Introduction

The Purpose of this Book

Over my years as a practicing artist and as an instructor in the arts, and specifically in my capacity as an instructor of perspective and composition, I have been asked repeatedly: “Can you recommend any books on composition?” Sadly, I have been able to refer students to books, but never to any *good* books. This is, I hope, the book that I can now recommend to students, and to readers everywhere interested in the subject.

By *good* I mean: books that are clear as to *why* images are composed the way they are, and what *effect* those arrangement, and as a result those pictures, have on an audience. In my experience, while most books on the subject have *some* useful information in them, none are comprehensive in the information covered. Worse yet, it is not often clear *what* they are speaking about, even less obvious *how* the tools they propose should be applied, and almost never apparent *why* they would be employed. This text is an effort to provide a clear and comprehensive answer to all of those questions.

I believe that artistic education starts from ‘the ground up’. The ‘ground’ in this case refers to having the ability of ordering objects in relation to each other, and creating space and feeling, with *intention*. Too often in contemporary art education, this foundational ability is overlooked, or overstepped, to grapple with theoretical or expressive concerns: “*An artist does not skip steps; if he does, it is a waste of time because he has to climb them later*.”[1] While the goals of theory and expression are essential to the creation of imagery, this leapfrogging of more fundamental concerns undermines the process of art education by not providing the tools for students to effectively communicate their message. Composition for two-dimensional art is essential for communicating effectively with an audience. Without this ability, important aspects of any visual message may be misinterpreted or simply lost.

When viewing an artwork (and I mean this in an exhaustive sense two-dimensionally: webpages, comic books, film, television, painting, drawing, etc.) we must consider it in *all* its capacities as an image. What the image is comprised of (i.e. people, places, things, splotches of paint, scribbled lines, etc.) is only part if it’s message. *How* things are arranged, coloured, and designed is at the very least equal in importance to *what* they are. I would go further and say that: what a thing *is* owes to as much to *how* it is *designed*, and how it stands *in relation to* other things.

Our audiences, whoever they may be, only know what we *show* them. What we see affects what we think: our sight fulfills our desire to interact with the external world and in turn aids in the development of our attitudes toward that world. When we make an image, we are creating opinions, beliefs, reactions, values…we are speaking directly to who and what our audience members most fundamentally *are* as people. We speak through imagery and metaphor, but we speak nonetheless, and the more power you have to control the imagery you present will help determine the success of that communication.

We may be speaking directly or indirectly to our audience. We may say what we wish to communicate in a purely superficial manner, or we may engage in metaphoric imagery that only *represents* an underlying message. But, if we make imagery of any kind, we are

---

[1] *Jean Cocteau. Art and Visual Perception, p.204*

engaging the viewer with representational *concepts* that must be interpreted. An image of any sort may ultimately represent *anything*. Communicating effectively with visual imagery is having your representational concepts resolve into an *intelligible* idea for your audience. This allows our audience to understand *what* it is we are trying to communicate. It is not always enough as image-makers to simply *self-express*; often, we must also have the ability to effectively *express an idea to others*, so that they may understand what we wish to say, and feel what we wish them to feel.

The pictures that we create are the objects of our audience's attention in a way that the natural world is not. A viewer goes out of their way to look at pictures: whether they be in a gallery, theatre, on a webpage, flipping through an illustrated book, or walking down the street surrounded by an array of advertising. I truly believe that if we are to put images out into the world, that we put out images of *value*. Images that are clear, comprehensible, and have both significance and meaning help create that value. Compositional control over imagery is the first, although most certainly not the last, step towards creating just this kind of imagery, and it is the aim of these pages to provide you with the tools to do so.

How to Use this Book

As all books, you may choose to read this book in a linear fashion as each chapter builds on the previous in a relatively straightforward way. This *Introduction* provides what I deem to be useful background information on our biology, psychology, and the history of the relevant subjects. I believe it will enrich an understanding of why we perceive and react to imagery, which will in turn enhance your ability to create effective pictures. That being said, this background information is less direct in its application to image-making, and can be safely skipped at any time if it is not of interest, and will not significantly depreciate your understanding of the subject.

