BLACK AND WHITE AND COLOUR

NOTE TO READERS: This paper uses a good deal of the neurophysiological research to argue for its conclusions. Many readers will find the neuroscience excessive and for this I apologize in advance. (I direct anyone who requires *more* detail to the footnotes. However, I do not wish to meet you.) The problem is an old one. To dislodge a firmly entrenched view, detailed arguments are required. But the details—of neurophysiology in this case—can quickly become overwhelming and to no apparent end. So how do you argue in detail for a conclusion without the details? This is a problem to which I have not yet found a good answer. Hopefully the next generation of neurophilosophers will find a better solution than I have. In an effort to make the chain of reasoning clear, this paper has been divided into two parts, each with (I hope) clear conclusions. They can be read separately, or in two rounds, which will reduce reader fatigue.

Part I: Luminance Information

I. Introduction

What is it like to be a human rod achromat?

This is not one of the more scintillating questions ever asked by a philosopher, I take it. Indeed, if you are reading this paper and you *are* a rod achromat, you might find the question both uninteresting and condescending. No, rod achromats are normal folks, not 'alien life forms'. They do not have strange, ineffable, visual experiences that are unimaginable by the average trichromat. On the contrary.

As result of a genetic anomaly, the retina of the rod achromat has only one kind of working photoreceptor, the rods. It thus lacks the three (functional) cones normally found in the human trichromatic retina, the "daylight" system of vision. Despite this, the achromat's retina is very similar to our own at least vis-à-vis the physiology of the rod system. In the normal human trichromatic retina, there are two main pathways of luminance information, one from the rods and one from the cones. (This often comes as a surprise to people who have been taught since grade school that "rods are for 'black-and-white' and "cones are for colour'.) Both luminance systems use the same outgoing pathway from the retina, the magnocellular pathway. Under bright light (photopic) conditions, cone luminance signals are sent onward via the magnocelluar pathway. Under dim light (scotopic) conditions, the rods send signals via this same pathway and presumably make use of whatever cortical resources for luminance processing lie downstream. So these two luminance pathways, have a sort of 'timeshare' arrangement for use of the magnocellular pathway, based upon the lighting conditions. In the retina of the rod achromat, the rods still use the magnocellular pathway despite the fact that there are no cones present with which to share the 'real estate'. So, surprisingly, given all

the ways that things could have been, the rod achromat's retina has roughly the same arrangement as our own.[1] The net result is that the rod achromat suffers the worst visual deficit under daylight conditions. Because the rod system is highly sensitive to light, the rod achromat is photophobic. Just as you find it unpleasant to be awoken in the dead of night by a sudden flick of the light switch, the rod achromat finds daylight painful. Moreover, because the rod system sacrifices spatial resolution in order to gain this sensitivity, the rod achromat has decreased visual acuity during the day (relative to trichromatic daytime vision and rod achromatic night vision). Just as a myopic trichromat is aided by large print or magnified illustrations, and so too is the rod achromat.

So what *is* it like to be a rod achromat? Based upon experiments in comparative anatomy and on psychophysical tests for acuity, luminance contrast, and sensitivity (Sharpe et al, 1993; Greenlee et al. 1988; Hess et al. 1987), many people have concluded that the experience of the rod achromat is very much like our own vision when we see 'in black and white'. First, given that both our retina and the achromat's retina have very similar rod systems, what the achromat sees at night is probably very close to what the trichromat sees at night. Second, the night vision of a trichromat is commonly described as seeing "in black and white". Making allowance for the lack of illumination at night and the "fuzziness" of rod-based vision, what we see at night is very much what we see when we looking at a monochrome photograph or watching an old black and white movie. By transitivity, then, the rod achromat sees just what you, a slightly myopic trichromat would see were you to watch an old black and white film without your glasses and perhaps a photosensitive headache.[2] (And speaking only for myself, this *is* something I can imagine.)

---

[1] This is a simplified explanation of the matter. We now realize that there are at least three different pathways from the rods to the parasol ganglion cells of the magnocellular system. In a normal trichromat, the rods connect *directly* with cones, thereby influencing the polarization of the cones without any intermediary connections. The second and most well-known route from the rods to the parasol ganglion cells is via the rod bi-polar cells of the rod system. These rod bi-polar cells, which have a center-surround receptive field, have only ON centers. This is in contrast to the ON and OFF centers of the *cone* bi-polar cells. This is where the third route comes in. In effect, in the normal trichromatic retina, an OFF center signal in the rod system is *created* through the input of the *cone* cells (Cao et al 2010). Obviously, this is one connection which the rods in the achromat cannot make. Very recently it has been shown that rod system of the achromat is less sensitive to light than the rod system of the trichromat (Moskowitz et al. 2009). So it is possible that the rod achromat suffers a deficit in the function *of the rod system* as a result of the absence in cone input.

[2] Here for example is an quotation from Knut Nordby, a rod achromat, vision scientist and philosopher: "A first approximation, then, in explaining what my colour-less world is like, is to compare it to the visual experiences people with normal colour-vision have when viewing a black & white film in a cinema or when looking at good black & white photographic prints (good here meaning sharply focused, high contrast with a long grey-scale, as in crisp, high quality, glossy, technical prints)." Although this appears to contradict what I have said above, note that Nordby is presenting the achromat's highest levels of spatial acuity and contrast sensitivity which occur at the highest light levels for the scotopic (dark adapted) illumination. The achromat's spatial acuity falls off sharply with slightly brighter light. So much depends upon whether you are talking about the achromat's visual abilities under the best lighting conditions or under the sorts of conditions in which a normal trichromat would look at a photograph.

My central interest in such arguments is not in the nature of achromatic visual consciousness per se, but, more generally, what we believe about luminance vision and its informational characteristics. And I am interested in this because I think that our assumptions about luminance vision—what we see in virtue of luminance processing—handicaps our understanding of colour vision. It leaves us little room to contemplate some of the more interesting scientific results out of contemporary neuroscience about chromatic processing or 'colour' vision.

