but are actually either properties of our mental states or are properties which are not instantiated at all. With regard to concluded colors, the first important point was that our ontology of concluded colors would depend on an account of sensed colors. Having a particular concluded color is indeed analyzed as displaying certain sensed colors under certain conditions. Subsequently, having a particular concluded color is being disposed to appear a certain way under certain conditions. In that sense, our theory of concluded color looks similar to dispositional theories of color.

However, this way of talking may be misleading. Dispositional theories of color usually state that it is the object that has the disposition. Yet it is at the very least unclear that this is a claim Zaidi wants to be associated with. It seems very cumbersome to say that an object being a certain concluded color is identical with its appearing to have a property (i.e., a sensed color) it does not really have. Nevertheless, it was not clear that there were resources in Zaidi's work which would be sufficient to resolve this issue.

In addition, it was unclear that our discussion of Zaidi's theory's ontological import fit the structure of our discussions of the two other theories. Whereas in the latter case we were able to identify a theory of constancy which could motivate certain claims about ontology, it is not clear that this is what we have done in the case of Zaidi. Barring further explanations on how the visual system determines the sensed colors it attributes to objects and how the perceptual-cognitive system determines the concluded color of objects on the basis of the sensed colors attributed to them, it is possible that what we have been doing is just discussing Zaidi's direct opinion on ontology.

The next part of the chapter was dedicated to a discussion of the way the proposed methodology sees the relationship between the theory of color constancy on the one hand and the ontological claims it may justify on the other[192]. Two options were canvassed. First, it is possible that the truth of the constancy theory's central claims entails certain ontological claims about color. Alternatively, the truth of those central claims may empirically justify the ontological claims. In the first case and only there is the relationship such that the ontological claims cannot fail to be true if the theory of color constancy is true.

Through an examination of the relationship between Maloney and Wandell's theory and reflectance physicalism, it was shown that the first option is a non-starter. Remember that the theory of linear models construes the output of the visual process responsible for color constancy as an attribution of a set of weighted basis functions to a visible surface. The process itself does not overtly refer to colors at any point and is argued to operate subpersonally. However, reflectance physicalism is about the colors we seem to see when we look, e.g., at fruits and vegetables. It is unclear how Hilbert and Byrne are supposed to bridge

---

[192] Remember that the caution exercised by the word "may" or by phrases such as "*could* represent justification" and so on is needed because it is not the existence of the theory that justifies ontological claims. Rather, it is the empirical confirmation of the central claims of the theory which would in due course perform such justification.

this gap and a few options were discussed. Nonetheless, note that the fact that we need an additional element to "bridge the gap" seems to be evidence that the theory of linear models on its own does not entail reflectance physicalism.

A few objections were then considered. First, it can be wondered why it is color constancy research, rather than any other part of color science, which is relevant to color ontology. Secondly, we surveyed Akins and Hahn's challenge, according to which it is at the very least unclear why we ought to focus on color vision when our primary focus is the nature of colors. Finally, *Revelation* was introduced as the thesis that the nature of colors is revealed to us in our usual color experiences, thesis which may very well be incompatible with the methodological proposal at issue.

To begin with, we saw that an astute reader may wonder about why our discussion is restricted to color constancy research when there are so many other theories out there in the field of color science. Those other theories may also have import for color ontology. We considered and rejected Davies' proposal, according to which color constancy is the central feature of color vision. In the end, it was clear that we ought to accept that other theories in color science besides those found in color constancy research could provide justification for particular ontological claims about color.

We also considered Akins and Hahn's challenge. The two authors argue that it is unclear why we ought to care about what color vision has to say when our primary focus is the ontology of colors. "Why not focus on colors themselves?", they say. There were two answers available to us. First, it could be claimed that, since we encounter colors through the exercise of color vision, it is expected that our account of the nature of colors will be able to explain why and how we come to have color experiences. A more pessimistic answer was that it is unclear what alternative there is. Barring an identification of color with SSR, and thus presupposing an answer to our research question, it is unclear how we could focus on colors themselves rather than our perception of it.

Finally, I addressed an objection to the effect that the proposed methodology is incompatible with the thesis known as *Revelation*. According to the latter, the nature of color is revealed to us in ordinary color experiences. We construed the thesis as the conjunction of two distinct ones, *Self-intimation* and *Infallibility*. The former states that, if it is in the nature of the colors that p, then p will seem true after careful reflection on color experience. *Infallibility* is essentially the converse and holds that, if after careful reflection on color experience, it seems to be in the nature of the colors that p, then it is in the nature of the colors that p.

Revelation is usually thought to entail that experience is the only source of evidence regarding the nature of colors. This is why it stands in conflict with the newly proposed methodology. According to the latter, empirical study of the phenomenon of color constancy could tell us about the nature of colors. But this possibility is precluded by *Revelation*. If the methodology at issue is to be viable, we have to reject *Revelation*.

I presented and endorsed Allen's criticism of the thesis. If *Revelation* is true, then, for any proposition p which is such that it is in the nature of the colors that p, it ought to be relatively obvious to *everyone* carefully reflecting on their own color experience that p. However, we found many deep disagreements concerning such propositions in the history of the field. As an example, I focused on the existence and extension of the unique-binary distinction. It was made clear that some hues which are thought unique (i.e., phenomenally uncomposed) by some researchers in color science or philosophy are considered binary by others. If *Revelation* were true, we ought not to find such deep disagreements.

Remaining issues

One aspect of the present methodological advice that has not been addressed in the last chapter concerns its relation towards the traditional methodology already in place in the field of color ontology. Such methodology notoriously includes *inter alia* the use of thought experiments, phenomenological reflection on color experiences and appeals to intuitions regarding the nature of the colors. For instance, Jackson's prime intuition is that red "is the property objects look to have when they look red" (Jackson, 1996). Alternatively, we are sometimes asked to pay attention to how our experience changes (or doesn't) when colored objects are brought indoors under a different illumination. For ease of discussion, let us call this methodology *Orthodoxy* and the methodology presented in the last chapter *New Proposal*. The question we now have to ask ourselves is whether *New Proposal* is supposed to replace *Orthodoxy* or whether it is to be seen as an element to be added to it[193].

At the outset, it seems that there is nothing in the present work which seems to speak to this issue in one way or another. Nowhere is it argued that there is something at fault with *Orthodoxy* or that some part of it should be deleted[194]. The minimal claim which is defended

---

[193] Of course, as *Orthodoxy* already appeals to color science as evidence for certain colors ontologies (evidenced by the Hilbert case), then *New Proposal* can be seen as propounding a change of focus that is of degree and not of kind. That is to say, *New Proposal* would involve the claim that we ought to focus *more* on color science and, in particular, on color constancy research and not that we should start focusing on it (which we have presumably already done). More on this below.

[194] Barring at least one exception. See below.

is that it at the very least cannot represent the whole picture methodologically speaking. This claim is consistent with *New Proposal* replacing or being added to *Orthodoxy*.

One possibility is thus that *New Proposal* will replace *Orthodoxy*. In this case, there is of course the issue that the best of our color constancy theory will not be able to answer certain questions we may have regarding color ontology. For instance, let us say that this theory tells us that colors are either relational properties or fictional ones. If there is no other source of evidence than color constancy theories[195], then our investigation into color ontology would just end up there. It would be known that colors are either relational or fictional properties. We would have narrowed down the options but would not have come up with a definite ontology of color.

Of course, some philosophers will be quite fine with that prospect. Others will not and it is open to them to argue that this is why we need the usual tools of *Orthodoxy*. The latter help us determine which of the remaining contenders is in all likelihood the correct one. In that case, although color constancy research (and perhaps color science as a whole) will represent one source of evidence with regard to color ontology, it will not be the only source of evidence. The likelihood that we will thus only narrow down the options is therefore reduced. Barring further arguments, it would seem advisable to opt for this weaker reading of *New Proposal*. Our tentative conclusion is therefore that color ontology research should evaluate the ontological claims motivated by the various main contenders in color constancy research but that this is not the end-all-be-all of our investigative endeavor regarding the nature of colors.

Naturally, it may be argued that philosophers in color ontology already pay attention to color constancy research. The best example is to be found in Hilbert and Byrne's work but other authors have also considered if not necessarily the research itself then at least the phenomenon of color constancy itself and how best to explain it (see, e.g., Cohen, 2004). So it may be claimed that *New Proposal* is actually not that new and there is nothing that is really added to the current methodology if we accept it.

First, it may be conceded that *New Proposal* does not bring anything categorically different to the table. It is a difference of degree, not of kind. Even though Hilbert and Byrne already believe that Maloney and Wandell's theory is relevant to the truth of reflectance physicalism, they believe that it is commonsense that first prompts us to accept color realism. However, the proposed methodology first puts the emphasis on the color constancy theory, in

---

[195] And maybe other fields in color science. As we saw in the last chapter, it is entirely possible that other areas of research besides color constancy research (such as ethology) may have implications for color ontology.

this case, Maloney and Wandell's, rather than on commonsense. Therefore, it asks us to focus *more* on the theory of constancy itself rather than on other potential sources of evidence, at least as a first step in research.

