BY **MIKE WOOD**

Color Rendering – Where are we?

IN THE LAST ISSUE, I promised that this time, I would take an overview look at all the options in color rendering and try to pull all the strands together. Well, that's what I said, but it's easier said than done. Frankly, the whole topic of color rendering metrics is a bit of a mess. It's actually always been a mess, but with continuous spectrum light sources and those that are close to it, such as HID, it's been a mess you could deal with. You could squint your eyes, hold your mouth just right, and convince yourself to believe what the datasheets were telling you. The problem is that the introduction of narrow-band, multi-source LED emitters means that you can't lie to yourself any more that the color rendering metrics are meaningful. CRI is just hopeless and can give you completely misleading information; CQS is somewhat better, but almost nobody uses it; and there are a raft of other metrics, all of which have utility, but nobody knows what they mean.

"... the whole topic of color rendering metrics is a bit of a mess ..."

Another problem is that, even if they were meaningful and not misleading, none of these existing metrics tell you the whole story. They are designed to evaluate a single light source that is used on its own to illuminate an object. That works fine for most people in the world. They want to know how their kitchen counters will look under this light source or how the office will look under another, and a single color rendering metric kind of works for that. However, what that metric doesn't tell you is how a scene will look when lit by many different light sources at the same time, each of which may have different characteristics, i.e. how will the light sources work together. However, I'm getting ahead of myself. Let's take a quick recap on what we know about color rendering and then look at how it works and where it falls down.

CRI – Color Rendering Index

CRI was the first and is still the best-known of all the color rendering metrics. It was originally developed as a way to distinguish and measure fluorescent lamps in the 1950s and 1960s. There are other stories around, but as far as I can make sense of it, lamp manufacturers drove the development of CRI as a tool to measure and combat consumer reluctance to their new lamps with the strange colors. Early fluorescents had appalling color rendering (as did early white LEDs), and a metric/marketing tool was needed to persuade consumers to try them again when the rendering improved. CRI was that marketing tool. It gave a hook to hang the message of improved color rendering on and a way to differentiate one product from another. Can you tell that I'm somewhat cynical about CRI? Although it ended up being used for all purposes, it was never originally designed for users or designers, it was purely designed to help differentiate and sell products.

The normal CRI metric you'll see published is Ra. This is a combined metric derived from a limited set of test colors, all of which are pastels. The CRI test color samples for Ra are designated TCS01 – TCS08 in **Figure 1**. (The corresponding results from these samples are known as R1 – R8.) These pastel shades tell you nothing about how the light source will look on more saturated colors.