Over the last 15 or so years, evidence has steadily been accumulating that chromatic vision plays a much larger role in human sight than we had previously thought. The 'paradigm' view, stemming from Livingstone and Hubel's seminal research in the 1980's (Livingstone & Hubel 1984, 1987, 1988; Hubel & Livingstone 1987), holds that the luminance system is responsible for the lion's share of visual work in human vision. Without luminance information, Livingstone and Hubel said, we cannot see—or see clearly—most of the features of the visual world: where one object ends and another begins, the relations in depth that one object bears to another, the basic shape of the furniture of the visual world, movement of objects from one place to another. On the paradigm view, then, what I have called 'colour-for-colouring', chromatic information was primarily in aid of colour perception, for seeing the colours.[3] In the intervening years, however, it has emerged that chromatic processes are involved in almost all of the tasks Hubel and Livingstone originally attributed to luminance vision.[4] So both systems, luminance and chromatic, partake in *seeing the world*. Each system discerns, via its own mechanisms and strategies, sometimes together, sometimes in parallel, almost every property that the normal human trichromats can see.

Here is how the case of the rod achromat connects with this research. If you think that seeing via luminance vision alone is seeing "in black and white"—and here the rod achromat is the 'pure' case, a person whose visual experience is untainted by chromatic information if there is a pure case to be found[5]—then the last 15 years of research on

---

[3] Livingstone and Hubel's considered view was more complex. Their psychophysical results suggested that the parvocellular system carried low resolution temporal information but high resolution spatial vision, even though the parvocellular system carried chromatic information. So Livingstone and Hubel concluded, consonant with their neurophysiological finding, that chromatic information was used to discern the fine-grained details of object surfaces. Their psychophysical experiments suggested something quite different however: if you look at an equiluminant image, an image in which the intensity is exactly the same at every point, it is very difficult to see surface detail. So while the official view credited chromatic processing with discerning texture, etc., the de facto view focused upon chromatic information for the purposes of ascribing hue to external media and their surfaces.

[4] Cf. Kingdom 2003, Kingdom et al. 2004, Kingdom & Kasrai. 2006, den Auden et al. (2005), Shimono et al. 2009, Michna et al. 2006, Zaidi & Li 2006,

[5] Even here, the case is not completely clear. Although the rod achromat cannot garner information about wavelength through a comparison of different cone types, there is another avenue that might provide wavelength information: chromatic aberration. When white light passes between two media, one of which is more dense than the other, the direction of propagation chances as a function of wavelength. Shorter wavelength light is 'bent' more than longer wavelengths, and the net result is the spatial segregation of the light by wavelength. White light entering the lens of the eye at a single point is thus spread across the

chromatic vision will not look very exciting. If luminance vision makes available all of the information that you, a trichromat, would get from a black and white photograph—the shape of objects, their size, depth relations and *more*—what does chromatic vision add? Colour. Just as you can take an old photograph and hand colour it with pigment, chromatic vision adds colour to luminance-based information. This is the proverbial 'mental paint', the mental pigment added to black and white visual phenomenology. So if we actually know what we think we know about luminance vision, this research is of little interest.

In this paper, I hope to nudge our assumptions about luminance processing just a little, so that we can, as my husband likes to say, *consider the alternatives*. We want to be able to ask about the consequences of this research without assuming that we know, *a priori*, what those consequences will be. Assume that the visual brain uses both chromatic and luminance information *to see*. If we applied the same metaphysical and epistemic conclusions to both systems—if we applied our philosophical and scientific beliefs about luminance vision *mutatis mutandis* to chromatic vision—how would chromatic vision and colour experience look then? Making this a genuine question is the purpose of this paper. And at the end of this paper I offer a small sketch of an answer.

For the sake of clarity in what follows, note that there are two basic moves in the argument about the rod achromat's experience. First, the ordinary visual experience of the rod achromat is identified with the night vision of the trichromat. Second, the trichromat's visual experience at night is identified with seeing "in black and white". By transitivity, the achromat experiences the world in black and white. I suspect that both steps are incorrect, but the denial of the second assumption is what this paper needs for philosophical leverage. (In passing, it will come out why achromatic vision might be quite unlike trichromatic night vision.) Most of what follows attempts to show why human visual experience, given luminance information alone, is not like "seeing in black and white". If you claim that your night vision is like watching a black and white film, with appropriate adjustments for the dim ambient light, the loss of spatial resolution and the small patches of coloured light from your accumulated electronic gadgets in your bedroom, then you are wrong about your own experience. And by *modus tollens*, you are wrong about the achromat and his or her experience. On this view, then, our own luminance phenomenology is no less 'alien' than the achromat's. We can be quite mistaken about what we see and how we see it. (But as I said, it is *colour* that drives this paper. Luminance vision is just the means to that end.)

Part I of this paper (Sections I-IV) concerns luminance *information*. (This is the more technical half of the paper.) The conclusion is that a luminance system does not contain, nor can it derive, the sort of information one would normally get from looking at a

---

retina, the image smeared. This is a huge problem for any visual system with a lens to overcome. However, Forte et al. (2006) suggest that chromatic aberration could be used to gain colour information about objects. Objects that reflect light with different predominant wavelengths will have different extents to their spatial 'smear' on the retina—spatial differences that are large enough to be encoded and used. Knut Nordby also reported that chromatic aberration might be responsible for his ability to judge object colours at rates greater than chance.

monochrome photograph. This is one reason, then, why seeing at night is not like looking at a monochrome print, a reason based upon visual information. In Part II, the focus is on *phenomenology*. This second half is an extended argument against *rendering*, the transformation of luminance-derived representations/encodings into images. Whatever shapes our phenomenology, our conscious experience of the visual world, the application of mental paint is not a good explanation of what we see. The final section returns to chromatic vision, and how we should think about it given what has come before.

II. Luminance Processing in Human Vision

Like many of terms that are commonly used in science (e.g. 'electron' 'energy' 'power'), we often use them without remembering (if we ever knew) their explicit definitions. The term 'luminance' is no exception. 'Luminance', we all know, has something to do with 'amount of light'. But if you look up the definition of 'luminance' and you will find the following puzzling statement: luminance is the radiant intensity of light as filtered by the human photopic luminance function. In turn, if you look up "photopic luminance function", you will learn that this is *a model* (for the normal trichromatic human observer, adapted to photopic conditions) *of the probability that an individual photon will be absorbed, expressed as a function of the wavelength of the stimulus*. Sooo…after suitable digestion, it seems that 'luminance' refers to the amount of light, at each wavelength, your visual system will absorb. Uh huh. While technically correct, there is something unsatisfying about this definition. We want to know what a luminance system *does*, but this definition does not explain the *function of a luminance system* or distinguish it from the obvious alternative, the *chromatic system*. Nor does it tell us what luminance vision *represents*. It appears to say only that a luminance system absorbs photons—presumably, for the purpose of seeing (although that is not a part of the definition!). However unsatisfying this may seem, I've come to realize that the consequences of this "empty" definition are far-reaching.