However, it has to be noted that the current methodology has actually ignored most of the field of color constancy research. Barring Chirimuuta's article we discussed in chapter 3 and 4, philosophers have concentrated on the one theory of Maloney and Wandell's theory, leaving out important theories, such as Foster's theory of cone-excitation ratios, and their potential consequences for color ontology. This is also evidenced by the fact that some philosophers act as though it was accepted by all the scientists involved that color constancy is to be understood phenomenologically, when there are a lot of dissenting voices in the field (see, e.g., Foster, 2003; Logvinenko et al., 2015; Valberg & Lange-Malecki, 1990; Zaidi, 2000). Consequently, *New Proposal* is also revisionary to current practice and perhaps not only as a matter of degree.

In addition, *New Proposal* is not compatible with certain claims found in the current research on color ontology. One main example is *Revelation*. Remember that, according to the latter, the nature of colors is presented to us in standard color experiences. We saw in the last chapter that, if *Revelation* is true, then *New Proposal* cannot be a viable methodology for color ontology and vice versa. *Revelation* entails that, if some claim C is in the nature of the colors, then upon careful consideration on our color experiences, we should come to the conclusion that C. However, if, say, Maloney and Wandell's theory is correct and justifies reflectance physicalism, then colors are (sets of) surface spectral reflectances. But this is not a conclusion we could come to upon consideration of our color experiences. So *New Proposal* does not entail that, in research on color ontology, 'anything goes'. Nevertheless, I would argue that rejecting *Revelation* is a small price to pay to endorse *New Proposal* and the various insights into color provided by color constancy theories.

We have arrived at the end of the present work. To conclude, I would like to point out that this project has been meant as an investigation of the methodological relationship there may be between color science, especially color constancy research, on the one hand and color ontology on the other. As has been said already (see, e.g., Hardin, 1988, 2014), the latter field has been more and more informed by color science in the last decades. However, we do not seem closer to an agreement than before we were thus informed. As Ross puts it, "the traditional divisions remain because, although constraints from visual science have recently been brought to bear on the problem of the constituting nature of color, this problem is not yet well constrained” (Ross, 2001, p. 43). It is my hope that the theoretical resources developed in this work help us see how to constrain the issue further.

Appendix A : linear models in color constancy research

In this appendix, I will explain the fundamental points of linear models as they are used in color constancy research, most notably in the work of Maloney and Wandell (Brainard et al., 1993; Maloney, 1986, 1999; Maloney & Wandell, 1986). This presentation may help the reader get a better grasp of why their theory has been used by Hilbert and Byrne to motivate reflectance physicalism if the material to be found in the main body of the present work is not enough to do so.

There are various points we have to cover in order to render justice to the theory of linear models. First, there are various assumptions which need to be in place in order for the algorithm described by the theory to function correctly. Secondly, the processes the visual system needs to perform to solve for constancy need to be described in a sequential way which allows for a clear understanding of the dynamics of the model. I address these points in turn below.

Maloney and Wandell's theory is set to describe a way to achieve color constancy in a very simplified environment. That environment is sometimes called a "flat-diffuse" environment in the sense that it is bidimensional and that the illuminant within it is diffuse so that there is no shadow or specular reflection present in the scene. As the complexity of the environment increases, the theory will have to be modified to accommodate the accompanying phenomena.

The theory also assumes that surface spectral reflectances can be approximated by a linear combination of a small number of basis functions. Algebraically, this gives us:

$$\mathrm{R}(\lambda) = \sum_{j=1}^{k} r_j \, B_j(\lambda)$$

Where $\mathrm{R}(\lambda)$ is the SSR of a surface, $\{B_j(\lambda)\}$ are the basis functions and r_j are the coefficients respectively assigned to those functions.

Similarly, the theory assumes that the same can be said for the scene illuminant's SPD:

$$L(\lambda) = \sum_{i=1}^{k} l_i \, E_i(\lambda)$$

Where L(λ)is the SPD of an illuminant, $\{E_i(\lambda)\}$ are the basis functions and l_i are the coefficients respectively assigned to those functions. Although the following explanations will also apply to this modelling of illuminants, we will restrict our attention to the case of reflectances.

The basis functions represent the fundamental dimensions along which an SSR can vary. For any wavelength λ along the visible spectrum, a basis function $B_j(\lambda)$ specifies how much light is reflected at that wavelength by a theoretical surface which would be correctly described by that basis function. Basis functions are fixed for all scenes and assumed known by the visual system. We can think of them as the building blocks from which more complex SSR can be constructed. Later, we shall see an example to make this notion clearer.

Linear models also assume that the number of basis functions needed to roughly approximate SSR for achieving color constancy is quite low. Of course, a model which presupposed an infinity of basis functions would perfectly replicate an overall reflectance. But an infinity, or even a large number, of such functions would be totally unmanageable by our visual system.

The coefficients or "weights" r_j stipulate how much of each basis function $B_j(\lambda)$ contributes to the overall reflectance R(λ). Simply put, they are numbers which scale the basis functions and thus indicate their preponderance in the overall reflectance. As the basis functions are known by the visual system and are fixed throughout scenes, knowledge of the coefficients represents complete knowledge of overall reflectance.

The equation above thus says that a reflectance is approximated by scaling each basis function (i.e., attributing a coefficient to it) and then summing all those functions together. Let us take a simple example to make all of it crystal clear. Let us say that we have three basis functions which are:

$B_1(\lambda)$: reflects mainly in the longer wavelengths (roughly 600 to 700 nm).

$B_2(\lambda)$: reflects mainly in the middle-length wavelengths (roughly 450 to 600 nm).

$B_3(\lambda)$: reflects mainly in the shorter wavelengths (roughly 380 to 450 nm).

Those three basis functions can be thought as the fundamental dimensions along which our real, complex reflectances can vary or, alternatively, as the building blocks out of which we can construct those reflectances.

Let us then imagine a surface whose reflectance we want to approximate. This surface reflects, say, 60% of longer wavelengths, 48% of middle-length wavelengths and only 12% of shorter wavelengths. We can thus express our real reflectance as:

$$\mathrm{R}(\lambda) = 0.6B_1(\lambda) + 0.48B_2(\lambda) + 0.12B_3(\lambda)$$

The coefficients in this case are $r_1 = 0.6$, $r_2 = 0.48$ and $r_3 = 0.12$. We have therefore reconstructed the real reflectance $\mathrm{R}(\lambda)$ using a sum of weighted basis functions.

Appendix B: A novel reason to pay attention to color constancy research

Introduction

In this section, I would like to spell out a reason we may have to consider color science in general and color constancy research in particular if what we are interested in are central arguments and claims in the field of color ontology. I will argue that a certain claim regarding the perceptual processes involved in color vision is required by many prominent arguments in color ontology. More specifically, the claim at issue concerns the form of the outputs generated by color vision processing according to many color constancy theories. The issue I will raise is that this claim entails that certain theories of constancy cannot be true, which is a problematic consequence of the arguments resting on the claim. In addition, strong evidence for one of those theories could make us rethink whether we want to accept those arguments in the first place. This strong relationship between color ontology and color constancy theories is in itself a reason to pay close attention to color constancy research as philosophers, or so I will argue.

To be clear, the goal of the following reasoning is not to defend any given theory in color ontology. Neither is it to defend a particular theory of visual processing. The aim is rather to argue that philosophers who are interested in using certain very prominent arguments and certain central claims made in color ontology ought to pay attention to what is happening in color science and more specifically in color constancy research. Of course, many arguments have been provided to that effect (see, as an important example, Hardin, 1988, 2014). However, to the best of my knowledge, the following reasoning is not one which has already been defended.

Color ontology and a very common assumption

Cohen's argument from perceptual variation and color contentfulness

The arguments we shall now look at are the argument from perceptual variation and the argument based on the existence of metamers, both of which are supposed to put pressure on color realism[196]. The strategy will be as follows. After presenting the arguments, I will try and show that they require a particular view of color experiences. In turn, the latter will be

---

[196] More specifically, the latter is usually directed at reflectance physicalism specifically. This feature will not be relevant for our current purposes.

argued to require an explanation and the best explanation we can come up with is that the perceptual processes responsible for color vision take a certain form. Finally, it will be argued that, if those processes actually take this form, then at least one particular theory of color constancy cannot be true, which is a problematic consequence of the arguments discussed.

The argument from perceptual variation is Jonathan Cohen's main tool for his attack on color realism and his defense of color relationalism (Cohen, 2004, 2009). In a nutshell, the argument is as follows. First, Cohen introduces the notion of perceptual variation when it comes to color and argues that it is widespread. In a second step, it is argued that color realism could be compatible with this phenomenon only by making some *ad hoc* stipulations. Since we ought to refrain from making such stipulations, by abductive reasoning, color relationalism is true. We will now dig deeper in order to get a very precise formulation of the argument. This will allow us to identify which claim it entails regarding color experiences.

To begin with, consider **figure 1**, which is thought to illustrate a case of simultaneous contrast[197]. If we were looking at a paper version of the figure, the two central strips would have the same SSR. However, we admittedly do not see those strips as having the same color. As far as I am concerned, the central strip on the left half of the figure appears greenish whereas the one on the right half appears mustard yellow. If we were to say that the two strips are really just one strip moved around (which you could do, armed with a pair of scissors), then we would conclude that the same strip looks greenish under certain viewing conditions (here, a certain background) and looks mustard yellow under other viewing conditions (here, a different background).