Each chapter is also structured so that it can be read independently from the others and still be comprehensible. This will provide you the opportunity to skip to the topic of your choice without fear of being lost for want of information previously covered. In the event that it would be useful to achieve a better understanding of a subject mentioned in the course of a chapter that is covered elsewhere in the text, there will be a notification of *where* to find that information: (*see Chapter X).*

Finally, this book has not been created with the intention of being a 'how-to-draw' book. Thus, all technical information regarding how to render perspective objects has been reserved for the *Appendix*. These are laid out in a very straightforward manner with sufficient examples to emphasize their individual concepts. Whereas there are not very many good books on composition available (in my humble opinion), there are a great many excellent texts available on perspective rendering. As part of the *Appendix*, I also provide a bibliography of those books that are available and their degree of usefulness.

I hope these pages will serve you well in your journey through the arts. It represents more years of experience and labour on my part than I care to mention, and exposure to the knowledge of a great many people…some of whom I have the pleasure to know and call friends, others whom I have had the good fortune to have as students, and still others that I know only through their work and words. My thanks to them all; and my good wishes to you.

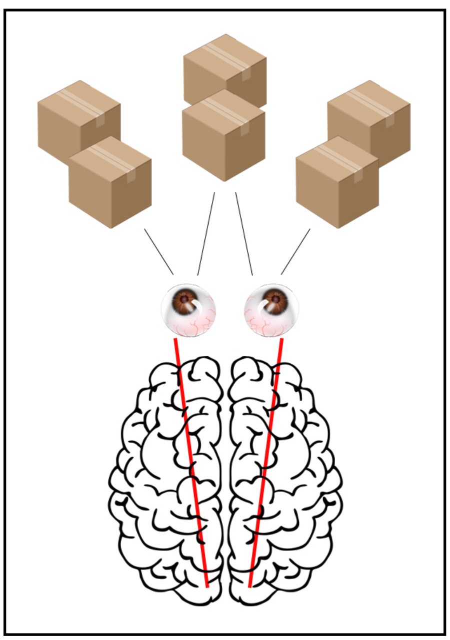
Enjoy.

Physiology of Vision

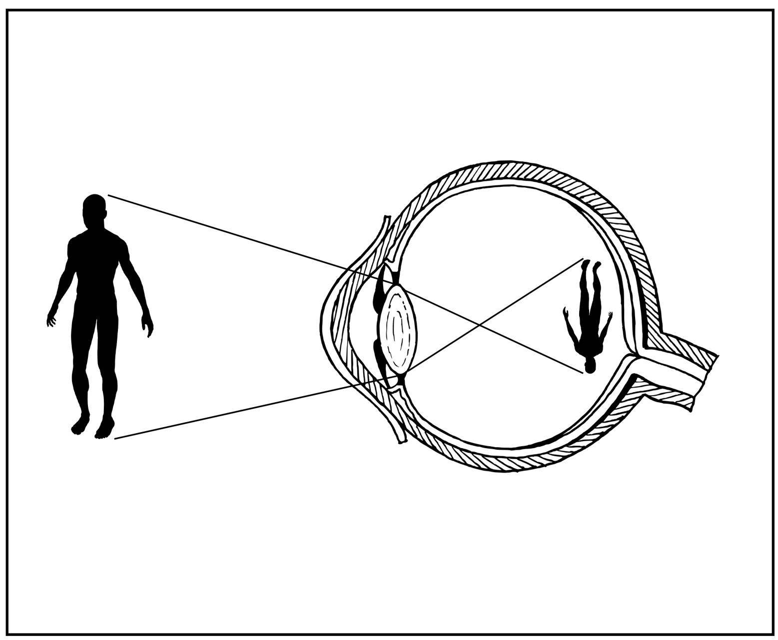
"All men by nature desire to know. An indication of this is the delight we take in our senses; for even apart from their usefulness they are loved for themselves; and above all others the sense of sight. For not only with a view to action, but even when we are not going to do anything, we prefer sight to almost everything else. The reason is that this, most of all the senses, makes us know and brings to light many differences between things."[2]

Of all of our senses, vision is the most studied and the most cherished in our ability to navigate the world. It is the front line of how we perceive and interpret the world both physically and psychologically. Our perception of the world ends in the mind, but it begins in the eye.

