Could an Old Theory of Color Vision Make Our RGB Devices Obsolete?

A HoosierScientist
Frederick J. Thomas
December 6, 2020

After centuries of scientific, technical and artistic investigation, our understanding of human color vision remains incomplete. In particular, two different theories of color vision have been under discussion for a very long time, with neither theory quite able to claim victory. This paper introduces the areas in which the two theories agree, ways in which the theories differ, and ways in which the difference might impact the present and future of color imaging. One question is whether our modern RGB cameras, TVs and computer screens might become as obsolete as the film cameras and analog projectors they replaced.

The *Trichromatic Theory of Color Vision* is the better known of the two theories and has the clearest links to modern RGB (red-green-blue) technologies. It is also known as the *Young-Helmholtz Theory*, since it is rooted in the work of physicists Thomas Young and Hermann von Helmholtz. Young proposed the basic idea in 1802 and Helmholtz developed it further in 1850. In 1861, Physicist James Clerk Maxwell recognized an important implication of the theory and worked with photographer Thomas Sutton to display what is generally regarded as the first permanent color photograph. Their technique of red-green-blue separation is the foundation for both color film photography and the RGB digital imaging which followed.

Opponent Color Theory also has deep historical roots, having been proposed by physiologist Ewald Hering in 1892. Although Hering's theory is less well known, experts generally agree it is superior to trichromatic theory in its ability to explain some aspects of human color vision. A complete understanding of how humans perceive color seems to require some yet-to-be-achieved synthesis of the two theories. Although not the default setting, Hering's theory remains with us today as the "Lab Color" alternative to RGB methods within image processing software such as Adobe Photoshop.

<u>The Human Eye</u>

The two theories agree that vision depends upon millions of photoreceptor cells in the retina, where they capture light photons in somewhat the same way as pixels in the sensor of a digital camera. "Rod cells" are the most numerous and most sensitive of the photoreceptors, but their activity is only obvious when they provide us with gray-scale images in very dim light.

Color vision relies upon three types of "cone cells." These are sometimes referred to as being red, green and blue sensitive, but the reality is more complex as shown in *Figure 1*. S-cones are most sensitive to short wavelengths of light which includes blues and violets, and they have very little sensitivity to greens, yellows and reds. M-cones (medium wavelength) and L-cones (long wavelength) are both sensitive to nearly the entire spectrum of visible light, but they respond differently to all wavelengths except their intersection point in the yellow-green region.[i]

Note that *Figure 1* is "normalized" to make the peaks equal. S-cones, however, are much less numerous than the others, and the overall peak sensitivity of human vision is to yellowish green light near the intersection point of the M- and L-cones. The fact that the peaks for M- and L-cones lie close

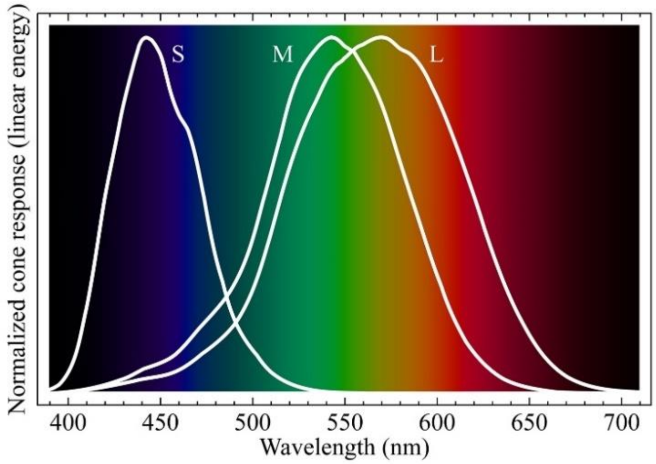


Figure 1: Normalized response of human cone cells vs wavelength Ben Rudiak-Gould, Public domain, via Wikimedia Commons

together in the yellow-green region enhances our ability to see those colors. It also seems to enhance our ability to distinguish quite well between different hues in that region of the spectrum.

Trichromatic theory assumes that each type of cone cell captures information about the strength of light in accordance with its spectral sensitivity and transmits that information to the brain as distinct red, green and blue signals. Maxwell recognized the implication that human vision could be fooled into experiencing a full-color image by combining just three monochromatic pictures. A spectrometer or prism can easily separate a red image from a green one, but the human eye sees that combination as being yellow. The effect can be achieved using nearly pure red and green images which a spectrometer shows to include no yellow at all. It is more efficient, however, to use broad spectrum sources—ideally sources which are close to the sensitivity curves for our cone cells.