Figure 1. The two strips appear to have distinct colors, although they are in fact identical.

Cohen's first point is that those cases of simultaneous contrast are not isolated, rare occurrences. On the contrary, simultaneous contrast is ubiquitous. In addition, it represents

---

[197] If the figure does not "work" for her, the reader is invited to look at other examples of simultaneous contrast on the web.

only one example of variation in color perception and many other cases can be engineered depending on whether you compare distinct color experiences of the same individual (intrapersonal perceptual variation), of distinct individuals (interpersonal perceptual variation) or even of distinct species (interspecies perceptual variation). The bottom line in all three cases is the same. There are distinct color experiences of the same object, each of those involving the object appearing a certain color and the latter varying with the experience in question.

If colors were nonrelational properties of objects, Cohen argues, then we would have to determine in which one of those experiences the object appears with its real color. That is to say, we would have to choose one experience as involving the correct appearance of the color of the object. To understand this point, note that, if colors were nonrelational properties of objects, then which color an object has would not involve factors beyond the object itself. In the case of the colored strips above, their colors would be determined by the strips on their own and not also by factors such as the properties of the background. Since the object is the same object in all experiences, the factor determining its color, i.e., the object itself, does not vary and so the color appearing in all experiences should remain stable. Yet it does not.

We therefore have to choose in which of those experiences the color of the strip appears as it truly is. In other words, is the strip green or mustard yellow? Unfortunately, there is no principled criterion by which we could make that choice, or so Cohen argues. That is to say, there is no unarbitrary set of reasons which could point in the direction of one experience rather than the other. Choosing one of the two backgrounds as *the* background against which the strip ought to be viewed would seem entirely unmotivated and *ad hoc*, for instance. If we can avoid it, we should avoid *ad hoc* stipulation.

We thus face a hard choice. Whatever we say regarding one of the experiences of the strip, we have to say regarding the other. For any of the two experiences, we could either say that the color the object appears in it is a color of the object or that it is not. That is, we can either say that the object has that color or does not have it. But then we have to say the same for the other experience. Subsequently, either the object has both colors or it has neither. The choice we face is thus between color eliminativism and color relationalism.

Although Cohen favors relationalism, the important point here is that color realism is not an option. As construed here, color realism is committed *inter alia* to the thesis that colors are nonrelational properties. As we saw above, if colors are nonrelational properties, then, if an object is colored, then this object has one and only one color. But the argument from perceptual variation entails that this can be so only if we are ready to arbitrarily stipulate under which conditions objects appear with their true color. As far as its first step is concerned (i.e., the one before selecting between eliminativism and relationalism), this argument is thought to be a pretty strong attack on color realism.

In order to more precisely focus on some elements later on, let us display the argument in a more formal fashion. Given two color experiences A and B involving a single object O:

1. In A, O appears a particular color C (say, green), in B, O appears a particular color D (say, mustard yellow) and C is not identical with D (green is not identical with mustard yellow)
2. O is both C and D (the strip is both green and mustard yellow).
3. If color realism is true, O is either C or D but not both (the strip is either green or mustard yellow but not both).
4. Therefore, color realism is not true.

We can see here that premise 3 restates the notion that the truth of color realism entails that color properties are nonrelational and that objects can therefore have only one color. Premise 1 is the description of a particular case of perceptual variation. As an example, we focused on the case of the strips against distinct backgrounds. The color experiences mentioned in premise 1 constitute a reason to accept premise 2, according to Cohen. More specifically, Cohen believes that if we do not have a reason to suspect that one of the experiences is deficient in some regard, then we ought to accept them both as presenting color correctly[198]. This is a consequence of the claim that we should avoid *ad hoc* stipulations whenever possible[199].

At this point, it is important to point out that we are not to focus on whether the argument goes through or on the truth of any particular premise. I would rather like to concentrate on the requirements of the argument if it is to work. That is to say, if we admit that the argument goes through, what does it require to be true and, more specifically, what kind of account of color experience does it require to be true? As the latter notion is really only used in the first premise, this is the premise we shall now focus on. So we have to ask ourselves how we ought to understand the notion of color experience used in premise 1 in order for the latter to represent a very good reason to accept premise 2[200].

---

[198] As far as the option of rejecting them both is concerned, Cohen has a distinct argument to evade that possibility and thus counter eliminativism.

[199] Combined with the claim that eliminativism is inferior to relationalism, which we will not cover here.

[200] In conjunction with Cohen's argument to the effect that we have no principled criterion to choose one experience over the other as containing *the* correct appearance of the object.

In order for Cohen's argument to work, I take it that it is safe to assume that there must be some kind of conflict between the color experiences at issue and color realism. The argument asks us to choose between the two experiences the one which correctly portrays the object. But it could not ask us to do so if the experiences did not purport to tell us anything about the object. In turn, that they tell us something about an object means that they have a content that is assessable for truth or falsehood. That is, in order for the argument to work, color experiences must have a content[201].

However, it will not be enough that color experiences have a content. The color part of the experience must consist in an attribution of a property, namely, the color, to an object. In other words, in a color experience, an object appears *to be* a certain color. Let us call this requirement *color contentfulness* (or CC, for short). My point here is that, if Cohen's argument goes through, then CC must be true. If CC were not true, no conflict could emerge between the color experiences at issue and the statement of color realism. In that case, it would be possible for color experiences to have a content in such a way that the color aspects of the experience do not purport to attribute a property to an object. Subsequently, Cohen's argument can go through only if CC is true.

The argument from metamers and color contentfulness

Before we analyze this requirement further, I would like to make a detour and argue that CC is also a main ingredient of another central argument in the philosophy of color, namely, the argument from metamers. According to the latter, color realism is incompatible with the existence of metamers, roughly, surfaces with distinct SSRs which appear to have the same color. Since we have very good empirical reasons to believe that metamers exist, we have to give up color realism. I suggest that we now take a closer look to the argument in order to see whether, like the argument from perceptual variation, it presupposes CC.

Remember that metamerism is a perceptual phenomenon that results from the fact that we have but a small number of kinds of photoreceptors. Because we are trichromat, the color signal our eyes receive is ultimately reduced to a three-value signal, the cone signal. Consequently, if two color signals result in the same cone signal, that is, have the same three values of cone activations, they will be indistinguishable.

More specifically, light metamerism refers to the fact that lights with different SPDs will be indistinguishable as long as they result in the same color signal. In our case, the latter

---

[201] I will henceforth assume that, if an experience has content, then that content is the kind of thing which can be true or false. In other words, I will assume that experiences which have contents have veridicality conditions.

will be a three-value signal but it will vary according to the number of cones present in the species under consideration (e.g., five for pigeons which are pentachromats, two for most mammals). Light metamerism is a very well-documented phenomenon (Jameson & Hurvich, 1989; Kuehni, 2013; Luo, 2016; Shevell, 2003; Worthey, 1985).

The argument from metamers does not rest on light metamerism and we shall thus not consider this phenomenon any further. We will therefore focus on the notion of surface metamerism. Surface metamerism refers to the fact that two surfaces with different SSRs can be indistinguishable if the color signals reflected from their surfaces result in identical cone signals. However, if the viewing conditions change, the respective cone signals might then differ and the surfaces no longer be indistinguishable[202].

Whereas in the case of light metamerism there is no sense in talking about a scene or background (the lights are directly projected in the observers' eyes and thus they do not see anything but the lights), surface metamerism requires that the surfaces be perceived under the same viewing conditions and in particular against a single background under the same illuminant. If a surface S is in a scene illuminated by a given illuminant, another surface T in a scene illuminated by another illuminant (that is, a spectrally different one) and both S and T yield the same cone signal, they will not usually be indistinguishable. This is due to the fact that the visual system, in supposedly ascribing a color descriptor to a surface, takes into account the spectral properties of the neighboring surfaces in the scene (and, probably, other scene statistics, such as the chromatic scene average) as well as the surface under consideration. Two identical cone signals in differently illuminated scenes will thus not usually be indistinguishable.

To understand how the issue for color realism arises, let us consider a simple case of metamerism. You are looking at two pieces of textiles under artificial light, perhaps part of a t-shirt and part of a shirt. Both items appear identical in color as far as you are concerned and appear a very vivid shade of green. Unbeknownst to you, however, the dye which has been applied to the t-shirt consists of different pigments than the one of the shirt. That is to say, even though both pieces of textile appear vivid green, their SSRs are different.

However, according to reflectance physicalism (RP, for short), colors are SSRs. In that case, if two SSRs are distinct, then they are distinct colors and *vice versa*. But this seems to be contradicted by our textile example. The t-shirt and the shirt appear the same shade of green but their SSRs differ. As RP entails that this ought not to occur and we have strong empirical

---

[202] Surfaces which are indistinguishable under any illuminant are called isomers and have the same SSR.

reasons to believe that surface metamerism is a common phenomenon, we have to abandon RP. Surface metamerism seems to represent a counterexample to the central claim of RP.