Evidence suggests that our perception of the world begins at a very early stage. Child psychologists have noted that the eyes of prenatal infants move independently beneath the lids at least 6 months before birth; and that the first 8 weeks of life outside the womb witnesses infants looking out onto the world absorbing information, while they remain relatively useless in terms of manual dexterity. This, among other examples, seems to be compelling evidence that we begin to learn about the world primarily through visual experience as opposed to manual interaction. Our ability to intake information from the external world, and apply a pattern to it is a fundamental precondition to perceive the world as three-dimensional. It is also a precondition (for the sighted) for our application of *meaning* to that world. How things *appear* to us affect how we *think* of them.



Ill.1: Stereoscopic vision unifies different images from either eye.



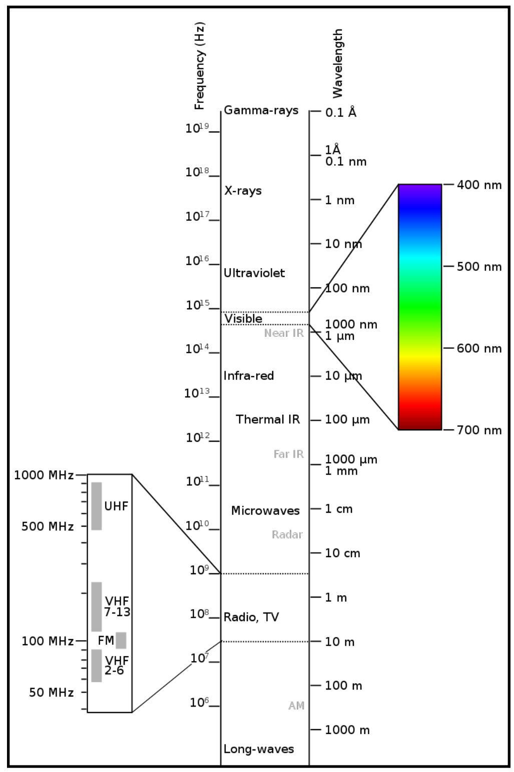
Ill.2: Light is refracted by the cornea onto the lens which projects an image onto the retina that is both reversed and inverted.

Problematically, vision is inherently ambiguous owing to the nature of how the mechanism of sight actually works. Even the basic fact that our vision is *stereoscopic*

[2] *Metaphysics, Aristotle, Book I, 980a.21.*

means that the image seen by the left eye is slightly different that what is seen by the right (*Ill.1*). The brain converges these slightly different images into a single image, in order to provide a simplified representation of the world. However, that does not necessarily mean that what we see is a *truthful* representation. The issue of *truthful* representation becomes even more apparent when one considers that the image reflected on the retina is both *reversed* and *inverted* from the one that we actually *perceive*. The image that we receive of the world *at the eye* is *ambiguous at best*, and means that there can be *no universal rules of vision*.

However, vision has been studied for just about as long as studying has been around. The biological structure of the eye was known even before the famous Roman physician, surgeon and philosopher Galen of Pergamon (129-199) made his discoveries known, but its *function* was not understood until Johannes Kepler (1571-1630) developed his *retinal theory* of 1604. Kepler is considered by neuroscientists to be the first to recognize that images are projected inverted and reversed by the eye's lens onto the retina (*Ill.2*). It is important to recognize, that we see thing right-side-up simply because our brains are used to doing so. A fascinating psychological study has shown that our vision can be retrained to see things upside down using special corrective lenses. However, that same study also showed that once we were used to seeing things upside down, our brains quickly forced them back into a right-side-up position...and then did so again once 'normal' vision was restored. This means that we see the world the way we do because it is easier for our brains to perceive it as such; and there are many more ways that the brain interferes with our perception of reality in order to make it more manageable. Kepler's theory of refraction in spherical lenses stated (and was accurate in so doing) that it was the function of the eye to focus an image on the retina. It does so by capturing packets of light called *photons* that are emitted or reflected by our surrounding environment. The cornea refracts light from the environment and focuses it on the lens, which in turn can both elongate and become more spherical in order to project this information onto the retina.