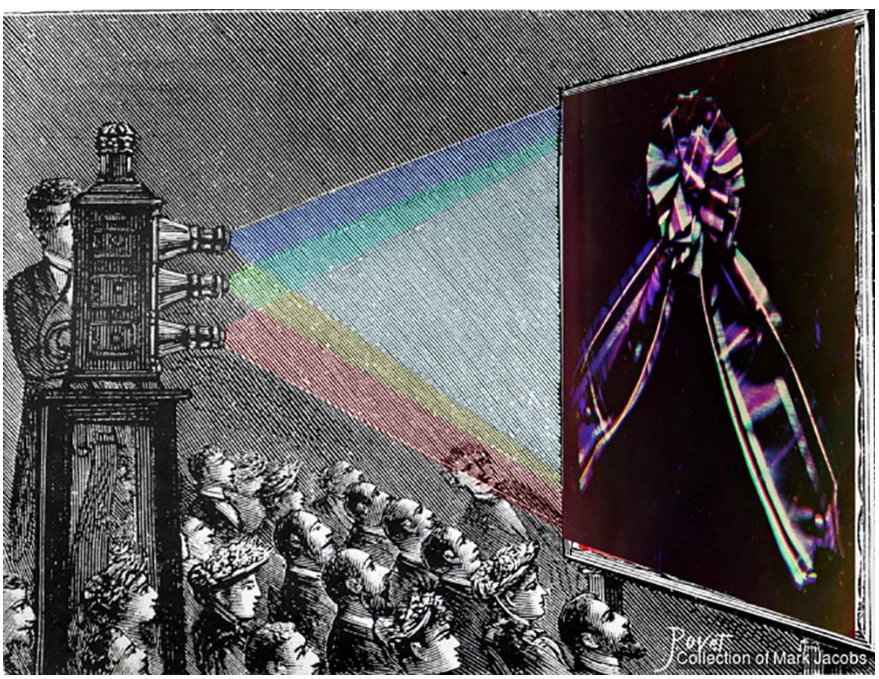


Figure 2: The Maxwell-Sutton demonstration of RGB photography

Maxwell instructed Sutton to take four gray-scale photographs of a tartan ribbon in bright sunlight, one each through red, green, yellow and blue filters. He was responding at the time to a debate as to whether yellow or green was the appropriate primary color, but used only the red, green and blue photos in his presentation as illustrated in *Figure 2*. His presentation at a conference on physiology and physics produced a flurry of interest, but more as a curiosity than the important step towards modern RGB technology which it was. It would be nearly five decades before the chemistry of photography developed enough to provide the first commercially viable color films—capturing three separate red, green and blue images on a single piece of film.

The Mathematics of Digital RGB

Particularly with the introduction of digital imaging, it became very important to develop a precise mathematical description of how the human eye responds to color variations. The topic remains important to scientists, engineers and others who work on the hardware and software which support medical, artistic and other imaging. It is also of interest to at least some visual artists and graphic designers. Serious users of Adobe Photoshop, for example, who venture into the menu for "Color Settings" as shown in *Figure 3* may find they are using the default setting for RGB conversions, "sRGB IEC61966-2.1." This standard comes from the International Electrotechnical Commission (IEC) and provides the complex mathematical conversion needed to match the output of an "average CRT display" to the color response of the average human eye. Especially since cathode ray tubes (CRTs) are largely obsolete, discerning users may be wise to consider one of the other fifteen or so alternatives for RGB conversion which Photoshop provides. There are also alternative modes including "CYMK Color" which translates the image into instructions appropriate for a printer using cyan, yellow, magenta and black inks.

