Rationale of Color Quality Scale

Yoshi Ohno and Wendy Davis, NIST

June 8, 2010

The purpose of this article is to solicit comments from the solid state lighting (SSL) industry on the Color Quality Scale (CQS) developed by National Institute of Standards and Technology (NIST). The CQS is being discussed as one of the proposals in the International Commission on Illumination (CIE) Technical Committee (TC) 1-69 and final discussions are in progress in the TC toward selecting a new metric. Since the SSL industry is not well represented in the TC, we hope to convey the opinions of the SSL industry to the committee members.

The color quality of SSL products is critical and is the subject of increasing attention. The CIE Color Rendering Index (CRI) [1] has been widely used for many years. However, the CRI is 40 years old and various problems with the CRI when used for light-emitting diode (LED) sources have been identified, as reported in many publications [2-7]. In particular, the report from CIE Technical Committee TC 1-62 “Colour rendering of white LED light sources” [8] summarizes several problems of the CRI when applied to white LED sources. The CRI score[1] does not correlate well with visual evaluation in many cases. In response to the conclusions of TC 1-62, a technical committee TC 1-69 (Colour rendition by white light sources, Chair: W. Davis) was established in 2006 to develop and recommend a new CIE metric. The new metric is intended to eventually replace the CRI (not immediately after publication of the new standard). It should work well for both traditional lighting technologies and SSL sources. At NIST, we have developed the CQS [9-11], which addresses the problems of the CRI for SSL sources yet maintains good consistency of scores with the CRI for traditional sources, and proposed it to TC 1-69. Our spreadsheet to calculate CQS has been distributed to many users in the SSL industry and continues to gain support in the USA. However, the CQS is currently not gaining good support in the TC. The TC is inclined to adopt a different metric, which is a pure fidelity metric very similar to CRI. The traditional lamp industry also has concerns that the concept underlying the CQS deviates from the CRI. We are concerned that another pure fidelity metric will not solve the problems for SSL sources, and we believe that a new concept metric, such as the CQS, is needed for the SSL industry.

Fixing the problems of the CRI is critical for the SSL industry. Traditional lamp companies have known about the problems of CRI for a long time and have found ways to work around these problems. The SSL industry is different. It is young and consists of numerous small companies. These companies often do not have expert knowledge on color rendering and colorimetry. They tend to design products that simply maximize the lumens per watt and the score of the color rendering metric. As products are optimized for the metrics, standards are important.

For those who would like to review the issues, we explain below the problems of CRI and how CQS works to solve the problems. We’d appreciate it if you could write to us your comments on the CQS and any experiences you have with CRI and/or CQS for SSL products. Please contact ohno@nist.gov.

[1] In many cases in this document, “CRI” or “CRI score” is used to refer to the R_a score.