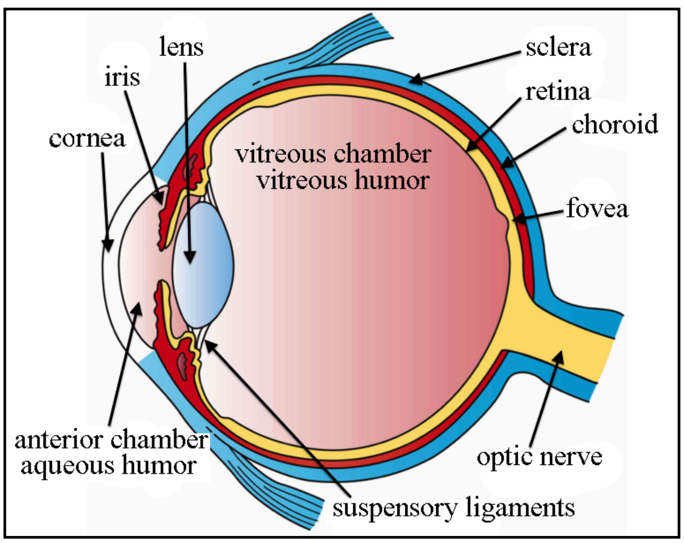


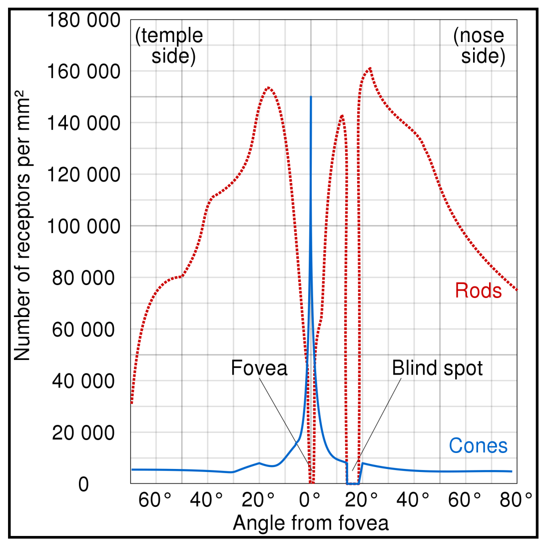
Ill.3: The colours we perceive constitute a very small portion of the electromagnetic spectrum as our visual system is only sensitive to frequencies within 360-760mm. Image by Victor Blacus. Licensed under the Creative Commons Attribution-Share Alike 3.0 Unported license.

Light has been described as both wave and particle, and sometimes as both. Let's assume that light is a particle for the time being. Light particles are called *quanta* and exist along a very small portion of what is called the *electromagnetic radiation spectrum (Ill.3)*. At the high energy end of this spectrum occur gamma-rays (what The Hulk is transformed by), and at the low energy end of the spectrum exist radio waves. Each *quanta* of *electromagnetic radiation* happens to have a specific *frequency*, but our visual system is only sensitive to a very small portion of the frequency range that exists between 360-760nm. Even though we can only see a fraction of that spectrum, we are are very good at doing so: if our visual system detects just 5 quanta of electromagnetic radiation, the normal percipient will perceive light...which is another way of saying that our visual system is *extremely* sensitive to stimuli.

This *electromagnetic radiation* is then transduced by millions of *photoreceptors* into *electrochemical* messages via *retinal ganglion* cells which ultimately passes that information along to be interpreted by the *visual cortex* in the brain. The optical image created by the projection of light on the *retina* stimulates around 130 million of these light-sensitive *receptors*, each of which respond to the frequency and intensity of light it is exposed to.