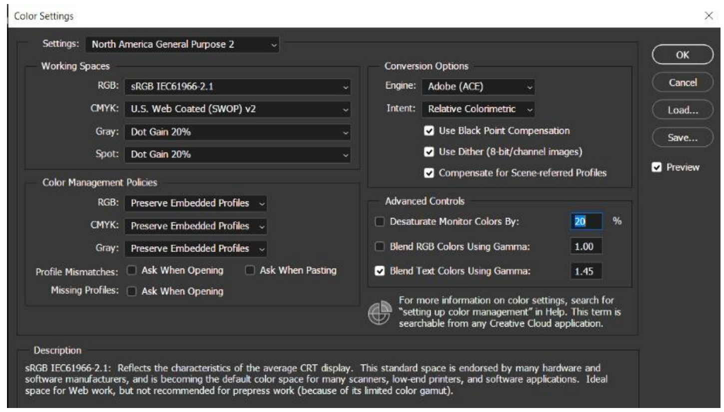


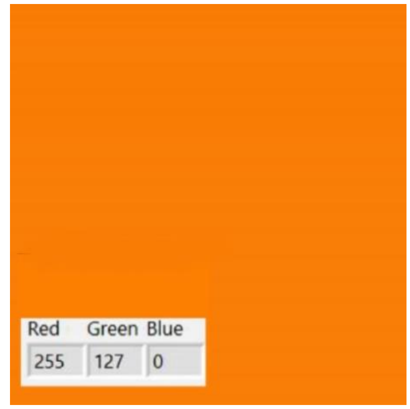
Figure 3 One of Several Options for RGB Conversions in Adobe Photoshop

All RGB and CYMK technologies are grounded in *Trichromatic* theory—overlaying three colorized gray-scale images (plus maybe the "no color" of black) to replicate the human experience of seeing a full rainbow of colors. There are many variations, but we focus here on RGB with "8-bit color depth" in which the strength of each RGB color component is represented by an integer from 0 through 255.

In this system, a specification of (0,0,0) calls on a single pixel of the output device to turn off its red, green and blue emitters, resulting in the darkest (blackest) output possible. A specification of (255,255,255) calls on the same red, green and blue emitters to provide their maximum possible output, which—if the system is adjusted correctly—combine to give a human viewer the experience of seeing bright white. A specification of (255,255,0) calls for full red and full green with no blue, which should appear to humans as a bright yellow. Intermediate values can shift the perceived hue. A specification of (255, 127,0), for example, stimulates the L-cones (red) more than the M-cones (blue). Humans see this as a shift towards red in the spectrum and perceive it as orange. There are 256^3 or 16,777,217 possible combinations of the RGB values, so the system is often described as capable of replicating over 16 million different colors.

Issues begin to arise, however, which require care in defining the word, "color." Physicists sometimes equate the word with wavelength, so that each "color" corresponds to a specific wavelength on the spectrum. Artists are more careful, using "color" to incorporate both location on the spectrum—which they call "hue"—and also "tint" or "shading." A pure orange paint, for example can be changed dramatically by mixing it with either white or black paint. It still retains its location on the color spectrum but it looks very different.

RGB systems successfully emulate tints and shading. They can start, for example, with the most brilliant orange which the output device can produce (255,127,0) and shift it towards white (255,255,255) or towards black (0,0,0). Video 1 shows these progressions using linear changes to red, green and blue, although it is not obvious that linear functions really provide the best match to the human perception of a true orange source in all its possible tints and shades. The video's progression through the RGB tints and shades of orange utilize over 500 RGB combinations, so the number of actual *hues* an 8-bit system can produce is closer to 30,000 than to 16 million. That still divides the 300-nm range of visible wavelengths into approximately 0.01 nm steps. This is far better than most human eyes can detect, and the spectral resolution can be improved further by shifting to a 16-bit or 32-bit RGB system.



Video 1 Tints and Shades of Orange

A significant problem with all RGB systems is the fact they are device-specific. The output is bounded by the darkest "off" and the brightest "on" which each color component can produce and it is influenced by the precise color spectrum which each emitter produces. The set of RGB values which produce a beautifully colored image on a CRT screen, for example, may display grossly distorted colors when sent to the tri-color light-emitting diodes (RGB LEDs) of an outdoor billboard or to the pixel elements in a very expensive medical display screen

Hering's Alternative

Hering's opponent process theory offers an alternative way of thinking about color, and that alternative approach is available today in image processing software. There is even a non-zero chance that his will be the basis for future types of cameras, TVs and other equipment which makes RGB systems as obsolete as film cameras and projectors. The reference "device" in a Hering-based system is not a CRT monitor or some other specific hardware but rather the actual human experience of color. Opponent process theory suggests that the color information which actually reaches our brains is ***not*** in the form of distinct R, G and B values.

Opponent process theory survives scientifically after 130 years because it seems superior to trichromatic theory in explaining some observations, including "afterimages." If humans stare at a bright blue spot for several seconds and then look at a white background, it is common for people to report they see a yellowish afterimage—something you can test for yourself using *Figure 4*. Opponent process theory does not deny we have red-sensitive L-cones and green-sensitive M-cones. It does content that the information delivered to our brains—perhaps after some sort of pre-processing in the optic nerve—is about the *relative stimulation of cone pairs* rather than separate red, green and blue values. Moreover, this relative value is a pure measure of hue as defined by wavelength, without regard to shading or tint. Light orange, for example, is still orange when defined by the relative stimulation of the red and green cones, regardless of the absolute strength of the light which stimulates each cone. Under this theory, there is a similar pairing of yellow and blue, perhaps weighing the combined L- and M-cones against the S-cones.