Rationale of CQS

Color fidelity is only one aspect of color quality

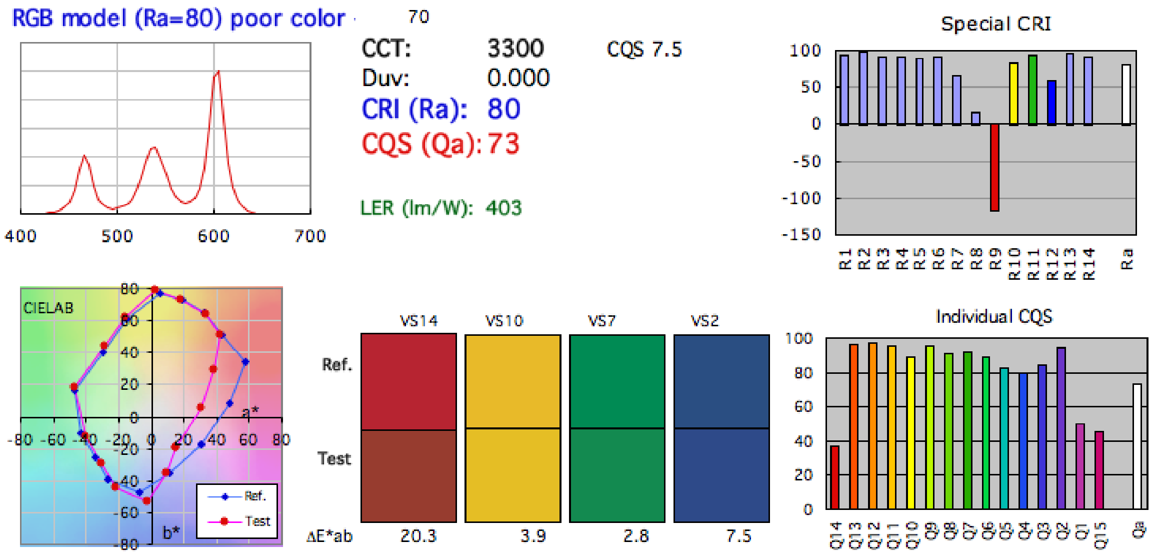
The CRI assesses the fidelity of color rendition (color fidelity) by calculating the color differences of a set of pre-defined test samples under illumination by a test source and a reference illuminant. It has long been known that color fidelity cannot serve as a sole indicator of color quality of white light. There are other aspects of color quality; in particular, object color shifts that increase color saturation can enhance visual preferences, color discrimination ability, and visual clarity [12-15], improving the subjective rating of color appearance of the illuminated objects. It is also known that perceived color saturation of objects is reduced at lower illuminance levels (Hunt Effect). Therefore, it seems that sources that enhance color slightly in indoor lighting may achieve higher color fidelity to real daylight [16]. There have been attempts to account for these effects by various proposed metrics such as the Flattery Index [12], the Color Preference Index [13], the Color Discrimination Index [14], the Feeling of Contrast Index [15], the Color Rendering Capacity [17], and the Gamut Area Index [5]. None of them, however, has been accepted as a standard, probably due to the fact that CRI has been so widely used (regardless of the problems) and that the industry refuses to use more than one metric to rate color rendering of products. The CQS is basically a fidelity metric, producing one number like R_a, but the CQS has improvements including the ability to account for the direction of object color shifts. The CQS thus agrees well with visual ratings for color-enhancing sources, while it works as an accurate fidelity metric for non-color-enhancing sources. The CQS score works to represent the overall color quality of products (perceptual color fidelity) for all types of light sources.

A good CRI score does not guarantee good color quality

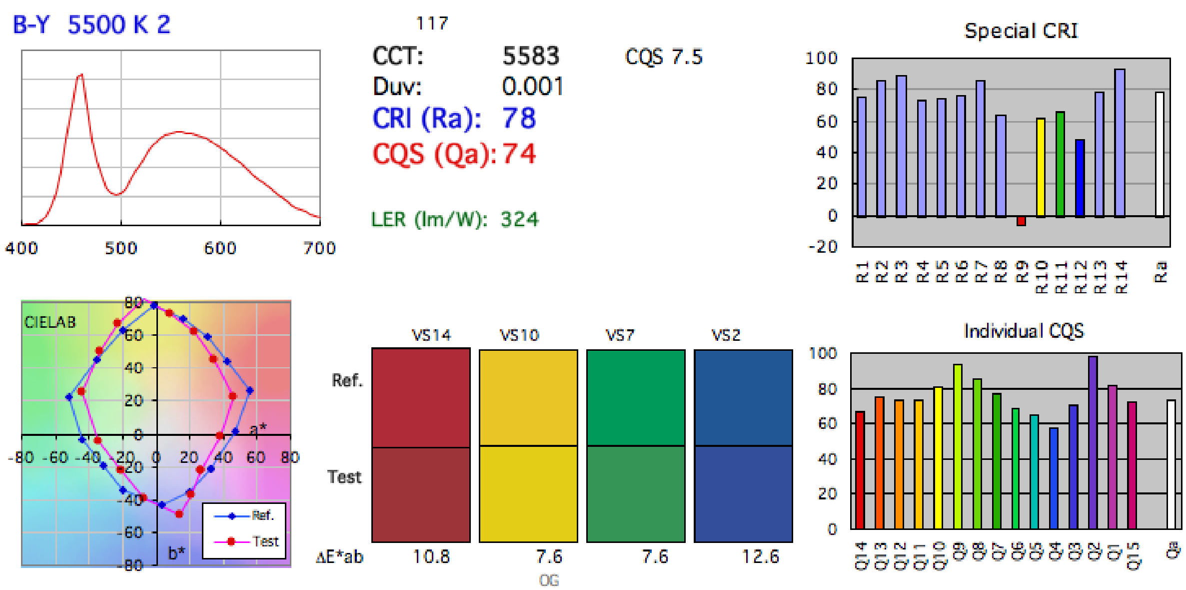
One problem with the CRI is that it can give fairly high scores to sources that render some saturated object colors very poorly. An example below shows a case of a red-green-blue (RGB) white LED simulation optimized for highest luminous efficacy of radiation (LER, lm/W) at CRI R_a=80. This light would qualify for Energy Star and achieve high luminous efficacy[2], but the actual color rendering is very poor. With R_9 = -114, skin tones under this light look pale and unhealthy. The CQS lowers the score of this lamp by seven points from CRI, making it ineligible for Energy Star (LED luminaires). The other leading proposal within the CIE TC, CRI-CAM02UCS, gives this light a score of 81, one point higher than the CRI.