Again, the point here is not to determine whether the argument is sound or whether its premises are true or false (but, for the interested reader, see Byrne & Hilbert, 2003, 2020; Hilbert, 1987; Kalderon, 2008). We want to concentrate on the requirements of the argument, more specifically, on what its premises tell us about color experience. To do so, I suggest we use a more formal rendering. If we stay with the features of the textile example described above, we can put it in the following way. Given a color experience A of an object O and another color experience B of another object R:

1. In A, O appears a particular color C (say, a particular shade of vivid green) and in B, R appears C.
2. If RP is true, if O and R have the same color, then O and R have the same SSR (if the shirt and the t-shirt are both vivid green, then they have the same SSR).
3. O and R do not have the same SSR (the pigment of the shirt and that of the t-shirt are spectrally different).
4. Therefore, if RP is true, O and R do not have the same color.
5. If O and R do not have the same color, then either O is not C or R is not C.
6. O is C and R is C (both the shirt and the t-shirt are vivid green).
7. Therefore, RP is false.

Premise 1 is the description of the particular case we have described above. Premise 2 is a consequence of the way we have construed RP here. Premise 3 is a stipulation and with premise 2 yields premise 4. The fifth premise merely represents the fact that if the shirt and the t-shirt do not have the same color, then only one of them can be vivid green. However, as far as we can tell, the evidence we have that the t-shirt is vivid green is the same kind of evidence we have in the case of the shirt. That is, premise 1 is a reason to accept premise 6. We thus ought to accept this last premise.

We can tell here again that if CC is required by the argument, it will be through the usage of the first premise that it will be so. So the question then comes down to how we are supposed to understand the color experiences mentioned in that premise in order for the argument to work. If we can show that the understanding of color experiences which was required for the argument from perceptual variation to work is the same which is needed here, then we will have demonstrated that CC is also a central presupposition of the argument from metamers.

In order to do so, we can look at the role which premise 1 is supposed to play in the argument. As we saw, it is supposed to motivate premise 6. Or, rather, it is supposed to do so in conjunction with the fact that we have no reason to select one of the two objects, O and R, as the one which ought to appear C. We should thus focus on the conditions required for there to be a need to perform that selection. To do so, I suggest we try and understand where the pull of the argument comes from.

As we understood RP, it entails that two objects with distinct reflectances have distinct colors. However, our color experiences in the textile example seem to be telling us that two such objects have the same color. So the choice is thus to either reject RP or claim that one of the experiences does not correctly portray the relevant object. In the latter case, either O or R would incorrectly appear as C. The pull of the argument comes from the fact that we have no reason to favor one color experience over the other. In a similar fashion to the argument from perceptual variation, there is no principled criterion by which we could make that choice.

I take it that it is clear by now that the argument from metamers requires the truth of CC. If the content of color experiences was not that an object (or some objects) appears *to be* a particular color, premise 1 could not motivate us to endorse premise 6. Appearing red means, in a color experience, appearing to be red. If this is true, then we indeed have to choose between characterizing one of the experiences (or both) as one in which an object appears to be a color which it in fact does not have or reject RP. The argument from metamers requires that CC be true.

Before we dig deeper into CC, I would like to suggest that, in all likelihood, it is found in many more arguments in the field of color ontology. In a nutshell, any argument which requires that appearing colored in a color experience just means appearing to have a particular color will automatically require CC. That is to say, the content of color experiences must be analyzed as "this object is C", where C is a color. Of course, CC will also be found in certain intuitions regarding color and particular conceptions of the problem of color. For instance, Jackson argues that our prime intuition about color is "simply that red is the property objects look to have when they look red"[203] (Jackson, 1996). Similarly, Boghossian and Velleman conceive of the problem of color in a way which essentially entails CC. It is best to quote them at length:

> "The role in which colors command attention, of course, is their role as the properties attributed to objects by a particular aspect of visual experience. They are the properties that objects appear to have when they look colored.

---

[203] And so on for the other colors.

> What philosophers want to know is whether the properties that objects thus appear to have are among the ones that they are generally agreed to have in reality." (Boghossian & Velleman, 1991, p. 68)

Here it is clearly stated that, if a color appears to one in one's experience, then it does so as a property of an object. In another article, the two authors go as far as saying that "to say that the phrase 'seeing a thing as red' describes a seeing event as having some adverbial property rather than as having the content that something is red […] does unacceptable violence to the concept of visual experience" (Boghossian & Velleman, 1989, p. 82). The right way to conceive of a red experience, they continue, is as an experience whose content is that "the object in question is red; and so the experience represents an object as having a property." A color experience is an experience with the content that an object *has* a color. CC therefore appears to run deep in central debates in the ontology of color.

The explanation of color contentfulness

The attributivity of color contentfulness

We are now in a position to try and determine whether CC entails any particular view of color vision, understood as the object of study of (part of) color science. Note that the point here is not to determine whether CC is true, false or which kinds of reason we could have for holding it. The point is rather to determine whether it implies particular claims regarding the perceptual processing responsible for color vision. The following should thus not be seen as an attack on CC or a defense of it.

At the outset, it has to be admitted that there does not appear to be any logical consequence of CC which directly concerns color vision as understood. There is no logical route from CC to a particular claim regarding color vision, its function or the perceptual processing responsible for it. A consequence of this is that there is no entailment from CC to the explanation we provide for color constancy. However, I would like to argue that, when we consider how to explain it, the best available explanation is one which involves a substantive claim regarding those subjects. Moreover, this claim may actually put pressure on certain theories of color constancy, which makes CC a heavy burden for whomever endorses it.

The first step, then, is to argue that, if CC is true, then it does not count as a brute fact. There is an explanation for CC. In other words, there is an explanation as to why color experiences are such that CC is true. Later on we shall restrict ourselves to only one of the properties color experiences have were CC to be true and to the explanation of color experiences having that property. However, for the time being, I believe that this point would be quite the uncontroversial one. CC would not count as the kind of facts which is an unexplainable element of our world. If it is true, it is true in virtue of other facts, and the latter represent an explanation as to why it is true.

I would like to argue that those facts concern the perceptual processes responsible for color vision. The phrase "perceptual processes responsible for color vision" refers here to the array of processes studied by vision scientists and which involve the use of the discovered chromatic pathways (such as the cones and their opponency due to the arrangement of retinal ganglion cells). For instance, when scientists propose an algorithm to account for our ability to be color constant or a mechanism explaining simultaneous contrast, they are studying those perceptual processes[204]. My proposal, then, is that CC could only be explained by those processes having certain properties.

More specifically, in order for CC to be true, at least some of those processes must have outputs which have a certain form, namely, one which mimics the form of the content of color experiences. If we consider CC carefully, we can see that it entails that the content of color experiences is of an attributive form. The content of a color experience is attributive iff that content is that an object (or a plurality of objects) has a particular property. To see this, note that CC claims that the content consists in an attribution of a property to an object. In the case of color experiences, that property is obviously thought to be a color but what I would like to focus on here is that, according to CC, the experience's content is attributive[205].

Note that, if CC entails that color experience are attributive, this is not a negligible entailment. In that case, it would run counter to some theories of color. Most notably, the attributivity of color experiences would be incompatible with adverbialism about color. According to the latter theory, colors are ways of seeing objects. For instance, even though we may think that we see an object as purple, our experience is better characterized as seeing an object *purply*. Purple is a way of seeing the object, not a property of the object itself. If this is true, then it appears that color experiences are not attributive after all.

It may similarly be claimed that the attributivity of color experiences would be incompatible with naïve realism. According to naïve realism, having a color experience consists in being directly presented with a color. The color is not represented in the experience, nor is it being seen indirectly through the mediation of a sense datum. The color

---

[204] Of course, if luminance vision is intertwined with color vision, then it is likely that many tasks which involve luminance vision will also involve color vision (Akins, 2014; Akins & Hahn, 2014). In that case, many if not all visual tasks would involve color vision. Even if that were true, the present argument would not be significantly affected.

[205] I will use "the content of the experience is attributive" and "the experience is attributive" as interchangeable in the present context.

itself is part of the experience. It might thus be thought that, since the experience involves the object itself and its properties, there is no attributive content in addition to the object itself. However, note that naïve realism can be conceived as an explanation of the fact that, in experiences, objects appear to us in particular ways. If one of those ways of appearing to us include appearing colored, then naïve realism may not be incompatible with CC, since the latter could be formulated in a way which implies that perceptual experiences do not have content. Naïve realism is therefore seen as an explanation of *inter alia* the fact that objects appear colored. In addition, naïve realists may want to claim that not only are the color and the object part of the experience, the color is presented as pertaining to the object. In that sense, the attributive component of the experience would be presented to the subject. If this is true, then this is another reason CC is not incompatible with naïve realism.

Color contentfulness and perceptual processing

When it comes to the perceptual processing of color vision, it can also readily be seen that some processes are argued to possess an output which is also attributive[206]. For instance, Maloney and Wandell argue that the output of the algorithmic process responsible for color constancy consists in an attribution of a particular set of weighted basis functions to a particular object (Maloney, 1986, 1999; Maloney & Wandell, 1986)[207]. Or if a process' output is that a particular edge in the scene viewed is of a certain degree, then that output would have an attributive form. Anytime the output of a perceptual process is that an object viewed has such-and-such property, the output is attributive.