There are various systems to describe these cone pairing mathematically, with the result that the variable. a*, represents the relative strength of red versus green and variable, b*, represents the relative strength of yellow versus blue. The two values a* and b* express everything there is to know about hue, but say nothing about brightness. Brightness (i.e., tints and shades) is made mathematical in Hering's system as the value of L*, perhaps related to the total stimulation of all cones and perhaps even involving rod cells.

In Photoshop, Hering's system is called "Lab Color," a name which reflects the letters used as variables and has no particular connection to the word, "laboratory." A better designation is L*a*b* (read as "L-star, a-star, b-star"). Each pixel of an image stored in "Lab" format contains values for L*, a*, and b* only, completely dispensing with R, G and B. The variable a* indicates placement on the red-green continuum, with positive values indicating redness predominates and negative values indicating that greenness is stronger. Similarly, positive values for b* indicate yellowness predominates and negative values indicate blueness. The zero point for a* and b* is a colorless shade of gray.

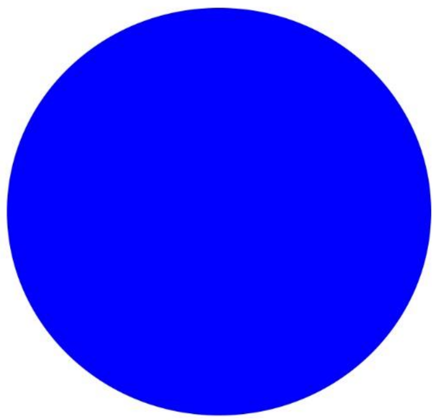


Figure 4: Afterimage? Stare at the blue spot for 20 seconds, then look at the white.

There are valid reasons to prefer the L*a*b* choice when manipulating images in Photoshop and similar software. When carefully adjusting a background color, for example, it can be conceptually and practically easier to manipulate just the value of L* to change brightness, rather than changing R, G and B simultaneously while struggling to maintain the desired hue. On the other hand, there are brightness controls available which take over that tedium when working with RGB images.

More importantly, histograms which let the user manipulate the range of a* or b* values within an image offer a different sort of control from that allowed by histograms controlling R, G and B values. A histogram for a photo of our dog, Meimei, is shown in *Figure 5*, for both RGB mode and L*a*b* modes. The top left "Red" histogram fills the horizontal scale, showing that the image includes a full range of possible red intensities, ranging from 0 to 255. This histogram also reflects the overexposure of an area of white in the upper left of the image, where pixel value that would more accurately have exceeded 255 are bunched together at the camera's RGB maximum. In L*a*b* mode, the initial histogram for a* with its red-green comparison shows the values covering a relatively small portion of the available range. The overexposure of the white is still present in the clipped peak of the histogram, but there is now room to spread the red-green range substantially. The final image shows Meimei with both a* and b* expanded.