The color patches shown in the figure are four of the 15 test samples used in the CQS. The upper boxes show the colors illuminated by the reference illuminant (in this case, blackbody radiation), and lower boxes show the colors lit by the test source. The presented colors are determined by colorimetric simulation [4].

[2] Referring to theoretical efficacy, LER=403 lm/W here, not the physical luminous efficacy of the source. RGB spectra can be produced not only by RGB(A) chips but also by green phosphor LED, UV excited narrowband phosphors, or quantum dots. Narrowband white light is theoretically more energy efficient than broadband spectra.



Though not as serious as the case above, the same problem with saturated object colors can occur for phosphor type white LEDs. The example below is a case where the CRI is nearly 80 (78) but color rendering of red objects is poor (R_9 is negative). The blue hue shift (to violet) is also significant. The CQS reduces the score of this particular light by four points. (But, note that CQS scores for most phosphor type LEDs do not change much from CRI - see the examples at the end of this document).



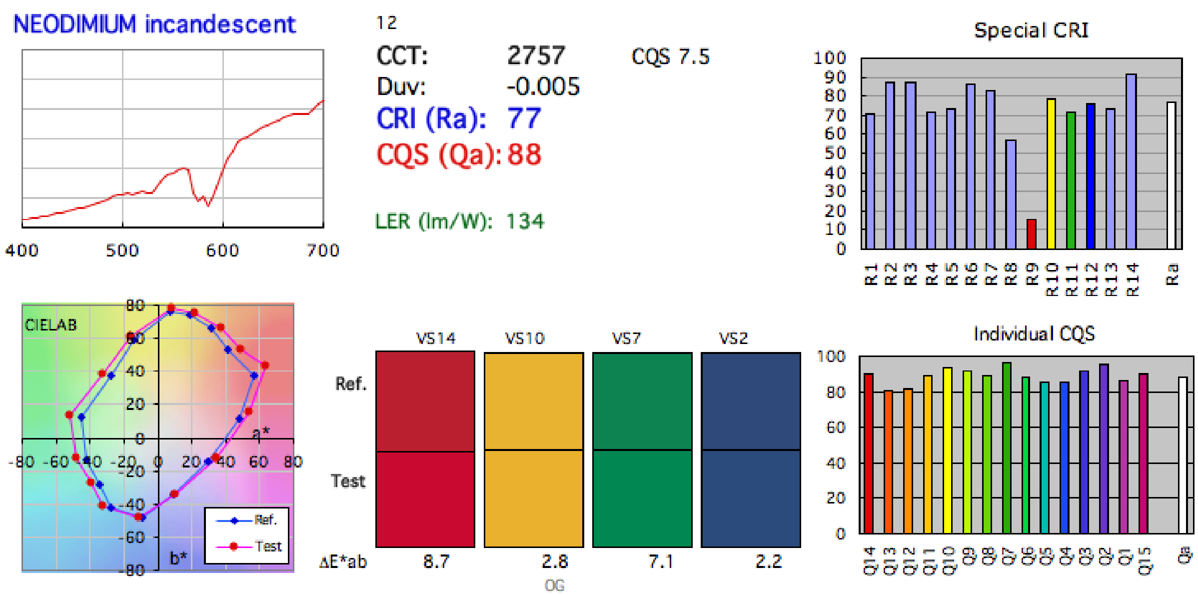
These problems occur because the CRI uses pastel color test samples, and fails to detect problems with saturated colors. Due to this problem with the CRI, some regulatory specifications[3] add requirements for R_9 and even R_{10}, R_{11}, R_{12} (special CRIs for saturated color