Consider first that all photopigments on earth function in fundamentally the same way. Importantly, each photopigment responds to light across a *restricted range of wavelengths*—what is commonly known as the cone's spectral range. Moreover, each cone responds in a wavelength selective manner. Fig. 1 illustrates this principle: *relative to a fixed intensity or amplitude of light* (and that is the part to always remember), the graph illustrates the probability that a photon will be absorbed at each wavelength of light. Each photopigment has a "preferred" wavelength, the wavelength of light that is most likely to be absorbed, at a fixed intensity. For example, the L or long wavelength cone of the human retina has peak absorption at 570 nm. Different classes of cones (e.g. an L from an M cone) differ from one another one largely in virtue of the *spectral range* over which they respond. What differentiates *cones* from *rods* is primarily their sensitivity to light, how much light is required to affect a response.

Again, if one thinks in the old terms of "rods are for 'black and white'" and "cones are for colour", this might come as a surprise. Above, in the case of the achromat, we saw

that cones are responsible for luminance signals under photopic conditions. But does it follow that rods are equally capable of producing colour vision? Yes and no, replied the philosopher. Rods, just like cones, respond within a specific spectral window; they are wavelength sensitive in exactly the same way as cones. The primary difference between rods and cones, as I have said, is the energy required for photon absorption: rods require far less energy and are thus ideal for low light conditions; cones need a higher energy stimulus and thus are most useful during daylight conditions. In general, rods are not used for colour vision because under low lighting conditions, the photon catch is so low that the rods must be wired to maximize photon catch. They *sum* their inputs and thereby increase the signal-to-noise ratio. Wavelength discrimination—a prerequisite of colour vision—requires the *comparison* of two different cone signals. In the dark of night, it is thus the low photon catch of the rods that usually disqualifies rods for participation in colour vision. (This is somewhat counterintuitive given that the rods are *more sensitive* than cones—i.e. they have a lower photon catch despite their greater sensitivity. In fact, rods are so sensitive that in daylight rods reach their maximum response—saturation—and their signals can no longer be used for contrast processing.) Still, there is nothing in the function of rods that intrinsically precludes them from chromatic processing. Indeed, the old view that the cones "shut off" just as the rods "come on" (and vice versa) is not true. We now know that at dawn, dusk and under the light of a full moon (mesopic conditions), the human visual system contains a sub-population of rods that contribute to chromatic vision. The world continues to look coloured even when the photon absorption of the cones is compromised. (This is also true of horse vision.) Moreover, very recent experiments suggest that even under scotopic conditions (in the dark of night without starlight), rods feed into the S cone chromatic pathway—which explains why at night things appear quite "blue" to the trichromat (Cao et al. 2010). In other words, in dim lighting, we *do* see colours partially as a result of *rod* processing.[6] In retrospect, the fact that rods and cones work together under mesopic and scotopic conditions is not surprising but it is predictable only against a general understanding of photoreceptor function. This is also one reason why trichromatic night vision may differ from the rod achromat's visual experience.

So human vision has (at least) two major luminance systems, a photopic (bright light) system that depends upon cone input and a scotopic (low light) system that sums rod signals. Importantly, neither of these luminance systems—indeed *no biological luminance system*—encodes light *intensity*. This is in stark contrast (sorry) to a monochrome photograph in which the intensity value of light at each point in the photographic image is represented using the grayscale. (One source of confusion for the reader may be that monochrome images, which represent light intensity, are sometimes called 'luminance images' in computer science and artificial intelligence.)

This point is the flip-side of one familiar to all who study colour vision. We all know that a visual system with but one cone cannot discriminate between two stimuli that differ in wavelength alone. In the above graph of cone function, recall that the graph plots the

[6] More interesting still, is the Bourne parrot. These birds forage at dawn and dusk; they have a population of rods that appear to contribute to chromatic vision under dim lighting conditions and to luminance vision under night conditions (Lind & Kelber 2009).

probability that a photon will be absorbed against the wavelength of light relative to *a set intensity of light*. Alter the intensity of the light at a given wavelength, and the probability that a photon will be absorbed is altered as well. So photon absorption, for each cone, is a function of both wavelength and intensity. Receptor response does not indicate or provide information about either property independently of the other. As I said above, for colour researchers, this is a widely accepted—one might even say 'shop-worn'—fact: a visual system with a single receptor (or, what comes to the same thing, without the ability to compare different photoreceptor signals) is 'colour blind'. But what is sauce for the goose is sauce for the gander: the same moral holds, *mutatis mutandis*, for the *intensity* of the light stimulus. If a single photoreceptor conflates the wavelength and intensity of the light stimulus, then each photoreceptor is *intensity blind* as well. This is a fact we don't hear repeated nearly as often. Without a signal comparison between two different types of photoreceptors—and luminance systems do not make such comparisons by definition—luminance systems do not distinguish between intensity and wavelength.

A rather brilliant demonstration of this fact, and of the nature of luminance systems in general, is an art exhibit, "RGB" by the design firm Carnovsky.[7] One of the images in the exhibit is wallpaper covered in 19th century illustrations of various animals (some of which are even recognizable as the species which they represent) (Figure 2a). Each figure is printed in a single colour from the printer's standard three-colour palette, cyan, blue or magenta. Under natural illumination or any light source that approximates a uniform spectral power distribution, the coloured figures are clearly visible to the human trichromat. The exception is the yellow figures, which can be quite difficult to see when overlaid with other creatures. (In human vision, perceptual yellow can never be made as bright as the other colours.) The magic begins when the wallpaper is illuminated by coloured lights (Figures 2b, 3a and 3b). The animal figures now appear in shades of black to gray. Each colour of light selectively highlights or makes visible one of the three groups of animals; the other groups disappear entirely or become shadowy background figures. (In figuring out this exhibit, I found it helpful to pick out three figures, one in each colour, from the original wallpaper to compare under the different lighting conditions. So let the fox be our magenta figure, the alligator be cyan or blue, and the large cockroach at the top, be yellow. I know. What cockroach?! But it is there, overlaid upon the elephant).