A quite straightforward proposal at this point would be that the attributivity CC claims color experiences display is explained by the presence of attributivity in the perceptual processes responsible for color vision. Color experiences would be attributive because some parts of those processes are attributive. More specifically, they would be attributive because the outputs of certain of those processes are attributive. One way this explanation could be fleshed out would be to say that, in having a color experience, I am becoming aware of some of the outputs of those processes. The content of color experiences is none other than (a particular set of) the outputs of certain perceptual processes. Or one may say that those

---

[206] Even though it might be said that it is the *content* of the output that is of an attributive form, I shall restrict the term "content" to experiences only.

[207] Note that Maloney and Wandell's theory does not address the difficult question of what it means for a neural process to refer to a particular object and to say of that object that it has such-and-such property. As their work does not operate at this level (i.e., the level of neurological implementation), this matter still remains to be elucidated.

outputs are co-opted in the creation of the content of color experiences and that the attributivity of the latter is based on but not identical with the attributivity of the former.

This would be one way to explain CC[208]. If that explanation were true, CC would follow from color processing having a certain form and an awareness of the output of this processing. However, as a *reductio*, consider what would happen if the perceptual processes responsible for color vision did not display any attributivity. In that situation, there is no perceptual output in color vision of the form "object O has property Q". No object is categorized or described as having a certain property[209]. For instance, Maloney and Wandell's theory is incorrect and no weighted set of basis functions is attributed to an object.

In addition, suppose we accept the very common assumption that color experiences are instances of perceptual experiences. It is usually admitted that such experiences occur prior to any belief formation. In that way, it can be understood how they come to justify beliefs, and not the other way round. I believe that the cat is on the mat because I have a perceptual experience with the content that the cat is on the mat, and not *vice versa*. If this is true, then the properties of perceptual experiences do not depend on the properties of the resulting beliefs since I could have a perceptual experience without the corresponding belief. In turn, the eventual attributivity of perceptual experiences cannot be explained by the attributivity we commonly think beliefs display.

If perceptual processing does not display any attributivity and the attributivity of color experiences cannot be explained by that of beliefs, then I would argue that it is utterly mysterious how CC could be true. If CC entails that color experiences are attributive in the sense at issue, and this attributivity is not explained by a similar property of perceptual processes, it is entirely unexplainable. However, I argued that CC is not a brute fact. So it must have an explanation. The outputs of the perceptual processes responsible for color vision must therefore be attributive.

If perceptual processing is attributive, then this constrains the form our theories of it can take. In particular, I would like to argue that it constrains the form our theories of color constancy can take since, if there is attributivity in perceptual processing, then it is most likely to be found in the processes responsible for color constancy. It is indeed sometimes assumed that color constancy consists in the constancy of the color descriptor which is *attributed* to an object. If a particular descriptor is attributed to an object under all scenes regardless of their composition (regarding, e.g., the illumination at issue), then that attribution is determined only

---

[208] Or at the very least, the attributivity of CC.

[209] This is not to say that there are no processes that will do so, only that they will not be at a perceptual level.

by the unchanging properties of the object and is therefore constant[210] [211]. Here it is clear that the outputs of the processes responsible for color constancy would be attributive.

However, this is only one way of conceiving of color constancy. There are other ways which do not involve the claim that descriptors are attributed to objects in the course of achieving constancy. For instance, Zaidi's original theory, which we touched on in the last two chapters, states that color constancy is achieved through the direct identification of the same object across illumination changes (Zaidi, 1998, 2000). The output of the processes of color constancy is not that an object has a particular property which does not change with, say, illumination. On the contrary, it is that one object in a given scene is qualitatively identical with another object in another scene, where this identification is not mediated by the determination of the unchanging properties of either of the objects.

Remember that Zaidi's theory only had the visual system operate on chromaticities. The latter refer to individual color signals and thus conflate information pertaining to the illuminant and the surface at issue. For instance, a blue surface under a white light would emit the same chromaticity as a white surface under a blue light[212]. In the case of two scenes with the identical number of chromaticities involved, the visual system would just order the chromaticities of the two scenes according to their cone-opponent excitations and match the chromaticity that scores the lowest on those excitations to the chromaticity scoring the lowest in the other scene, the second lowest with the second lowest, etc.[213].

In the case in which the set of chromaticities in one scene is a subset of the set of the other scene (and the number of chromaticities is thus less in the first than in the second scene), the visual system will first perform a hypothetical matching of one chromaticity in one set with some other chromaticity in the second set. Having done so, it will be able to make an estimate of the shift in illuminant chromaticities[214]. Using this estimate will allow for the transformation of the other chromaticities into the relevant chromaticities under the second illuminant. The visual system will estimate the distance between this transformed

---

[210] Of course, this is consistent with the fact that, in situations in which we are not perfectly constant, the attribution does not only depend on the unchanging properties of the object. The attribution is so to speak "infected" by other variables. The important point here is that, even in cases of imperfect constancy, a descriptor is still attributed to an object.

[211] The term "unchanging" is used here instead of the more contentious "intrinsic" although they do not differ in spirit.

[212] Provided the SSR of the blue surface is described by a curb that is indistinguishable from the curb of the white illuminant's SPD and similarly for the white surface and blue illuminant.

[213] This is quite effective since there is a very strong correlation between the respective cone-opponent excitations of an object across illuminant changes. For more information on this, please see the discussion of Zaidi's theory in chapter 3.

[214] This shift being equivalent to the shift in chromaticities.

chromaticity and the corresponding actual chromaticity observed. The sum of the distances for all chromaticities will represent the error of the model. As this sum increases, so does the error of the model. The visual system will then iterate this procedure by matching the same chromaticity in the first set to another one in the second for all possible combinations. The model minimizing the error will be chosen as the one describing the correct equivalence of chromaticities.

I take it that it is clear that this procedure does not involve the determination of the unchanging properties of any object in the two scenes[215]. Even when it comes to the illuminants, the only conclusion the visual system reaches concerns the relative shifts in cone-opponent excitations due to the illuminant change. Zaidi's theory therefore does not appear to involve attributive processes.

However, it may be countered that Zaidi's algorithm does include attributivity. The first step actually consists in the attribution of a chromaticity to a given object. Moreover, the rest of the process will refer to objects by using the respective chromaticities which have been associated with them. So the process involves attributive elements. However, this reasoning assumes that we can be aware of not only the outputs of our perceptual processes (the end products) but also of the inputs and/or intermediate steps.

But this is not usually believed to be the case. Since Helmholtz (Peddie, 1926), it has usually been admitted that, if we are aware of any part of our perceptual processes, then we are aware of their conclusions. This is the traditional thinking behind the notion of unconscious inference. According to Helmholtz's original idea, all of the inferential steps of visual processes are parts of subpersonal functioning to which we are not privy. When the inference is completed, however, we become aware of the conclusion of the process, unconscious of the complexity of the steps leading up to it. If this is an accurate picture of what is going on, then the attributivity found in the processing steps prior to the outputs of Zaidi's algorithm cannot be used to explain the attributivity color experiences display according to CC.

If CC is true and it is explained by the attributivity of color constancy processes, then this at the very least puts pressure on Zaidi's theory, making it an unlikely candidate for the explanation of color constancy[216]. I take it that this is an unwelcome result of our investigation into some central arguments in color ontology. Since those arguments appear to

---

[215] For more information on this particularity of Zaidi's theory, please see chapter 5.

[216] Of course, there might be ways to modify or add to Zaidi's theory in a way that this becomes a non-issue. But this is exactly my point: the best explanation we can have of a central requirement of prominent arguments in color ontology (that is, CC) constrains what answers we are allowed to get in our research on color constancy.

require that our explanation of color constancy takes a certain form (namely, involves an attributive element as part of some of its outputs), this represents a very good argument for philosophers working on color ontology to pay close attention to the scientific research on color constancy and how it turns out.

References

Akins, K. (1996). Of sensory systems and the "aboutness" of mental states. *The Journal of Philosophy*, *93*(7), 337–372. https://doi.org/https://doi.org/10.2307/2941125

Akins, K. (2014). Black and white and colour. In R. Brown (Ed.), *Consciousness inside and out: Phenomenology, neuroscience, and the nature of experience* (pp. 173–223). Springer. https://doi.org/http://dx.doi.org/10.1007/978-94-007-6001-1_16

Akins, K., & Hahn, M. (2000). The peculiarity of color. In S. Davis (Ed.), *Color Perception: Philosophical, Psychological, Artistic, and Computational Perspectives*. Oxford University Press. https://www.sfu.ca/~kathleea/docs/PeculiarityOfColour.pdf

Akins, K., & Hahn, M. (2014a). Colour perception. In M. Matthen (Ed.), *The Oxford handbook of the philosophy of perception*. Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199600472.013.042

Akins, K., & Hahn, M. (2014b). More than mere colouring: The role of spectral information in human vision. *British Journal for the Philosophy of Science*, *65*(1), 125–171. https://doi.org/10.1093/bjps/axt060

Allen, K. (2011). Revelation and the nature of colour. *Dialectica*, *60*(2), 153–176. https://doi.org/10.1093/ACPROF:OSO/9780198755364.003.0007

Amano, K., & Foster, D. H. (2004). Colour constancy under simultaneous changes in surface position and illuminant. *Proceedings of the Royal Society B: Biological Sciences*, *271*(1555), 2319–2326. https://doi.org/10.1098/rspb.2004.2884

Arend, L., & Reeves, A. (1986). Simultaneous color constancy. *J. Opt. Soc. Am. A*, *3*(10). https://doi.org/http://dx.doi.org/10.1364/JOSAA.3.001743

Arend, L., Reeves, A., Schirillo, J., & Goldstein, R. (1991). Simultaneous color constancy: papers with diverse Munsell values. *J. Opt. Soc. Am. A*, *8*(4).