These *photoreceptors* are called *rods* and *cones*, and are densely packed around a small area on the *retina* called the *fovea*. These two different *receptor cells* are what gives rise to the term *duplex structure of vision.* There are about 6-7 million *cones* and 127 million *rods*. Inside the *fovea* there are only *cones,* and as soon as one leaves the fovea the ratio of *cones* to *rods* decreases quickly until one reaches the extreme periphery of the *retina* where there are only *rods (Ill.4)*.



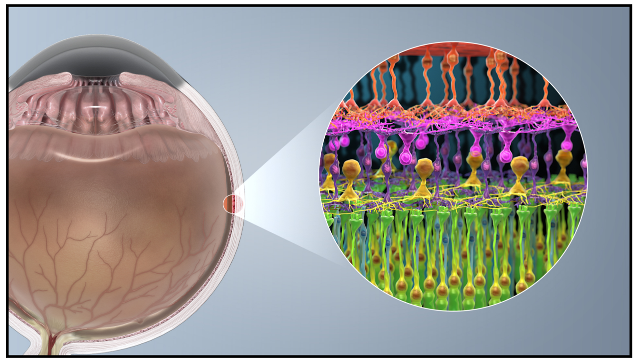


Ill.4: Photoreceptors are called rods and cones, and are densely packed around a small area on the retina called the fovea. Inside the fovea there are only cones and at the extreme periphery of the retina there are only rods. Image (left) by Holly Fischer; Image (right) by Cmglee. Both licensed under the Creative Commons Attribution 3.0 Unported license.

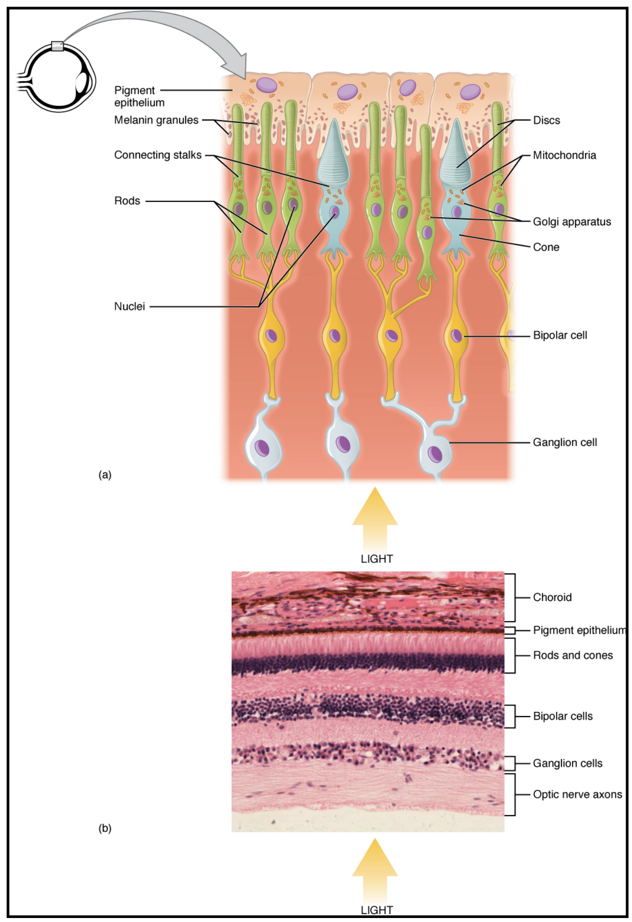
Rods and *cones* are found in the outermost layer of the retina, and are connected via *bipolar cells* to *retinal ganglion cells*, which are the first true *neuron* in visual system (*Ill.5*). The *axon* of every *retinal ganglion cell* is a fiber in the *optical nerve*, which contains about 1 million fibers overall. This information ultimately culminates in the *occipital cortex*, the area of the brain associated with vision. But *how* the *rods* and *cones* of the visual system operate does much to explain why we perceive the world the way we do.