There are two basic principles at play in this exhibit. First, the narrow-band filters, which "colour" the light by transmitting only a narrow range of wavelengths, render the viewer "more or less" monochromatic, as if he or she had a single-cone luminance system. In effect, the three colours of light mimic—*very roughly*—three possible luminance systems, each composed of a single cone. Recall from above, that each cone responds across only a limited spectrum of light. In vision, it is the "receiver" that limits the spectral range of visible light given the inherent properties of the photopigments: cones *absorb* only a narrow spectrum of light. In the Carnovsky exhibit, it is the "sender" that narrows the spectral range of vision: each illuminant *produces* light across only a narrow spectrum, red, blue or green. Despite being trichromatic, the normal observer is

[7] RGB is currently on exhibit in Berlin. For more images, go to http://www.carnovsky.com.

restricted to whatever the stimulus provides, namely the stimuli of red, blue or green light. So, the consequences are roughly the same. The experience is roughly that of a monochromatic system. (Although *only* roughly because the filtered lights do not mimic the selectivity of the three human cones.)

The second principle in play is that the filters are designed to enhance or diminish contrast, the *luminance* contrast of each set of figures, R, G, or B, against the white wall. When the wallpaper is bathed in red light, for example, only red light can be reflected back from the figures. Thus, the white wall, which normally reflects light across the entire spectrum, reflects only red light. (Indeed, under these conditions, everything that one can see is seen in virtue of *red* light.) To the trichromat, the wall appears red. And indeed, for the trichromat, every thing that is seen appears in shades of red from bright red to red-black. What then is visible against a bright red wall? A magenta figure (fox) will reflect a good amount of red light. It will 'appear' as a red fox against a red wall—or rather, it would if we could see it. A red fox on a red wall lacks sufficient contrast with its background to be seen and the same holds true for all of the magenta figures. It is the *blue* figures that will be most visible under the red light. A blue figure, which reflects very little red light under any lighting conditions, will now reflect very little light *at all*. The blue alligator will thus appear as a *black* figure against a red wall. Finally, the yellow figures will also be largely invisible. We are not told the spectral power distributions (SPDs) of the filtered lights. But suppose that the red light source contains some "yellow" light and the yellow pigment reflects, in addition to yellow light, a bit of red light as well. (Note that all natural pigments have an surface spectral reflectance that is continuous across the visible spectrum, so yellow objects usually reflect a good bit of both red and green light.) This lack of contrast would render the yellow figures all but invisible.

The two principles used by the Carnovsky exhibit are, in fact, two of the most important principles that have shaped the evolution of vision. This is why the RGB exhibit *works* on us. First, every known photopigment acts as a wavelength filter, responding to light as a function of both wavelength and intensity. Two different pigments may produce profoundly different levels of excitation in response to one and the same stimulus. In the evolution of any visual system, then, the type of photopigments/filters in place will directly effect what can be seen. Figure 4 (Gegenfurtner 2003) demonstrates the evolutionary outcome for the human case. It shows what effects the wavelength-selective absorption of our S, M, & L cones will have, individually, on the brightness of objects represented in the visual image. Here the scene is one of fruits and vegetables, most of which reflect light predominantly from the middle- to long-portion of the spectrum. (Most fruits and vegetables are yellow, orange and red. Blue vegetables are rare, not to mention somewhat unsettling.) Note, for example, the lightness of the images of the banana and the orange pepper in illustration (c) as compared to their images in illustration (e). Because the S cones are highly sensitive to 'blue' light, the reflected light of the yellow pepper and banana have very little effect. Similarly, the two oranges, directly below the banana appear black and dark grey respectively in image (e): they reflect very little light that an S cone will absorb. So the amount of 'effective' light reflected from an object depends upon the *spectral* facts of the environment—the *colour* of the light and of

the object—and on the sensitivity of the photoreceptor. This is why, in the RGB exhibit, the red, blue, and green lights produced such distinct effects.

Second, the primary concern of evolution in vision—i.e. what natural selection hinges upon—is not the filtering properties of cones *per se* but what the organism can *see*. Here, seeing requires that the organism be able to segregate the item of interest from its background via the responses of one or more cones. At bottom, *seeing requires contrast* along one dimension or another. Whether an object has positive or negative contrast with its background or whether the contrast arises from a difference in reflected wavelength or a difference in intensity (or both) is irrelevant. In the evolution of a luminance system what matters is whether *the cones produce luminance contrast relative to the spectral power distribution of the light source and the spectral reflectance profiles of the objects of interest.* In other words, in a luminance system, contrast occurs along a single dimension, the absorption of photons. But *whether and to what degree* contrast occurs depends in part about the spectral facts—the colour facts—of the world. Again, this is why the figures in the RGB exhibit were sometimes visible, sometimes not. When the red illuminant is trained upon the wallpaper, for example, visibility depends upon the difference in the reflection of red light, between the figures and their background. A magenta fox on a white wall lacks visual contrast because both reflect approximately the same amount of red light. But a blue elephant on a white wall will be highly visible—a dark figure on a red wall. (And not just because it is an elephant!)

It is worth emphasizing that the above two principles, taken together, yield highly unintuitive results for most people. It is unlikely that the Carnovsky exhibit emerged *de novo* without extensive experimentation. Nor could even a seasoned visual psychophysicist accurately predict what image would result from applying, to a natural scene, a filter that acted like a human cone. One reason for this is simply the complexity of the problem. Looking at a natural scene, we cannot eyeball each and every luminance and chromatic contrast in the scene and then imagine whether and to what degree each edge and coloured area will be transformed. (It is a safe bet that even Ansel Adams exhibited only his most successful guesses and his unintentional successes.)