Aston, S., Radonjić, A., Brainard, D. H., & Hurlbert, A. C. (2019). Illumination discrimination for chromatically biased illuminations: Implications for color constancy. *Journal of Vision*, *19*(3). https://doi.org/10.1167/19.3.15

Averill, E. W. (1992). The relational nature of color. *The Philosophical Review*, *101*(3), 551–588. https://doi.org/https://doi.org/10.2307/2186057

Averill, E. W. (2005). Toward a projectivist account of color. *The Journal of Philosophy*, *102*(5), 217–234. https://doi.org/10.5840/JPHIL2005102525

Beare, A. C., & Siegel, M. H. (1967). Color name as a function of wavelength and instruction. *Perception & Psychophysics*, *2*(11), 521–527. https://link.springer.com/article/10.3758/BF03210259

Berryman, S. (2022). Ancient atomism. In E. Zalta & U. Nodelman (Eds.), *Stanford Encyclopedia of Philosophy*. https://plato.stanford.edu/archives/win2022/entries/atomism-ancient/

Boghossian, P., & Velleman, D. (1989). Colour as a secondary quality. *Mind*, *98*(389), 81–103. http://www.jstor.org/stable/2255062

Boghossian, P., & Velleman, D. (1991). Physicalist theories of color. *The Philosophical Review*, *100*(1), 67. https://doi.org/10.2307/2185515

Boyaci, H., Doerschner, K., & Maloney, L. T. (2004). Perceived surface color in binocularly viewed scenes with two light sources differing in chromaticity. *Journal of Vision*, *4*(9), 664–679. https://doi.org/10.1167/4.9.1

Boyaci, H., Maloney, L. T., & Hersh, S. (2003). The effect of perceived surface orientation on perceived surface albedo in binocularly viewed scenes. *Journal of Vision*, *3*, 541–553. https://doi.org/https://doi.org/10.1167/3.8.2

Brainard, D. H. (1998). Color constancy in the nearly natural image. 2. Achromatic loci. *Journal of the Optical Society of America. A, Optics, Image Science, and Vision*, *15*(2), 307–325. https://doi.org/10.1364/JOSAA.15.000307

Brainard, D. H., Brunt, W. A., & Speigle, J. M. (1997). Color constancy in the nearly natural image. I. Asymmetric matches. *Journal of the Optical Society of America. A, Optics, Image Science, and Vision*, *14*(9), 2091. https://doi.org/10.1364/JOSAA.14.002091

Brainard, D. H., Cottaris, N. P., & Radonjic, A. (2018). The perception of colour and material in naturalistic tasks. *Interface Focus*, *8*(4). https://doi.org/10.1098/rsfs.2018.0012

Brainard, D. H., Kraft, J. M., & Longère, P. (2003). Color constancy: Developing empirical tests of computational models. In R. Mausfeld & D. Heyer (Eds.), *Color perception: Mind and the Physical World*. Oxford University Press. https://doi.org/https://psycnet.apa.org/doi/10.1093/acprof:oso/9780198505006.003.0010

Brainard, D. H., & Maloney, L. T. (2011). Surface color perception and equivalent illumination models. *Journal of Vision*, *11*(5), 1–18. https://doi.org/10.1167/11.5.1

Brainard, D. H., Wandell, B. A., & Chichilnisky, E.-J. (1993). Color constancy: From physics to appearance. *Current Directions in Psychological Science*, 165–170. https://doi.org/https://psycnet.apa.org/doi/10.1111/1467-8721.ep10769003

Brill, M. H. (2003). "Color realism" shows a subjectivist' mode of thinking. *Behavioral and Brain Sciences*, *26*(1), 23–24. https://doi.org/10.1017/S0140525X03240012

Brown, R. O., & MacLeod, D. I. A. (1997). Color appearance depends on the variance of surround colors. *Current Biology : CB*, *7*(11), 844–849. https://doi.org/10.1016/S0960-9822(06)00372-1

Buchsbaum, G. (1980). A spatial processor model for object colour perception. *Journal of the Franklin Institute*, *310*(1), 1–26. https://doi.org/https://doi.org/10.1016/0016-0032(80)90058-7

Byrne, A., & Hilbert, D. (2004). Hardin, Tye, and Color Physicalism. *The Journal of Philosophy*, *101*(1), 37–43. https://doi.org/https://doi.org/10.5840/jphil2004101113

Byrne, A., & Hilbert, D. (2020a). *Objectivist Reductionism*.

Byrne, A., & Hilbert, D. R. (2003). Color realism and color science. *The Behavioral and Brain Sciences*, *26*(1), 3. https://doi.org/10.1017/S0140525X03000013

Byrne, A., & Hilbert, D. R. (2007). Color primitivism. *Erkenntnis*, *66*(1–2), 73–105. https://doi.org/10.1007/S10670-006-9028-8/METRICS

Byrne, A., & Hilbert, D. R. (2020b). Objectivist reductionism. *The Routledge Handbook of Philosophy of Colour*, 287–298. https://doi.org/10.4324/9781351048521-22

Chalmers, A. (2019). Atomism from the 17th to the 20th century. In *Stanford Encyclopedia of Philosophy*. https://plato.stanford.edu/archives/spr2019/entries/atomism-modern/

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), 563–582. https://doi.org/10.1080/00048400802215364

Chirimuuta, M. (2015). *Outside color*. The MIT Press. https://doi.org/https://doi.org/10.7551/mitpress/9780262029087.001.0001

Cohen, J. (2004). Color properties and color ascriptions: A relationalist manifesto. *The Philosophical Review*, *113*(4), 451. https://doi.org/http://dx.doi.org/10.1215/00318108-113-4-451

Cohen, J. (2009). *The Red and the Real: An Essay on Color Ontology*. Oxford University Press. https://doi.org/10.1093/ACPROF:OSO/9780199556168.001.0001

Cornelissen, F. W., & Brenner, E. (1995). Simultaneous colour constancy revisited: An analysis of viewing strategies. *Vision Res*, *35*(17), 2431–2448. https://doi.org/https://doi.org/10.1016/0042-6989(94)00318-1

Craven, B. J., & Foster, D. H. (1992). An operational approach to colour constancy. *Vision Research*, *32*(7), 1359–1366. https://doi.org/https://doi.org/10.1016/0042-6989(92)90228-B

Davies, W. (2018). Colour vision and seeing colours. *British Journal for the Philosophy of Science*, *69*(3), 657–690. https://doi.org/10.1093/bjps/axw026

Delahunt, P. B., & Brainard, D. H. (2004). Does human color constancy incorporate the statistical regularity of natural daylight? *Journal of Vision*, *4*(2), 57–81. https://doi.org/10.1167/4.2.1

D'zmura, M., & Lennie, M. (2001). Mechanisms of color constancy. *Journal of the Optical Society of America A*, *41*(10), 1662–1672. https://doi.org/https://doi.org/10.1364/josaa.3.001662

El-Bizri, N. (2005). A philosophical perspective on Alhazen's optics. *Arabic Sciences and Philosophy*, *15*(2), 189–218. https://doi.org/10.1017/S0957423905000172

El-Bizri, N. (2009). Ibn al-Haytham et le problème de la couleur. In *Oriens-Occidens: sciences, mathématiques et philosophie de l'antiquité à l'âge classique* (Vol. 7, pp. 201–226). https://www.academia.edu/35877004/Nader_El_Bizri_Probleme_des_Couleurs_Alhazen_pdf

Foster, D. H. (2003). Does colour constancy exist? *Trends in Cognitive Sciences*, *7*(10), 439–443. https://doi.org/10.1016/j.tics.2003.08.002

Foster, D. H. (2011). Color constancy. In *Vision Research* (Vol. 51, Issue 7, pp. 674–700). https://doi.org/10.1016/j.visres.2010.09.006

Foster, D. H., Amano, K., & Nascimento, S. M. C. (2005). Psychophysical estimates of the number of spectral-reflectance basis functions needed to reproduce natural scenes. *Journal of the Optical Society of America. A, Optics, Image Science, and Vision*, *22*(6), 1017–1022. https://doi.org/10.1364/JOSAA.22.001017

Foster, D. H., Craven, B. J., & Sale, E. R. H. (1992). Immediate colour constancy. *Ophthalmic and Physiological Optics*, *12*(2), 157–160. https://doi.org/10.1111/j.1475-1313.1992.tb00280.x

Foster, D. H., & Nascimento, S. M. C. (1994). Relational colour constancy from invariant cone-excitation ratios. *Proceedings of the Royal Society B: Biological Sciences*, *257*(1349), 115–121. https://doi.org/10.1098/rspb.1994.0103

Foster, D. H., Nascimento, S. M. C., Craven, B. J., Linnell, K. J., Cornelissen, F. W., & Brenner, E. (1997). Four issues concerning colour constancy and relational colour constancy. *Vision Research37*, *37*(10), 1341–1345. https://doi.org/https://doi.org/10.1016/S0042-6989(96)00285-4

Fuld, K., Werner, J. S., & Wooten, B. R. (1983). The possible elemental nature of brown. *Vision Research*, *23*(6), 631–637. https://doi.org/10.1016/0042-6989(83)90069-X

Fuld, K., Wooten, B. R., & Whalen, J. J. (1981). The elemental hues of short-wave and extraspectral lights. *Perception & Psychophysics*, *29*(4), 317–322. https://doi.org/https://doi.org/10.3758/bf03207340

Funt, B. v. (2003). Imprecise color constancy versus color realism. *Behavioral and Brain Sciences*, *26*(1), 29–30. https://doi.org/10.1017/S0140525X03300019

Gert, J. (2006). A realistic colour realism. *Australasian Journal of Philosophy*, *84*(4), 565–589. https://doi.org/10.1080/00048400601079128

Gert, J. (2008). What colors could not be: An argument for color primitivism. *The Journal of Philosophy*, *105*(3), 128–155. https://doi.org/http://dx.doi.org/10.2307/20620083

Gilchrist, A. (2006). *Seeing black and white*. Oxford University Press.