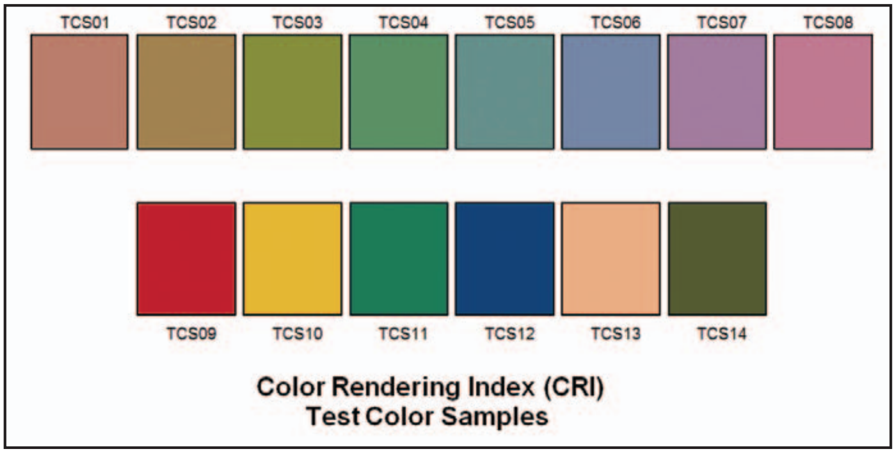


Figure 1 – CRI Test Color Samples

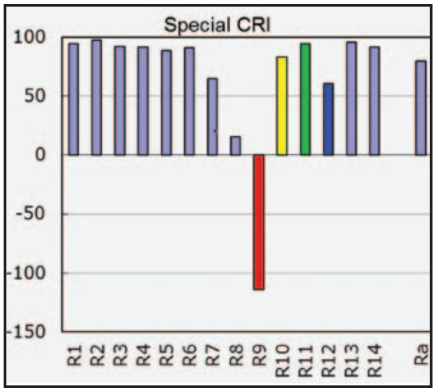


Figure 2 – Special CRI of RGB LED set to 3,300 K White


Figure 4 – CQS Test Color Samples

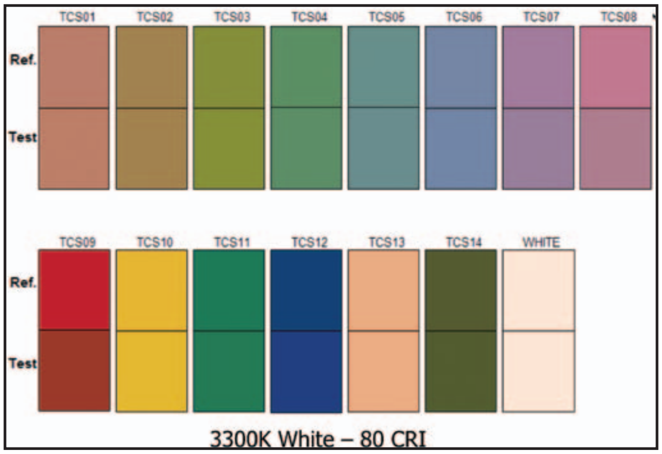


Figure 3 – CRI of RGB LED set to 3,300 K White

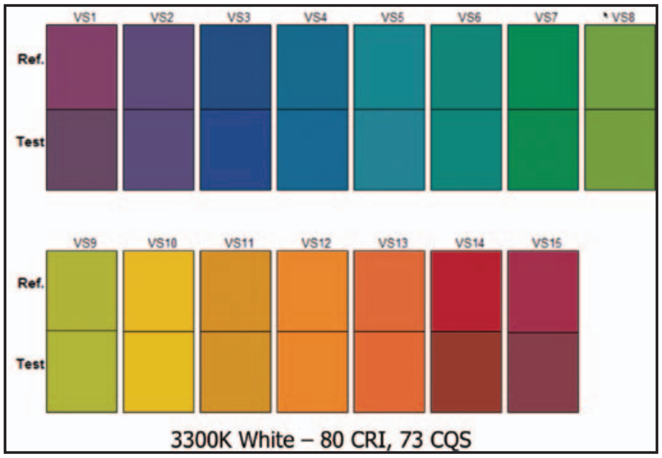


Figure 5 – CQS of RGB LED set to 3,300 K White

CRI is capable of qualifying more saturated colors, and the additional samples TCS09 – TCS14 are occasionally reported. TCS09 (R9) is particularly useful as it shows the response with deep reds. It's also a metric that's often left out with LED and fluorescent products as they both commonly produce appalling results with this color! Very low or even negative values are not uncommon, showing that the light source cannot render deep reds at all. **Figures 2** and **3** show the CRI of a particular RGB LED source with dreadful red rendering that still manages to achieve an overall CRI of 80.

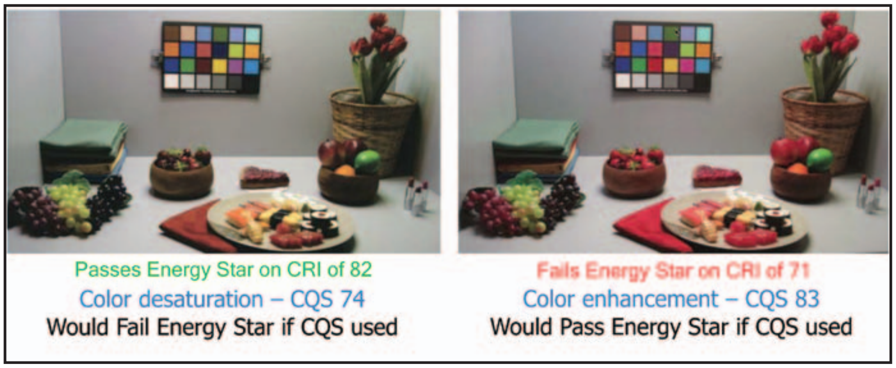
A further issue with CRI is that the Ra value reported is a simple average of the results for R1 – R8 and a single bad score easily can be hidden in that average. You could have a light source that scored 96 on seven of the test colors, but scored zero on the eighth, and it would still have a CRI Ra of 84. In practice, such a light source would be a poor choice. In other words, it is possible for a light source with an apparently good CRI to render a critical color poorly. There is more to say about CRI, particularly in how it can be gamed by less scrupulous manufacturers, but I refer you back to the Winter 2010 edition of this journal for more information. Suffice it to say that it's an almost useless metric for narrow-band emitter sources such as RGB LEDs. It's a little better with broad-band sources such as white LEDs but still not great. Lamp and LED manufacturers still like it for the same reasons they liked it when they were making fluorescents. Again, it's a manufacturer's marketing tool and not one that's useful for end-users.