More importantly, it is very easy to systematically misunderstand or to underestimate the effect on perception of these principles. Certainly, I catch myself doing this over and over again. For example, a filter limits the range of wavelengths that form an effective stimulus for the observer. But applying a filter is not like training a spotlight onto objects one colour at a time, onto all the blue animals or all the red ones, even if you also remember to imagine that the red or blue figures will appear in 'black and white'. Because *contrast* defines visibility, not absolute reflectance, there is no guarantee that we will see *any* of the objects that reflect light within your targeted (filter) range. Often, only the objects that do *not* reflect the 'effective' light will be visible qua dark objects against light backgrounds, as the RGB exhibit so nicely illustrated. Thinking in terms of a "spectral spotlight" is a very bad strategy for imagining what will be visible.

A second common pitfall of imagination (and again, I often do this) is to discount just how different one's experience would be as a 'luminance only' subject of visual

experience. One thinks: I know that luminance vision conflates intensity and wavelength but the world would still look *more or less right*. But it wouldn't. Photographic film and paper have been designed over a long history to eradicate this conflation, to represent light intensity alone. This is because a monochrome photograph, taken using a narrow band filter, looks 'unreal' to the trichromatic viewer. Add an 'L-cone' filter to your camera, and take a photograph of a vase of light blue hydrangeas and dark red roses. The hydrangeas will appear almost black, so dark that none of their individual petals or surface features are visible, while dark red roses are to be much brighter, their features clearly discernible. To the trichromat, the scene will look *plain weird* (here a technical philosophical term), not anything like a black and white photograph.

Let me try to draw together, from above, the most relevant points about luminance vision. Just as the name suggests, luminance vision *registers net photon absorption*. At the level of the ganglion cell, photon absorption is encoded in terms of the *relative rates* of photon catch, luminance contrast.. This measure, luminance contrast, forms the basic informational currency of all luminance processing. Importantly, because the absorption rate of each photopigment is wavelength sensitive, photopigment absorption does not correlate with light intensity (as it would if the photopigment were equally responsive across the range of visible light.) This is why the information used by a luminance system is nothing like the information represented in a monochrome photograph: in informational terms, a measure of *relative photon catch* is a far cry from the encoding of *absolute light intensity* at each place in the visual image. Finally, the fact that photopigments act like narrow-band 'colour' filters means that the informational characteristics of any luminance system are very much a function of the spectral facts of the world—the wavelength composition of the light source (Spectral Power Distribution), the surface spectral reflectance of objects, and so on. This fact also results in substantial loss of information, just as illuminating the Carnovsky wallpaper results in our inability to see much of what was visible under 'white' lighting. It results in the conflation of wavelength and intensity information; it renders ineffective *most* of the light within the normal human range of visible light ineffectual. In other words, luminance vision is a system that needs an informational helpmate, chromatic vision.

III. Luminance versus Chromatic Processing

It is beyond the score of this paper to explain chromatic processing and in fact, while we are making daily progress on the subject, I do not think we yet understand the nature of chromatic processing. Still, a few words should be said about the difference between chromatic and luminance processing, what chromatic information might add to a luminance system.

A central problem with a 'pure' luminance system is that it is rather primitive *qua* a source of information about the distal world. As light travels through the atmosphere, its interactions with bulk matter will affect, in specific and law-like ways, both its wavelength and intensity. For example, the intensity of light diminishes as it travels further and further from its source or as it makes contact with bulk matter; transparent

substances act as wavelength filters thus changing the wavelength composition of the emerging light; opaque objects cast shadows, resulting in areas of light of diminished intensity but of the same wavelength composition. And so on through the laws of optics applied to a natural environment. Thus, each dimension of the light stimulus, if encoded separately, can act *as an independent source of information about distal bulk matter*. By definition, a luminance system conflates these dimensions of light. So a luminance system cannot make use of the vast informational resources of the multi-dimensional stimulus that is light. Of necessity, a luminance system limits what can be perceived.

The above fact explains, very roughly, why all mammals that are active during the day have two cones and chromatic processing. Through genetic drift, the amino acid chains of existing photopigments are altered, an event that yields a new chain, with a new sensitivity to light. In effect, existing photopigments evolve into new ones, new "filters" of the spectral world. Very occasionally, on the order of .01 duplications per gene per million years, an entirely new photopigment is created. And *that* photopigment will itself become subject to mutation and drift. Over the eons, then, the existing photoreceptors of each species are altered and a species may gain (or loose) entirely new classes of photoreceptors. Often over time, a stable set of photoreceptors emerges, a set that maximizes *visual contrast* relative to the environment and the other constraints of the system.

Again, as my husband says, this makes for some interesting possibilities. How might you wire up these various cones now that you have them? The bedrock principle of any system of object vision is contrast encoding *of some kind of other* as we saw above. So all wiring that supports object vision must eventuate in the comparison of two spatially contiguous populations of neurons, in visual contrast. But within this confine, there are many possible permutations.

A *luminance* system uses a single type of cone in contrast processing or it sums together two or more types of receptor signals *prior to* contrast processing. More generally, a luminance system compares the photon absorption of two neural populations, both of which act as the same type of spectral filters. In human trichromatic human vision, for example, the central luminance channel takes M and L signals, and sums them prior to comparison. In this way, nature yields *a diverse array of natural luminance systems*, often within the same brain, each of which produces somewhat different measures of luminance contrast.[8] This is yet another reason why one should not equate a luminance system with one that tracks light intensity or intensity contrast. 'Colour-corrected' black and white photographic film was a great success because it made possible the representation of image intensity over a wide spectral range, while preserving good

[8] Again consider the RGB exhibit. In the monochromatic condition, the filtered light has a significant effect upon what you see (in effect, what type of luminance contrast is available). In the same way, there is a huge difference between the luminance contrast of the human photopic luminance function system and the human scotopic luminance function, or between the scotopic luminance functions of human and cat vision. When your cat goes downstairs for a little midnight snack, you do not need to get out of bed and turn on the lights for her. But when *you* head downstairs to raid the refrigerator, turning on the lights is a very good idea. This would be so even if, contrary to fact, the cortical contribution of the two systems, cat cortical luminance processing and human cortical luminance processing, were identical.

intensity contrast. *But a successful luminance system must make objects and their properties visible given a specific environment and the visual needs of the species.* And this goal need not coincide with the accurate representation of intensity information. (To see this point clearly, ask yourself this: which filtered light in the RGB exhibit—red, green or blue—is the correct one?)