Gilchrist, A., Kossyfidis, C., Bonato St Peters College Tiziano Agostini, F., Cataliotti Ramapo College Xiaojun Li Lucent Technologies Branka Spehar, J., Annan, V., Economou, E., Cataliotti, J., Arend, L., Bruno, N., Emmerich, D., Gal-monte, A., Gerbino, W., Knau, H., Pessoa, L., Ross, W., Sedgwick, H., Todorovic, D., & Whittle, P. (1999). An anchoring theory of lightness perception. *Psychological Review*, *106*(4), 795–834. https://doi.org/https://psycnet.apa.org/doi/10.1037/0033-295X.106.4.795

Golz, J., & MacLeod, D. I. A. (2002). Influence of scene statistics on colour constancy. *Nature*, *415*, 637–640. https://doi.org/https://doi.org/10.1038/415637a

Gori, S. (2015). Afterimage. In L. Ming Ronnier (Ed.), *Encyclopedia of Color Science and Technology* (Springer Reference, pp. 1–4). Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-27851-8_271-1

Granzier, J. J. M., Brenner, E., & Smeets, J. B. J. (2009). Reliable identification by color under natural conditions. *Journal of Vision*, *9*(1), 39–39. https://doi.org/10.1167/9.1.39

Hardin, L. (1988). *Color for Philosophers: Unweaving the Rainbow*. Hackett Publishing Company.

Hardin, L. (2014). More Color Science for Philosophers. In D. Stokes, M. Matthen, & S. Biggs (Eds.), *Perception and Its Modalities*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199832798.001.0001

Hilbert, D. (1987). *Color and color perception: A study in anthropocentric realism*. CSLI Press. https://philpapers.org/rec/HILCAC

Hilbert, D. (2005). Color constancy and the complexity of color. *Philosophical Topics*, *33*(1), 141–158. https://doi.org/http://dx.doi.org/10.5840/philtopics20053315

Hilbert, D. (2008). Spraying color. In D. Hilbert & N. Last (Eds.), *Katharina Grosse: Atoms Inside Balloons*. University of Chicago Press.

Hilbert, D. (2012). Constancy, content and inference. In G. Hatfield & S. Allred (Eds.), *Visual Experience: Sensation, Cognition, and Constancy* (pp. 199–211). Oxford University Press. https://doi.org/http://dx.doi.org/10.5964/ejop.v10i1.731

Jackson, F. (1996). The primary quality view of color. *Noûs*, *30*, 199. https://doi.org/10.2307/2216244

Johnston, M. (1992). How to speak of the colors. *Philosophical Studies 1992 68:3*, *68*(3), 221–263. https://doi.org/10.1007/BF00694847

Joost, U., Lee, B. B., & Zaidi, Q. (2002). Lichtenberg's Letter to Goethe on "Färbige Schatten": G. C. Lichtenberg. *Color Research & Application*, *27*(4), 300–303. https://doi.org/10.1002/COL.10069

Kalderon, M. E. (2008). Metamerism, constancy, and knowing which. *Mind*, *117*(468), 931–968. https://doi.org/10.1093/mind/fzn043

Kelly, J. M. F., Daugirdiene, A., Kulikowski, J. J., & Murray, I. J. (2018). Chips in the sunshine: color constancy with real versus simulated Munsell chips under illuminants adjacent to the daylight locus. *Journal of the Optical Society of America A*, *35*(4), B100. https://doi.org/10.1364/josaa.35.00b100

Khang, B.-G., & Zaidi, Q. (2002). Cues and strategies for color constancy: perceptual scission, image junctions and transformational color matching. *Vision Research*, 211–226. https://doi.org/https://doi.org/10.1016/S0042-6989(01)00252-8

Kozaki, A., & Noguchi, K. (1976). The relationship between perceived surface lightness and perceived illumination: A manifestation of perceptual scission. *Psychol. Res*, *39*, 1–16. https://doi.org/https://doi.org/10.1007/BF00308942

Land, E. H. (1983). Recent advances in retinex theory and some implications for cortical computations: color vision and the natural image. *Proceedings of the National Academy of Sciences of the United States of America*, *80*(16), 5163. https://doi.org/10.1073/PNAS.80.16.5163

Land, E. H., & Mccann, J. J. (1971). Lightness and Retinex theory. *Journal of the Optical Society of America*, *61*(1), 1–11. https://doi.org/https://doi.org/10.1364/JOSA.61.000001

Levin, J. (2000). Dispositional theories of color and the claims of common sense. *Philosophical Studies*, *100*(2), 151–174. https://doi.org/10.1023/A:1018660204635

Ling, Y., & Hurlbert, A. (2008). Role of color memory in successive color constancy. *Journal of the Optical Society of America. A, Optics, Image Science, and Vision*, *25*(6), 1215. https://doi.org/10.1364/JOSAA.25.001215

Linnell, K. J., & Foster, D. H. (1996). Dependence of relational colour constancy on the extraction of a transient signal. *Perception*, *25*(2), 221–228. https://doi.org/10.1068/p250221

Logvinenko, A. D. (2013). Object-colour manifold. *International Journal of Computer Vision*, *101*(1), 143–160. https://doi.org/10.1007/s11263-012-0555-2

Logvinenko, A. D., & Demidenko, E. (2016). On counting metamers. *IEEE Transactions on Image Processing*, *25*(2), 770–775. https://doi.org/10.1109/TIP.2015.2504900

Logvinenko, A. D., Funt, B., Mirzaei, H., & Tokunaga, R. (2015). Rethinking colour constancy. *PLoS ONE*, *10*(9). https://doi.org/10.1371/journal.pone.0135029

Logvinenko, A. D., & Maloney, L. T. (2006). The proximity structure of chromatic surface colors and the impossibility of asymmetric lightness matching. *Perception & Psychophysics*, *68*(1), 76–83. https://doi.org/http://dx.doi.org/10.3758/BF03193657

Maloney, L. T. (1986). Evaluation of linear models of surface spectral reflectance with small numbers of parameters. *Journal of the Optical Society of America*, *3*(10). https://doi.org/https://doi.org/10.1364/josaa.3.001673

Maloney, L. T. (1999). Physics-based approaches to modeling surface color perception. In K. R. Gegenfurtner & L. T. Sharpe (Eds.), *Color vision: From genes to perception* (pp. 387–422).

Maloney, L. T., Boyaci, H., & Doerschner, K. (2005). Surface color perception as an inverse problem in biological vision. *Computational Imaging III*, *5674*, 15. https://doi.org/10.1117/12.598463

Maloney, L. T., & Wandell, B. A. (1986). Color constancy: a method for recovering surface spectral reflectance. *Journal of the Optical Society of America*, *3*, 29–33. https://doi.org/https://psycnet.apa.org/doi/10.1364/JOSAA.3.000029

McLaughlin, B. (2003). The place of color in nature. In *Colour Perception: Mind and the physical world* (pp. 475–502). Oxford University Press. https://doi.org/https://doi.org/10.1093/acprof:oso/9780198505006.003.0016

Morimoto, T., Kusuyama, T., Fukuda, K., & Uchikawa, K. (2021). Human color constancy based on the geometry of color distributions. *Journal of Vision*, *21*(3), 1–28. https://doi.org/10.1167/jov.21.3.7

Morimoto, T., & Smithson, H. E. (2018). Discrimination of spectral reflectance under environmental illumination. *Journal of the Optical Society of America A*, *35*(4), B244. https://doi.org/10.1364/josaa.35.00b244

Nascimento, S. M. C., de Almeida, V. M. N., Fiadeiro, P. T., & Foster, D. H. (2004). Minimum-variance cone-excitation ratios and the limits of relational color constancy. *Visual Neuroscience*, *21*(3), 337–340. https://doi.org/10.1017/S095252380421327X

Nascimento, S. M. C., Ferreira, F. P., & Foster, D. H. (2002). Statistics of spatial cone-excitation ratios in natural scenes. *Journal of the Optical Society of America A*, *19*(8), 1484. https://doi.org/10.1364/josaa.19.001484