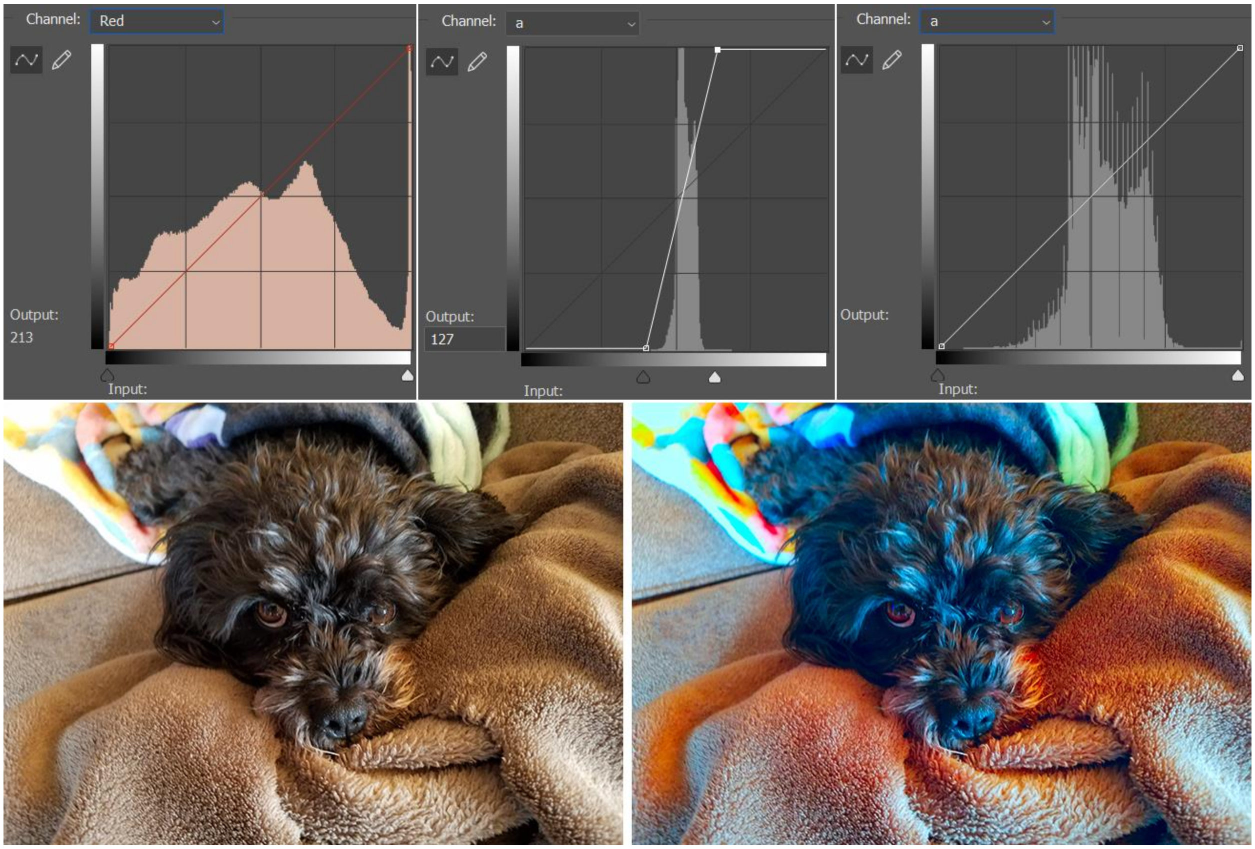


*Figure 5: Using L*a*b* to add "punch" to colors*

Switching modes does not eliminate failures caused by the limits of the photographer, limits of the camera or limits of the display screen. Even in this implementation, it does provide additional possibilities for creative artists and others who want to do more. Photoshop's current version of L*a*b* limits the range of a* and b* in 8-bit mode to an integer between -128 and +127, but this could change. Unlike RGB where 0 is a true minimum and 255 is a true maximum for the selected input or output device, the limits on a* and b* are more for the same sort of practical reasons that a piano keyboard has maximum and minimum frequencies. Expanding both a* and b* to fill Photoshop's available range yields the image shown in the bottom left of *Figure 5*. Meimei (who has dichromatic vision) might not recognize herself, but people wishing to add "punch" to their images may like this option.[ii]

A more fundamental advantage of L*a*b* is its device independence. L*a*b* uses a standard human as its reference frame, not a specific type of camera, computer screen, projector or other device. L*a*b* values must today be derived from the RGB values captured by current cameras and they must be translated back to RGB for use with current RGB output devices. Still, the image data itself is designed to depend only on fundamental human capabilities for color perception. As mentioned above, an RGB image intended for one application such as a CRT monitor may not translate well to other types of displays. If an image needs to be displayed in multiple ways—perhaps even on display devices which have not yet been invented—it is likely best to use the human reference frame. Perhaps, too, there will someday be cameras and display screens which more completely replicate the full power of human color vision.

Summary

To the extent there is a broad scientific consensus as to the “correct” theory of human color perception, it seems the trichromatic theory and opponent process theory compliment rather than contradict each other. Trichromatic theory seems to do a better job explaining what happens in the retina itself, while opponent theory may do a better job explaining the type of information which our brains actually use in the perception of color. There is still a great deal of room for additional research into the physiology and psychology of human color vision. There is also still room for continuing development of image capture and image display devices which better align with the functioning of human eyes and brains. It is unlikely our current RGB cameras, TVs and other devices will become obsolete in the near future. It is even less likely that RGB will rule forever.

---

[i] Cone cell - Wikipedia

[ii] How to Use LAB Color in Photoshop to Add Punch to Your Images (digital-photography-school.com)