A *chromatic* system compares the signals of two *different* populations of photoreceptors, two distinct wavelength filters. (Think of it this way. If you took a picture of one scene twice, first using Filter A and then using Filter B, you would have two very different monochrome images. Chromatic processing is like comparing two spatially adjacent portions of these two images.) In our own case, the blue-yellow chromatic channel compares the signals of S (blue) cones with the combined signals of the M and L cones which, when summed, act like a filter selectively responsive to 'yellow' light. The human system also contrasts mixed populations of M and L cones, each population of which contains a slightly different ratio of M to L cones. This strategy yields neural populations with slightly different filtering characteristics—a very 'intelligent' strategy for a system that requires fine-grained wavelength discrimination. Just as in luminance processing, there are as many different kinds of chromatic systems as there are species *and more*, for a species may have more than one chromatic system.

This brings us to the question of the informational contents of chromatic processing. It is no accident that the standard definition of 'chromatic' is 'about or concerning colour' and not something more helpful, for herein lies a very confusing tale. Chromatic cells are so-called because they respond to a spatial boundary defined by wavelength alone. This is the definition of colour vision, wavelength discrimination. But to maintain this definition, one must conveniently overlook the fact that any luminance cell also responds to one very specific chromatic stimulus. Add to this that 'chromatic' cells, which contrast different filter types, do not respond to only wavelength differences. They respond to wavelength differences *of a certain kind* and intensity differences *of a certain kind*. This fact is very easy to forget in our hurry to identify where, in human vision, wavelength and intensity information are *finally* encoded separately—as if the *point* of a chromatic system is to encode light wavelength and the *point* of a luminance system were to encode intensity. The real story—the 'there-is-no-simple-story' story—is that chromatic systems have proven useful because they respond to certain spectral discontinuities to which the luminance system does not. And even this markedly increases the processing power of the system as whole.

Let me give one example of how chromatic information increases the informational 'stash' of a visual system, here, via edge processing. Start with a standard colour photograph. We can illustrate the contributions of wavelength and intensity to any retinal image with two separate illustrations, an isochromatic image and an isoluminanct image (meaning, literally "all the same colour" and "all the same intensity").[9] These two

[9] Isochromatic and isoluminant displays are used because of the problems of *illustrating* a light array by *using* a light array. As light just is a wave, and hence has with both amplitude and intensity, we cannot make a *display* with only one of those dimensions. Rather, the trick, for both of these displays, is to use stimuli that are *perceptually indiscernible to people* along those two dimensions. In the first illustration,

dimensions of light are objectively present in any retinal image. So there are two sources of potential information that, at least on visual inspection, appear quite distinct (Figure 5). For example in the isoluminant image, you can see that all of the shadows that so nicely define shape in the original photo are lost, but they are still present in the isochromatic (monochrome) image. What would a specific chromatic system paired with a specific luminance system make of these objective facts? What edges would each system pick out? The same edges or different ones?

In an experiment by Hansen and Gegenfurtner (2009), 700 images of the natural and artificial world were examined to find out what the human visual system can see when it looks for edges defined by chromatic and luminance contrast. The image in Figure 6(a) depicts the original image or what they called the 'Input Image', a full colour image; the second image, 6 (b), shows the 'Luminance Image' or the total photon catch by the system, calculated by adding together the absorption of the S, M and L cones; Figure 6(c) shows the M-L Image, namely the contrast between the responses of the L and M cones at each point in the image (using a continuum between red and gray to illustrate the relative responses). Finally (e) and (f) show the edges that can be computed from these two types of contrast data. Here you can see that the edges determined using luminance contrast were often distinct from—and different strengths than—the chromatic edges. Obviously an image of red fruit against green leaves is the best possible subject matter for showing the rewards of adding red-green (M-L) contrast to luminance contrast. But importantly, when joint edge histograms were computed, the result turned out to be a general one. *Chromatic edges and luminance edges are statistically independent of one another in the visual image*. This means that, potentially, there is a huge gain in information whenever an M-L chromatic system is added to an M + L + S luminance system: new information about edges.

The RGB exhibit also gives a clear example of how our own chromatic and luminance systems work in tandem to see the world better. Look at the tri-color wallpaper under the red light and you immediately notice the loss of *colour*. You cannot see the creatures as cyan, magenta or yellow. But in terms of information, we have lost more than that. Recall that when we view the wallpaper under the red light, the red fox no longer stands out from its red background. So, with only one cone, sensitive to red light, we would loose *chromatic contrast*. As a monochromat, with only luminance information, we would also loose *intensity contrast*. If you take the Carnovsky wallpaper and render it as a monochrome photograph, you can see clearly that the fox is much darker than the white background (Figure 7). All of the creatures are. In fact, they are rendered in coloured pigments of roughly equivalent intensity. In the trichromatic system, somehow or other, this information seems to be present at some level of processing, even if it is not always available at the personal level. (If you have ever tried to sort paint samples on multiple chip cards according to whether they differ in hue or simply differ in saturation, you will

---

the light intensity has been made uniform across the array (relative to human luminance sensitivity): every pixel has the same *perceived intensity* but differs in wavelength across the display. In the second illustration, there is no perceptual difference in the wavelength across the array, but there is a difference in intensity. It is called 'isochromatic" because, to the human observer, there are no *perceived differences in wavelength*.

see what I mean. Inevitably, you must ask the paint salesperson to know what system of ordering has been used, if any.) Exactly how the human trichromatic system extracts this luminance information is an open question. Certainly by combining the M and L signals, the human luminance system increases its share of intensity information over the spectral range of these cells. But the important point here is that chromatic information allows our visual system to discern *more of the world* than its colours. It gains basic visual information in the form of chromatic and luminance contrast, in virtue of which other properties can be discerned. Finally, note how difficult it is to articulate what is lost, when the light turns from 'white' to red, except in terms of the *features of the world* that we can no longer see. We say: 'I can see only *red* and the *fox* has disappeared as well.' Without knowledge of what *information* is missing from normal trichromatic processing, this is as far as we get—unless we latch onto a metaphor, 'seeing in black and white', in order to 'explain' our visual phenomenology.