Nascimento, S. M. C., & Foster, D. H. (1997). Detecting natural changes of cone-excitation ratios in simple and complex coloured images. *Proceedings of the Royal Society B: Biological Sciences*, *264*(1386), 1395–1402. https://doi.org/10.1098/rspb.1997.0194

Nascimento, S. M. C., & Foster, D. H. (2000). Relational color constancy in achromatic and isoluminant images. *Journal of the Optical Society of America A*, *17*(2), 225–231. https://doi.org/https://psycnet.apa.org/doi/10.1364/JOSAA.17.000225

Noguchi, K., & Kozaki, A. (1985). Perceptual scission of surface-lightness and illumination: An examination of the Gelb effect. *Psychological Research*, *47*, 19–25. https://doi.org/https://psycnet.apa.org/doi/10.1007/BF00309215

Olkkonen, M., & Allred, S. R. (2014). Short-term memory affects color perception in context. *PLoS ON*, *9*(1). https://doi.org/10.1371/journal.pone.0086488

Olkkonen, M., Saarela, T. P., & Allred, S. R. (2016). Perception-memory interactions reveal a computational strategy for perceptual constancy. *Journal of Vision*, *16*(3). https://doi.org/10.1167/16.3.38

Oxtoby, E. K., & Foster, D. H. (2005). Perceptual limits on low-dimensional models of Munsell reflectance spectra. *Perception*, *34*(8), 961–966. https://doi.org/10.1068/P5186

Padgham, C. A., & Saunders, J. E. (1975). *The perception of light and colour* (Bell). Bell. https://search.worldcat.org/title/954246317

Palmer, S. E. (1999). *Vision science : photons to phenomenology*. MIT Press.

Peddie, W. (1926). *Helmholtz's treatise on physiological optics*. Nature Publishing Group. https://doi.org/10.1038/118074a0

Putnam, H. (1974). How to think quantum-logically. *Synthese*, *29*(1), 55–61. https://doi.org/https://doi.org/10.1007/BF00484951

Putnam, H. (2005). A philosopher looks at quantum mechanics (again). *British Journal for the Philosophy of Science*, *56*(4), 615–634. https://doi.org/10.1093/bjps/axi135

Quinn, P. C., Rosano, J. L., & Wooten, B. R. (1988). Evidence that brown is not an elemental color. *Perception & Psychophysics*, *43*, 156–164. https://doi.org/https://doi.org/10.3758/bf03214193

Radonjić, A., & Brainard, D. H. (2016). The nature of instructional effects in color constancy. *Journal of Experimental Psychology: Human Perception and Performance*, *42*(6), 847–865. https://doi.org/10.1037/xhp0000184

Radonjić, A., Cottaris, N. P., & Brainard, D. H. (2015). Color constancy in a naturalistic, goal-directed task. *Journal of Vision*, *15*(13). https://doi.org/10.1167/15.13.3

Radonjić, A., & Gilchrist, A. L. (2021). Large depth effects on lightness in the absence of a large luminance range. *Journal of Experimental Psychology: Human Perception and Performance*, *47*(1), 134. https://doi.org/10.1037/XHP0000874

Reid, T. (1764). *An inquiry into the human mind : on the principles of common sense*. Edinburgh University Press.

Robilotto, R., & Zaidi, Q. (2004a). Limits of lightness identification for real objects under natural viewing conditions. *Journal of Vision*, *4*(9), 779–797. https://doi.org/10.1167/4.9.9

Robilotto, R., & Zaidi, Q. (2004b). Perceived transparency of neutral density filters across dissimilar backgrounds. *Journal of Vision*, *4*(3), 183–195. https://doi.org/10.1167/4.3.5

Robilotto, R., & Zaidi, Q. (2006). Lightness identification of patterned three-dimensional, real objects. *Journal of Vision*, *6*(1), 18–36. https://doi.org/10.1167/6.1.3

Ross, P. W. (2001). The location problem for color subjectivism. *Consciousness and Cognition*, *10*(1), 42–58. https://doi.org/10.1006/ccog.2000.0473

Rutherford, M. D., & Brainard, D. H. (2002). Lightness constancy: A direct test of the illumination-estimation hypothesis. *Psychological Science*, *13*(2), 142–149. https://doi.org/https://psycnet.apa.org/doi/10.1111/1467-9280.00426

Shevell, S. K., & Wei, J. (1998). Chromatic induction: border contrast or adaptation to surrounding light? *Vision Research*, *38*(11), 1561–1566. https://doi.org/10.1016/S0042-6989(98)00006-6

Slowik, E. (2023). Descartes' Physics. In E. N. Zalta & U. Nodelman (Eds.), *Stanford Encyclopedia of Philosophy*.

<https://plato.stanford.edu/archives/win2023/entries/descartes-physics/>

Smithson, H. E. (2005). Sensory, computational and cognitive components of human colour constancy. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *360*(1458), 1329–1346. https://doi.org/10.1098/rstb.2005.1633

Smithson, H., & Zaidi, Q. (2004). Colour constancy in context: Roles for local adaptation and levels of reference. *Journal of Vision*, *4*(9), 693–710. https://doi.org/10.1167/4.9.3

Soranzo, A. (2016). Simultaneous color contrast. *Encyclopedia of Color Science and Technology*, 1149–1152. https://doi.org/10.1007/978-1-4419-8071-7_268

Speigle, J. M., & Brainard, D. H. (1996). Luminosity thresholds: effects of test chromaticity and ambient illumination. *Journal of the Optical Society of America A*, *13*(3). https://doi.org/https://doi.org/10.1364/josaa.13.000436

Sternheim, C. E., & Boynton, R. M. (1966). Uniqueness of perceived hues investigated with a continuous judgmental technique. *Journal of Experimental Psychology*, *72*(5), 770–776. https://doi.org/10.1037/H0023739

Thompson, E. (1995). *Colour vision: a study in cognitive science and the philosophy of perception*. Routledge. https://www.routledge.com/Colour-Vision-A-study-in-cognitive-science-and-the-philosophy-of-perception/Thompson/p/book/9780415117968

Troost, J. M., & De Weert, C. M. M. (1991). Naming versus matching in color constancy. *Perception & Psychophysics*, *50*(6), 591–602. https://doi.org/https://doi.org/10.3758/bf03207545

Valberg, A., & Lange-Malecki, B. (1990). “Colour constancy” in Mondrian patterns: A partial cancellation of physical chromaticity shifts by simultaneous contrast. *Vision Research*, *30*(3), 371–380. https://doi.org/https://psycnet.apa.org/doi/10.1016/0042-6989(90)90079-Z

Weiss, D., Witzel, C., & Gegenfurtner, K. (2017). Determinants of colour constancy and the blue bias. *I-Perception*, *8*(6). https://doi.org/10.1177/2041669517739635

Westland, S., Shaw, J., & Owens, H. (2000). Colour statistics of natural and man-made surfaces. *Sensor Review*, *20*(1), 50–55. https://doi.org/10.1108/02602280010311392

Witzel, C., van Alphen, C., Godau, C., & O’Regan, J. K. (2016). Uncertainty of sensory signal explains variation of color constancy. *Journal of Vision*, *16*(15). https://doi.org/10.1167/16.15.8

Wright, W. (2019). The unique hues and the argument from phenomenal structure. *Philosophical Studies*, *176*(6), 1513–1533. https://doi.org/10.1007/s11098-018-1076-9

Wuerger, S., & Xiao, K. (2016). Color vision, opponent theory. In M. Ronnier Luo (Ed.), *Encyclopedia of Color Science and Technology* (pp. 413–418). Springer, New York, NY. https://doi.org/10.1007/978-1-4419-8071-7_92

Yang, J. N., & Maloney, L. T. (2001). Illuminant cues in surface color perception: tests of three candidate cues. *Journal of the Optical Society of America A*, *41*(10), 2581–2600. https://doi.org/https://psycnet.apa.org/doi/10.1016/S0042-6989(01)00143-2

Zaidi, Q. (1998). Identification of illuminant and object colors: heuristic-based algorithms. *Journal of the Optical Society of America*, *15*(7), 1767–1776. https://doi.org/https://doi.org/10.1364/josaa.15.001767

Zaidi, Q. (2000). Color constancy in a rough world. *Color Research and Application*, *26*(Suppl.), 192–200. https://doi.org/http://dx.doi.org/10.1002/1520-6378%282001%2926%3A1%2B%3C%3A%3AAID-COL41%3E3.0.CO%3B2-M

Zaidi, Q., & Bostic, M. (2008). Color strategies for object identification. *Vision Research*, *48*(26), 2673–2681. https://doi.org/10.1016/j.visres.2008.06.026

Zeki, S. (1983). Colour coding in the cerebral cortex: The reaction of cells in monkey visual cortex to wavelengths and colours. *Neuroscience*, *9*(4), 741–765. https://doi.org/10.1016/0306-4522(83)90265-8

Credits for the first page's picture: Kemp, W. (2014). *Apple study, acrylic on canvas* [Painting]. Retrieved from https://willkempartschool.com/how-to-acrylic-paint-light-and-shade-part-3/ with permission from the author.