[3] Energy Star for integral LED lamps specifies the minimum R_9 value as well as R_a. DOE Energy Alliance Specification (LED refrigerated show case) lists the minimum average scores for R_9 to R_{12}.

samples) in addition to R_a. But, normally only the R_a value is communicated to end-users. Generally, industry and consumers do not want to deal with multiple numbers for color rendering ratings. The CQS uses saturated color samples as test samples to calculate the score, thus its one number output (Q_a) represents the color rendering quality for all colors including saturated colors. The CQS is also designed so that its score is notably reduced if one or two colors appear very poorly even when all other colors are rendered well[4] (as is the case of the RGB LED example above).

CRI penalizes color-enhancing lights

As mentioned above, the CRI score does not correlate well with visual evaluation when the chroma (color saturation) of an object is increased by the light source. This is well demonstrated by the case of a neodymium lamp[5], as seen in the example below. This lamp produces a slight increase of color gamut, and the color contrast of objects is slightly enhanced. Object colors under this light look vivid and beautiful, and this lamp is often preferred over normal incandescent lamps (R_a=100). But this lamp is given a CRI score of only 77. Visual evidence suggests that this lamp deserves a much higher score. The CQS gives this lamp a score of 88.

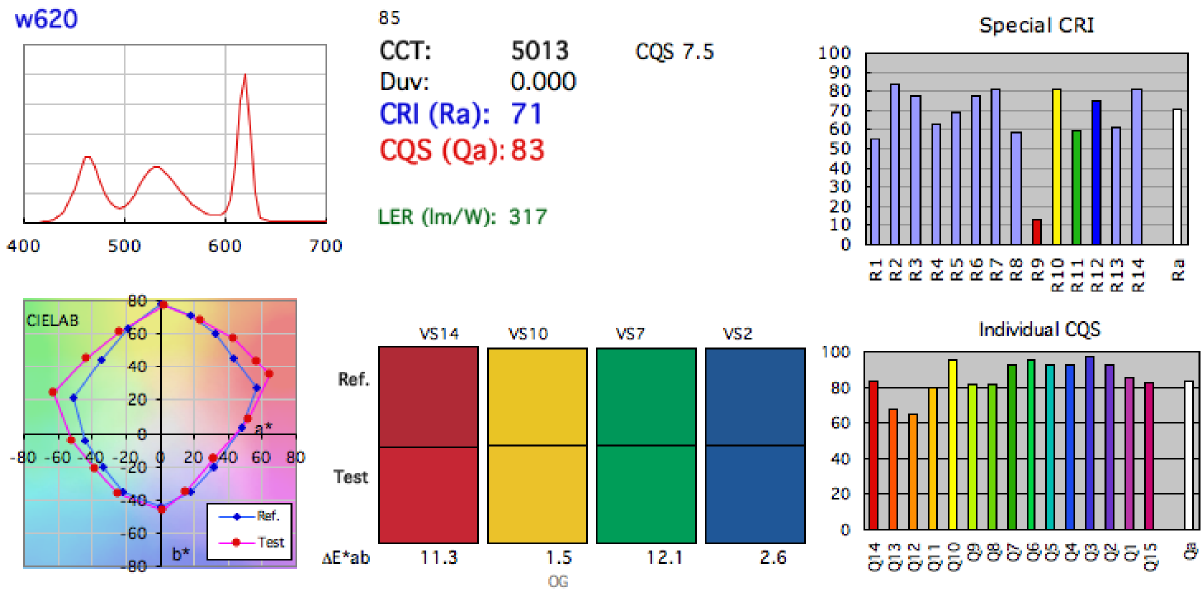


---

[4] This is achieved by using root-mean-square (RMS) averaging of color differences of saturated test color samples.

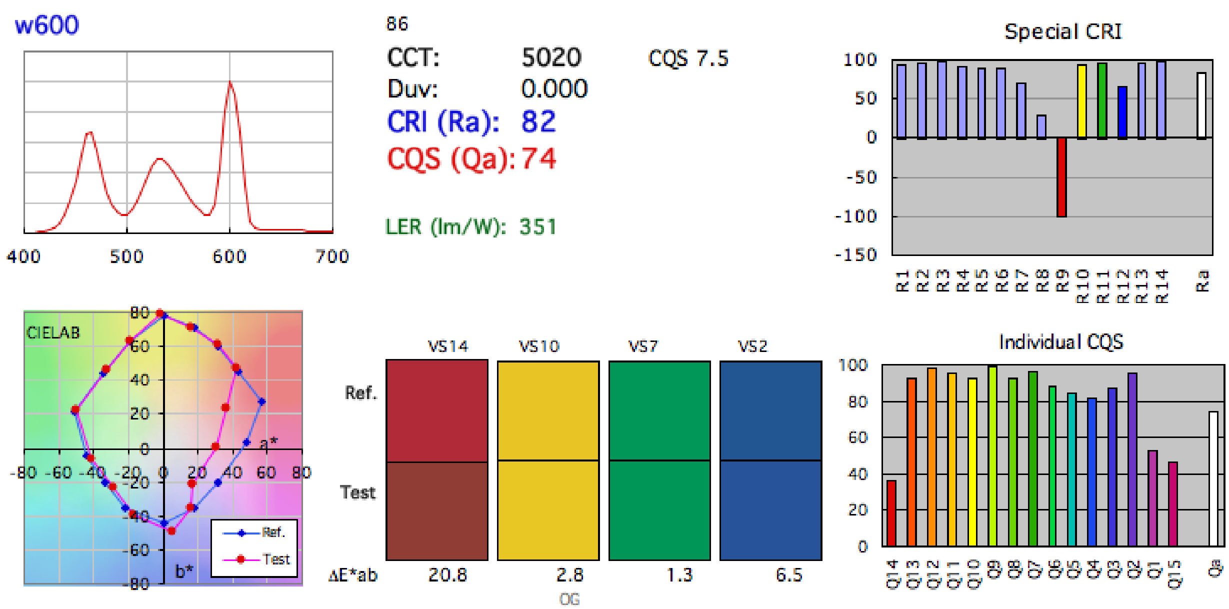
[5] Incandescent lamp with its glass envelope doped with neodymium, which absorbs much of the yellow portion of the lamp spectrum. This type of lamp is popularly sold in the USA and other countries.

Narrowband RGB spectra tend to produce similar color-enhancing effects. See the example below of a cluster of real high power RGB LEDs used in the NIST color rendering booth.



This light is given the CRI score of only 71, but we visually verified that objects under this light looked beautiful and still natural. This light source deserves a much higher score.

We also built another combination of RGB LEDs for the booth as shown below, by changing only the red peak. Below is a color-desaturating example. This light has R_a=82 but we verified that objects under this light looked very poor.