All of the above arguments have been in the service of one conclusion: the information to which a luminance system has access, luminance contrast, is far less useful than the information, absolute intensity, contained in a black and white photograph. But what about further downstream? Does anything like an intensity value for each 'pixel' of the visual image emerge from later processing? Here the answer, at least in neuroscientific circles, is unequivocal: no. As the reader will know, the vision sciences do not know exactly how luminance contrast information is used in humans. So it might seem surprising that a univocal answer would be given. A general agreement about the broad outlines of the luminance processing rules out this possibility however. Luminance processing begins with the encoding of luminance contrast signals of the retina, which are then 'massaged' to yield the multiple visual capacities/representations of luminance vision. Importantly, every property discerned is encoded, often implicitly, in terms of the dimensions of the relevant property space (or space of properties, given that encoding does not usually occur for single properties). Consider the problem of how to guide the hand visually, such that its grasp correctly aligns with say, the branch of a tree. Because grasping is a task that will be repeated over and over again, the visual system must represent (or systematically embody) the spatial relations between two arbitrary points within a prescribed area of the world, roughly that area which is both within arm's reach and can be seen. Such a mechanism defines or encodes the spatial relations between the opened hand and the object-to-be-grasped. In the visual domain, the visually-guided hand moves through a space circumscribed by the visual field of view, and that, in turn, is dependent upon the eye, neck and trunk co-ordinates. In the motor domain, the same space may be described by the range of motion possible, given the triad of the shoulder, elbow and wrist joints. So peri-personal visual space, encoded for hand guidance, will bear little relation to the space of luminance encoding. [10] The representational space of properties discerned *from* luminance contrast information is not the representational space *of* luminance contrast itself. The general story is, then, that we discern multiple properties of the visual environment using luminance information, and each of these

[10] Although we think of luminance as a two-dimensional space—'more or less' absorption—the actual encoding of luminance contrast is more complex. At the retina, luminance contrast is encoded in two separate spaces, positive and negative, in order to increase the signal-to-noise ratio, a fact has repercussions downstream.

resultant properties is represented in a space appropriate to it. These are not the dimensions of an intensity encoding.

The above story about luminance vision does not leave any room for a subsequent encoding of 'intensity space'…well, except perhaps for one, for there is one very reasonable response to the above criticisms of luminance vision as 'black and white' vision. 'Look, my dear, I never claimed that luminance *information* is like looking at a black and white photograph. That is obviously a mistake. Rather, at night, we see in 'black and white' in the following sense. Via luminance information we perceive the numerous properties of a visual scene, one of which is *albedo*—the surface reflectance of an object. When we look at an object, we see it as light or dark. It is *this* perceptual experience, the perception of albedo plus whatever other properties of the visual scene discerned via luminance information, that give us an experience much like looking at black and white photograph.' This is certainly a much more sophisticated version of the view we have been considering and requires close attention

IV. The Computational Problem of Determining Albedo

Although the reader may be unfamiliar with the term 'albedo', we are all familiar with the concept. Coal is dark, chalk is light and this is the case whether the chalk and coal are viewed in indoors in dim light outside in the sun. Surface reflectance or albedo is a constant property of an object, then, for no matter how we illuminate a surface, with a 20-Watt bulb or a theatre spotlight, the same percentage of light is always reflected from an object's surface. For the purpose of clarity, note that vision researchers have a strict use of the terminology for albedo perception and related problems. Albedo is the *physical property* of objects, the proportion of incident light reflected. 'Lightness' is *perceived albedo*, how albedo appears to the observer. Thus we perceive coal, potting soil, and the best kind of chocolate as *dark*; copier paper, pastry flour, and bleached cotton are *light* substances. The term "brightness" is reserved for our perception of the intensity of light or of a light source. When you open the curtains, the sun shining in appears blindingly *bright*; when you enter the house on a sunny day, the interior can appear very *dim*.

The diagram in Figure 8 illustrates the problem of albedo as it is standardly understood in the sciences of vision. An illuminant (light) shines upon a three dimensional object. The intensity of light at each point of the object's surface—the object's *illuminance*—is a function of the brightness of the light source and the particular shape of the object. (A memory trick for the reader: Illuminance is how the object is *illuminated*.) Each object, in virtue of its surface qualities, absorbs and reflects a certain percentage of the light cast by the illuminant, a property known as *surface reflectance* or *albedo.* Thus, the total light reflected from the object is a function of both the intensity of the illumination at each point on the object's surface and the percentage of that light which is absorbed. Unfortunately—particularly so in the context of this paper—the total reflected light is usually called the *luminance* of the scene. In this context then, "luminance" is a measure of the *stimulus*, of the intensity of light at each point of the retinal image prior to absorption by the cones. In computer science, it is also common to explain the problem

of albedo as one of image "decomposition". So, to solve the problem of albedo, we start with the 'Luminance Image' which represents the intensity of the reflected image, which is then "decomposed" into "layers" or into representations of the individual properties of the scene that have contributed to the Luminance Image. This is why in Figure 8, the problem of albedo is illustrated using two layers, the 'Reflectance Image' and the 'Illuminance image'. (In order not to fall into terminological chaos, I am going to use the term 'luminance' in only the physiologist's sense, and talk instead about 'intensity' or an 'Intensity Image' when referring to monochrome images, an image that depicts light intensity using a grayscale.)

More formally, the problem of albedo is usually represented as follows. Let I(x,y) equal the light intensity for any arbitrary point in the Intensity Image, R(x,y) be the value of that point the Reflectance image, and E(x,y) its value in the Illuminance image. Then:

$$I(x,y) = R(x,y)\, E(x,y)$$

Note that for any value of I(x, y), no unique value of R(x,y) can be computed without prior knowledge of the value of E(x,y). If you do not know the properties of light source, an infinite combination of Reflectance and Illuminance values are possible. Thus we cannot determine the contribution of an object's surface reflectance to the Intensity Image. The problem of albedo is undecidable in the mathematical sense. Of course, lack of decidability is not a good property for a computational problem. But the human visual system meets this same hurdle when it discerns almost any visual property that humans can see. What is required is background knowledge of some sort, be that scene statistics that are acquired during visual development or 'assumptions' that are hard-wired into human vision in some other way.[11]