CQS – Color Quality Scale

Color Quality Scale (CQS) is an improved version of CRI that tries to address some of the shortcomings and make it more useful for the consumer. In particular, it uses a wider range of more saturated color samples (**Figure 4**) and combines results in a way that doesn't hide single poor results.

Figure 5 shows the CQS results for the same RGB LED as in **Figures 2** and **3**, this time the inadequate red performance is penalized and the corresponding poor skin tone performance is clear. Look at color test samples VS1, VS14, and VS15 in particular.

I'd like to show a real example of how misleading CRI metrics can be. **Figure 6** shows two photographs of an identical scene. The image on the left is lit with an RGB light source that has a CRI of 82 and a CQS of 74. The main reason for the difference between CRI and CQS in this case is that the colors are all desaturated, which CRI doesn't penalize as much as CQS. There are also some true rendering errors that both metrics pick up. Under current regulations, this light would pass Energy Star (which requires a minimum CRI value). The image on the right is lit with a different RGB arrangement that has a lower CRI, only 71, but a higher CQS of 83. This is because the colors have been chosen so that they fail by over saturation, or color enhancement, rather than under saturation. CQS allows some over saturation, as it's often unobjectionable, but CRI penalizes



COURTESY AND COPYRIGHT OF WENDY DAVIS – NIST

Figure 6 – Which do you prefer?

it heavily. There are actually fewer color rendering errors in this image than in the first image other than that over saturation. Under current regulations, this light source would fail Energy Star because of a too low CRI. This is clearly nonsense. Which do you prefer? All light spectra that enhance color like this also induce hue shifts. Therefore, a color-enhancing source can never receive a CQS of 100. CQS does not favor color-enhancing sources or those producing over saturation but limits the extent to which they are penalized.

In my opinion, CQS is a much more useful metric than CRI for entertainment lighting. However, even with that said, is CQS alone enough information for us? Unfortunately, no.

The missing information

We still have missing information here. CRI and CQS both give us an indication of how well, or how poorly, a light source will render colors on average, and, if you are using that light source on its own, that may be enough. Unfortunately, what none of these color rendering metrics tell you is precisely where the shortcomings are. You could have two light sources, both with a CRI (or CQS, this problem applies to both) of 85. That level of rendering is acceptable for many purposes, so you would rightly decide that either of them on their own would probably be adequate, not great but adequate, for the job. However, this tells you nothing about how they would look when used together. Perhaps one of them is deficient in red, while the other is deficient in green. Thus, each on their own may be acceptable, but using them together could be quite horrible. In our business, this is a situation you are likely to meet regularly: your wash lights providing fill and side light are from manufacturer A, while your spots providing key light are from manufacturer B. Both products use narrow band LED emitters, and they both have identical CRI of 85. Each on its own does a reasonable job of lighting the performer, but use them together, and the colors on stage change depending on which source is dominant. What is the designer to do? How could they possibly know about this in advance before the moment in the lighting rehearsal when they are both seen together for the first time on the set? It's enough to give you sleepless nights. To some extent, we've seen this problem before, every time we use arc source moving lights in conjunction with incandescent. But now we risk seeing it in every scene with every cue. We seemed to have learned to live with this when using moving lights, but I still think it looks awful!

The 2013 Version 9 of CQS attempts to address this problem by introducing the Color Saturation Icon. This is a small image that is supposed to accompany the raw CQS value in datasheets that indicates in a simple manner where the light source performs well and where badly.

Figure 7 shows the Color Saturation Icon for an incandescent source; the rainbow colored area showing the light's output almost completely fills the outer circle, which represents a perfect source, just failing slightly in the deep blue where we know incandescent is lacking. This works because the CQS test color samples have been deliberately chosen to have equal chroma values. **Figures 8** and **9** show two RGB LED sources, both of which have a CRI of 67 and are therefore indistinguishable by use of that metric alone. However, the icons clearly show that one of them, **Figure 8**, oversaturates red and green, while the other, **Figure 9**, undersaturates them. With this extra information, it's much easier to make a judgment that these two lights would likely not work well together.