In the *fovea* there are approximately 3 *foveal cones* to every 2 *bipolar cells* to every 3 *retinal ganglion cells*. In other words, every *cone* has its own *retinal ganglion cell* devoted to transferring its information along the *optical nerve* into the brain for processing. However, as soon as one leaves the *fovea*, this ratio changes wildly, so that hundreds or thousands of *rods* may be sharing dozens of *bipolar cells* which will in turn be funnelling all of the received information into a single *retinal ganglion cell (Ill.6)*.

Understanding how our biology and psychology function perceptually allows us to better control how we can best construct our imagery, because these ratios have predictive effects on how our visual system operates. There are aspects of our perception that are hard-wired into our biological and thus psychological make-up. This means that when any person with a normally functioning perceptual system is exposed to certain kinds of stimuli, they *must* respond in a predictable manner because the system (both biological and psychological) is *can only* to respond in this manner due to its structure. *Rod-* and *cone-meditated* vision are two of exactly these kinds of perceptual systems. *Rod-mediated vision* is best described as being responsible for the *summation of energy* in the periphery of the visual system. Because of the pooling of information of thousands of *rods* into a single *retinal ganglion cell*, *rods* will respond to much lower levels of energy than will *cones*. This also means that *rods* work best in low-light situations: we can see much more clearly out of our peripheral vision at night or in dimly lit situations, than we can with *foveal vision*. In addition to this, vision mediated by *rods* is not colour vision, as the *rod* system does not feed information into the part of the brain responsible for colour vision. Evolutionarily this makes sense: because our peripheral vision is very sensitive to low thresholds of movement and light, we are much more likely to react to minimal sensations that may pose a threat to our person when perceived by this area of our vision. For purposes of this text, this has massive implications on predicting viewer response to an image. Our primitive hard-wired response to changes in lighting situations and movement means that we will be *biologically* affected by these percepts. This means that levels of *value*



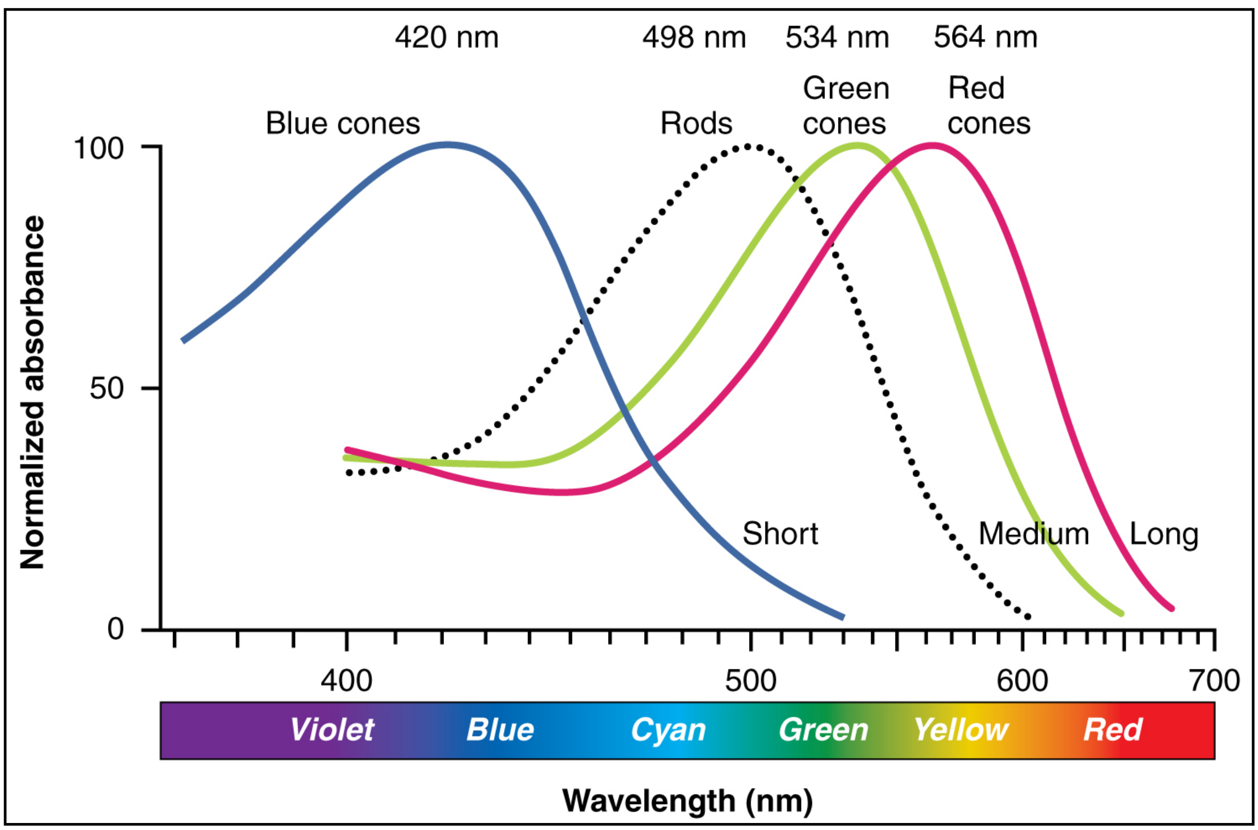
Ill.5: Rods and cones are connected via bipolar cells to retinal ganglion cells. The axon of every retinal ganglion cell is a fiber in the optical nerve. Image courtesy of www.scientificanimations.com. Creative Commons Attribution-Share Alike 4.0 International license.