One thing of which we can be reasonably certain is that a 'pure' luminance system is very unlikely to perceive albedo. First, from the last section, we know that a luminance system does not start with an encoding of the Intensity Image, information about the intensity values at each point in the visual image. The retina is presented with a physical stimulus, a spatially arrangement of light each ray of which possesses all of the dimensions of light. These are the properties of the stimulus prior to transduction. The first layer of 'encoding'—or perhaps a better term might be 'registration'—is merely photon absorption. The second layer of registration is luminance contrast. This is important because it complicates an already exceedingly complex computational problem. Light of a certain intensity and wavelength illuminates an object. That object then reflects and absorbs a certain percentage of light depending upon its surface properties; what sort of light (its wavelength composition) is reflected and its intensity will depend upon its surface pigmentation (i.e. the object acts as a wavelength dependent filter). The total reflected light is then absorbed by a luminance system, which itself acts as a wavelength-dependent filter. What emerges, at the other end of the retina, is luminance contrast. There are thus two points in the causal chain that conflate intensity with wavelength information, surface reflectance and photopigment absorption. Yet the

[11] For an excellent overview of current research on albedo perception, see Kingdom in press.

task of albedo perception is to figure out whether the object is 'light' or 'dark'—the percentage of total light (intensity) that the object reflects. Yet by definition, the luminance system does not have a *chromatic* system that might help it to resolve these conflations, to encode intensity independently of wavelength. The net result is that a rod achromat or a 'pure' luminance system cannot perceive *albedo*, if by that one means an accurate representation of surface reflectance, an objective property of objects.

Still, the rod achromat is not without some resources for classifying the surfaces of objects. If you ask a rod achromat whether he or she sees objects as light or dark, the answer will be 'yes'. Assuming retinal adaptation for light intensity, the rod achromat can make judgments about an object's surface 'luminance' reflectance, the percentage of light reflected by an object as-filtered-by-the-rods-of-the-achromat. This is a dispositional property relative to the filtering properties of the rod photopigments. And this primitive system will work, i.e. it will track a stable property of the environment, much of the time. As long as the light source does not change 'colour', i.e. as long the light has a stable spectral power distribution of the light over time, the achromat will make consistent judgments. But change the spectral power distribution of the illuminant, and the rod achromat will no longer be able to consistently order objects from 'light' to 'dark'. The objects will change their appearance (to the achromat) under different light sources. And without chromatic processing, the achromat will not be able to compensate for those changes. Of course, the property perceived by the achromat will not be *albedo*. Put ten differently coloured objects in a row and ask the achromat to arrange the objects in order from 'light' to 'dark' and this ordering will not be their correct ordering for albedo. But as I said, the strategy is stable enough to do *some* work, to allow the achromat to correctly select a desired shirt—to re-identify a particular—if the shirt is viewed under the same light source each time.

Let me try to sum up the argument so far. At the very outset of the paper, I said that there is a common view about the nature of luminance processing, namely that in virtue of luminance vision alone, a person is able to see more or less what she would see when looking at a monochrome photograph. This is the assumption a luminance system gives us access to an 'intensity image' at either the initial level of luminance encoding (at the receptor or ganglion cell level) or at a later stage in processing. I have argued that there is no such encoding, at any level. A defender of mental paint might then respond as follows: it is the end of result of luminance processing that is of interest. In virtue of luminance processing, plus albedo perception, we see more or less what we would if we looked at a black and white photograph. Again, this seems highly unlikely given the restricted nature of luminance information (last section) without the aid of a chromatic system. More specifically, a rod achromat or a cone monochromat is very unlikely to see albedo, as we do, without chromatic information. In other words, we cannot see objects as light or dark without chromatic information. This is a quite counterintuitive result for many people. Why would you need to 'colour' to see 'on black and white'? As I hope is now clear, this way of putting the question precludes an informative answer.

There is a deeper general point to the above, however. At this point in our scientific research, we do not know very much about how, in the course of normal trichromatic

visual processing, chromatic information and luminance information are used to see the world. We can present stimuli to the trichromat that do not contain any chromatic contrast, i.e. monochrome photos or displays, and measure what the trichromat's responses. We can compare the trichromat's performance on these monochromatic displays to her performance with full colour displays, i.e. measure whether additional chromatic information enhances vision along various dimensions. But such experiments do not tell us how chromatic information is normally used in tandem with luminance information. (Think of it this way. If you want to find out how gait is co-ordinated in aa human bi-ped, you won't learn a lot by first watching him hop on the right leg, and then 'subtracting' the hopping from his normal two-legged gait.) What the foregoing should make very clear, however, is we have very little introspective access to the information that each system brings to normal vision. Merely asking the normal trichromat how things appear at night is not a fail-safe guide to what she actually sees (e.g. whether she sees albedo). Nor could the introspection of one's night vision reveal how her visual system would work under daylight conditions. When the cones are active and providing luminance information to visual cortex, it is an open question how the luminance and chromatic systems will divide up (or cooperate on) various tasks.

What we do know is this: it is unwise to equate luminance vision in the rod achromat, night vision in of the trichromat, and whatever happens when a trichromat looks a monochrome photograph. A trichromat who looks at a monochrome photograph receives intensity information about a scene; because she is a trichromat, her visual system also infers that the photograph does not contain chromatic contrast. At night, the trichromat has access to only luminance contrast, as opposed to intensity contrast. So from environmental information alone, she is probably unable to perceive albedo accurately. However, she does have access, via memory, to albedo information (which objects are light and which are dark) as well as memories of their surface colour. This knowledge is something that could affect her visual phenomenology.[12] Finally, a rod achromat has access, whether at night or during the day, to only luminance information. She does not know the intensity of light across the retinal image. She cannot perceive albedo either, but she can order objects by surface 'luminance', a property that remains stable under natural nighttime illumination. She may well call things 'light' or 'dark' but she does not pick out the same property to which a trichromat refers using the same words. Given these facts, and the differences in cortical functioning that each type of system would develop, it is very unwise to identify these three types of experience. We should not elide the differences with the phrase 'seeing in black and white'. 'In passing', then, this is another good reason not to identify the rod achromat's visual experience in general with the experience of trichromatic night vision.

[12] For a very interesting experiment on the perceived colour of objects in monochrome photographs, see Hansen et al. 2006.