I've not seen any manufacturers using the CQS Color Saturation Icon on datasheets yet, but I could see it being a very useful addition. It's easy to calculate and print and conveys a lot of information in a single small image. NASA's Johnson Space Center is considering using these icons when evaluating light sources for use in spacecraft for both lighting and as colored indicators. Color is commonly used to indicate different controls and safety features of spacecraft, and the lighting needs to be good enough to allow you to see those colors. Whether a warning is yellow or orange could be important! Just like us, NASA has had problems when a supposedly good light source with a high CRI has rendered one

Figures 7, 8, and 9

specific critical color poorly and quickly realized the shortcomings of CRI as a measure.

What else?

CRI and CQS may be the most commonly seen color rendering metrics, but there are many, many more out there. Some other suggestions for improving the agreed shortcomings in CRI include metrics that use many more test colors. One current proposal uses many theoretical simple color band samples that overlap to cover the entire spectrum. This would give a much more granular result than either of the current tests. **Figure 10** shows the idea with multiple overlapping colors. None of these pure colors actually exist as pigments in real life, of course, but that doesn't matter, as all tests are done in the computer. However, even with this granularity, we still wouldn't know from a single number metric where any failures in rendering were.

Color Gamut Index (CGI) is another strongly supported metric for color rendering. This time, it is proposed as an addition to CRI or CQS and not as a replacement for them. The proponents of CGI understand and recognize the strengths of CRI and CQS as well as their inherent shortcomings in being unable to distinguish between under and over saturation. Both CGI and CQS produce a drop in the metric with either of these situations, but you don't know which it was that caused the change. CGI is a means of quantifying the test shown by the CQS Color Saturation Icon. It takes the same color samples as either CRI or CQS and plots the resultant colors when illuminated by the test source on a u',v' chromaticity chart. This creates polygonal shapes whose boundary is the gamut of the test colors when lit with that source. The area of these polygons, or Color Gamut Area (CGA), represents how saturated the produced color range is and thus gives a measure of the under or over saturation of the source. In general, the larger the gamut area, the more saturated the color samples are, and the easier it is for the human eye to discriminate between them. The CGI is then derived from the CGA by comparing it to the CGA of a perfect source of the same color temperature. This, combined with the CRI or CQS, gives a better feel for the overall performance of the light. However, yet again, even these two numbers can't tell you in which colors the light performed well or badly, so I'm not convinced it provides everything we need in our industry. **Figure 11** shows example Color Gamut Areas for various light sources.

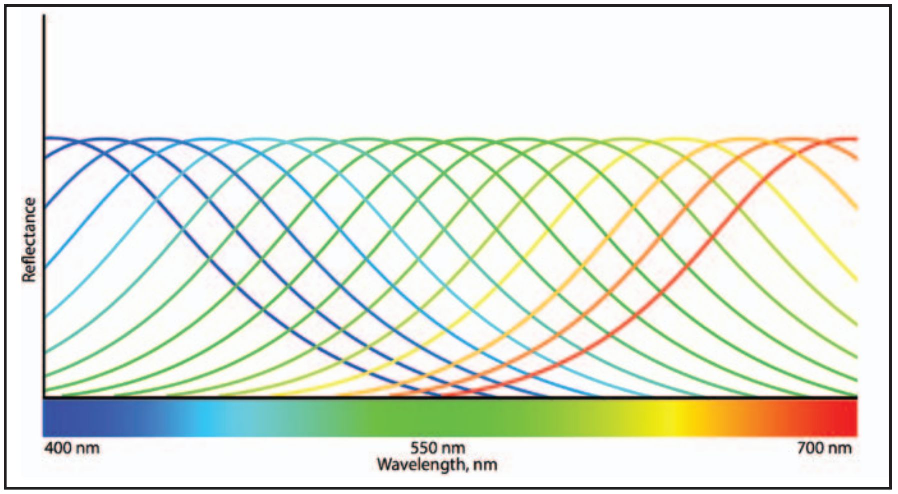


Figure 10 – Multiple overlapping color test samples

Unlike the other metrics we've talked about, Color Gamut Index (CGI) can actually be too high as well as too low, so there will be an acceptable range of CGI rather than just a lower bound. The currently suggested range of acceptability for normal lighting is to have a CGI more than 80 but less than 100 coupled with a CRI (or CQS) above 80. (I recommend the work done by the Lighting Research Center at Rennsselaer Polytechnic Institute for more information on CGI and its proposed use.)