Ill.6: Every cone has its own retinal ganglion cell, while hundreds, and sometimes thousands of rods will share a single retinal ganglion cell. Image courtesy of OpenStax College. Creative Commons Attribution-Share Alike 3.0 Unported license.

contrast can be manipulated in order to affect how an audience responds to an image on a deep-seated emotional level triggered by a physiological reaction (*see Chapter 5*). The *cone* system is different altogether. It is common to refer to the *cones*, and vision that corresponds closely with the area around the *fovea*, as being responsible for the extremely precise visual acuity we possess. In other words, *cones* are more active in *focal vision*. When we create *focal points* (each chapter in this text refers to different ways in which to do this) we are cuing the viewer to pay attention to that object, or this objects, in the same way as when we physically bring an object into focus. It has become important enough to us to focus upon. When we look *at* something we physiologically focus on it, and it is our *cones* that are doing this work. Because each *cone* has a dedicated *retinal ganglion cell,* it is capable of transferring massive amounts of information, including colour.

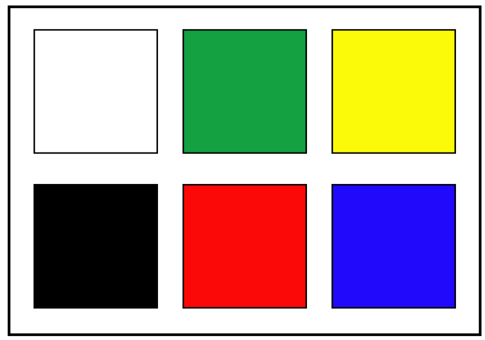
There are three different kinds of *cones*, each of which contains a different *photopigment*. The *photopigments* in each of these *cone* types respond to different *frequencies* of *electromagnetic radiation*, and in turn are responsible for our experience of the colour associated with those frequencies. *S-cones* respond preferentially to *high frequencies*, or *short wavelengths*, and are responsible for our experience of the blue/violet end of the visible spectrum; *M-cones* respond to *intermediary frequencies* and *wavelengths*, and are responsible for the experience of the yellow/yellow-green part of the spectrum; and *L-cones* respond to *low frequencies*, or *long wavelengths*, and mediate our experience of red. These receptors have a peak sensitivity in each case: *S-cones* are most sensitive from 420-440nm, *M-cones* from 534-555nm, and *L-cones* from 564-580 (*Ill.7*). However, these



Ill.7: S, M, and L-cones all have a range and peak level of sensitivity. Their overlapping nature allows us to experience the entire visible spectrum. Image courtesy of OpenStax College. Creative Commons Attribution-Share Alike 3.0 Unported license.

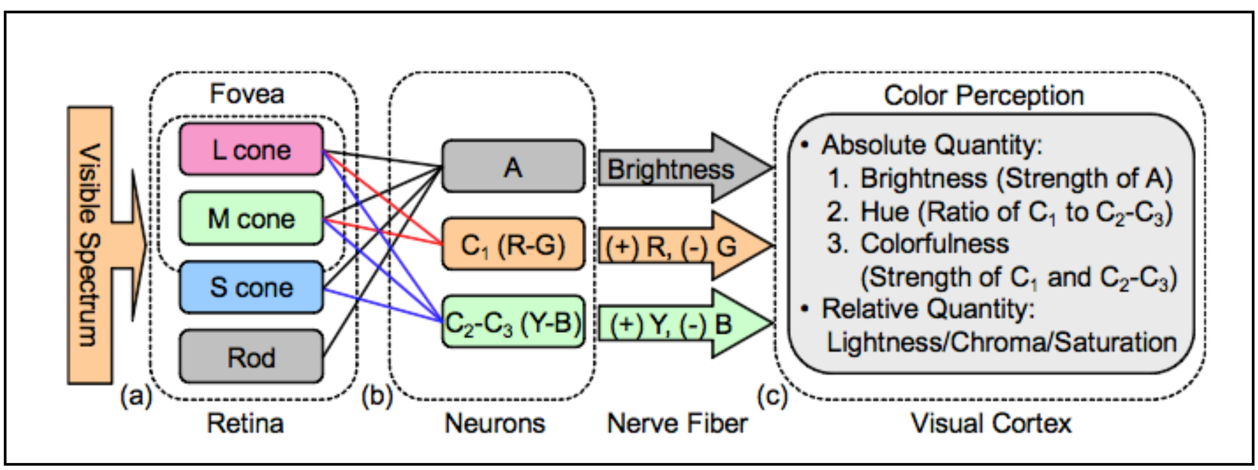
cones are active across a wider range of the spectrum outside of this peak performance, and this wider range creates areas along the spectrum where two, or all three receptors are activated (*S-cones* from 400-500nm, *M-cones* from 450-630nm, *L-cones* from 500-700nm). The overlapping nature of *cone* sensitivity means that with proper stimulation of the *receptors*, every colour of the visible spectrum may be experienced. This *cone-receptor* function was first suggested by Thomas Young (1773-1829), was later developed by Hermann von Helmholtz (1821-1894), and is the basis for the *trichromatic* (RGB, or RGV) *theory of colour vision* (or *Young-Helmholtz theory*). The theory, like all theories, is not without its imperfections, but does much to explain our experience of colour and visible world.

However, colour perception does not end at the eye. The image at the eye is not continuous, it contains only dots. The processing of that information occurs in the brain where the received information is made continuous for our perception (i.e. we do not 'see' in a straight line, but interpret a straight line as a simplified path through dots of pigment that zig-zag). While the *pigment chemistry* of the *retina* follows the *trichromatic theory of colour vision,* the electrophysiology of colour vision adheres to the *opponent process theory of colour vision.* Ewald Hering (1834-1918) developed the theory in 1872 and it is needed to compliment the *triple-receptor theory* of Thomas Young in order to account for the facts observed in colour vision. Initially believed to be antagonistic towards each other, it is now thought that both theories are valid and operative in the perception of light and colour as each describes different stages in visual physiology. *Trichromacy* arises at the level of the receptors, and *opponent processes* arise at the level of *retinal ganglion cells* and beyond.



Ill.8: Opponent Process theory identifies three antagonistic colour pairings: black/white, blue/yellow, red/green.

The *opponent process theory* states that our visual system interprets colour in an antagonistic way: red vs. green, blue vs. yellow, black vs. white (*Ill.8*). A range of *wavelengths* of light stimulates each of these receptor cells (both cones and rods) to



Ill.9: The antagonistic colour pairings of Opponent Process theory hypothesizes that fatiguing one colour is the same as exciting the other. Image courtesy of Googolplexbyte. Creative Commons Attribution-Share Alike 3.0 Unported license.