Other color rendering metrics being considered include Full Spectrum Color

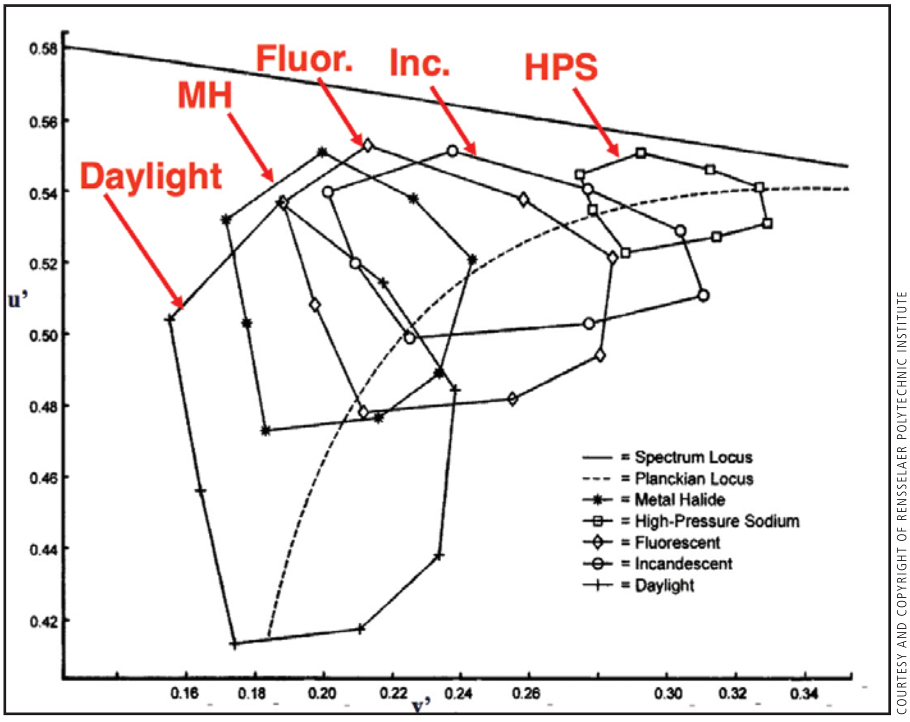


Figure 11 – Color Gamut Area

Index (FSCI), Color Rendering Map (CRM), Color Discrimination Index (CDI), Color Rendering Capacity (CRC), Feeling of Contrast Index (FCI), Flattery Index (FI), Color Preference Index (CPI), and many others. The goal of them all is essentially the same: to provide a single figure (or diagram) that rates a light source on how well it renders color to the human eye. What the word "well" means here varies from metric to metric, but in general, it is a comparison of the test light source to a perfect hypothetical light source, with daylight often used as the example. They all do the job, at least to some extent, but at the moment, the world cannot agree on which, if any, of these to use. There have been numerous international committees throughout the last few years trying to find a replacement for the flawed CRI, and, so far, every committee has broken up without agreement. (A notable exception being the PLASA Technical Standards Program. We determined what to use for our industry, but we can't wait forever on others to decide!) There are too many vested interests at work to achieve consensus: the manufacturers have commercial concerns and don't want a metric that makes their products look bad, the academics won't agree to anything that isn't perfect (which, with a metric that relies on psychology and statistics, is probably impossible to achieve), and end-users and designers just want something—anything—that helps them today.

> " . . . end-users and designers just want something—anything—that helps them today."

So where does that leave us? Everyone agrees that CRI is useless, but because we can't agree on a replacement, the useless CRI is perpetuated as the only codified color rendering metric. Sadly, I'm not sure the decisions we make in our tiny industry can make much difference here, but it's important to at least know what's going on and realize that (as I seem to say so often in these articles) the only true judge of color rendering right now is your eyes. ■

Mike Wood runs Mike Wood Consulting LLC, which provides consulting support to companies within the entertainment industry on product design, technology strategy, R&D, standards, and Intellectual Property. A 35-year veteran of the entertainment technology industry, Mike is the Immediate Past Chair of the PLASA Governing Body and Co-Chair of the Technical Standards Council. Mike can be reached at mike@mikewoodconsulting.com.