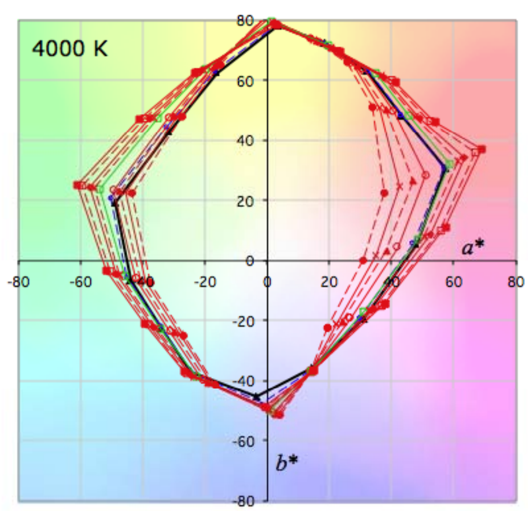


The two sources in the above examples were visually compared as shown in the pictures below (though colors in the pictures are probably not accurately reproduced). We verified that the color-enhancing light (b) was more acceptable than the color-desaturating light (a) in spite of the CRI scores. Note that (a) would be accepted by Energy Star and (b) would not, but that would be reversed if CQS is used.

(a) Color-desaturating light (R_a=82, Q_a=74)

(b) Color-enhancing light (R_a=71, Q_a=83)

The effects of the degree of color saturation have been studied further in the NIST Spectrally Tunable Lighting Facility [18, 19] (picture below). Visual evaluation experiments were done with several subjects to rate the color rendering of the objects in the room and their skin tones under illumination by spectra that induced various levels of object color-desaturation or saturation. The results showed clearly that subjects did not accept de-saturating lights, and accepted saturating lights much better [19]. For example, the subjects rated the slightly saturating light (R_a=82, Q_a=90) equally high to the broadband reference light having Ra=98. This experiment has demonstrated how important it is to take into account the direction of color shifts in color rendering evaluation.



CQS maintains consistency with CRI for traditional lamps

The CRI has been widely used in the lighting industry for over 40 years, and traditional lamp manufacturers are concerned about any changes in scores from the CRI by a new metric for existing lamp products, in particular, triphosphor fluorescent lamps [20]. The CIE's attempts to revise the CRI have failed several times since the 1980s, most recently, by TC1-33 [21]. The CQS is designed to solve the problems of CRI while maintaining consistency with the CRI's scores for existing lamp products.

Other improvements by CQS

The CQS includes several other improvements over the CRI. People often experience that they cannot distinguish navy blue socks and black socks under incandescent light. This is because short-wavelength (blue) energy is reduced at low correlated color temperatures (CCTs) and the gamut area of rendered object colors shrinks. The CRI, for example, would give a perfect score Ra=100 to a 2000 K blackbody radiator but such reddish light would not be acceptable for general lighting. The CQS introduced a "CCT factor" based on the color gamut area of the reference illuminant, to account for this effect and reduce scores for very low CCT sources. This factor affects only sources having CCTs lower than 3000 K. For example, 2800 K incandescent lamp scores (and any other sources having the same CCT) are reduced by only two points.[6]

The CQS also replaces outdated formulae in the CRI. The object color space used in the CRI (W*U*V*) is very non-uniform and does not accurately measure object color differences. It is replaced by the latest CIE standard color space, CIELAB. The old chromatic adaptation formula used in the CRI (von Kries) is also replaced by CMCCAT2000. As a result, the CQS slightly penalizes the lights that have chromaticity with significantly positive Duv (above the blackbody locus – yellowish or greenish light), while CRI does not. The scores agree well with our visual experience that greenish lights are less acceptable than pinkish lights having the same magnitude of Duv shifts.[6]

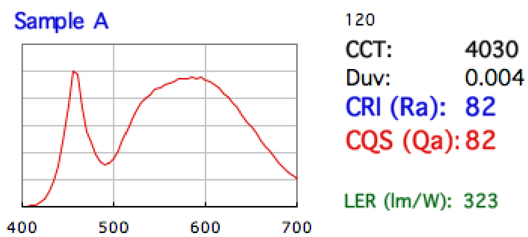
In addition to one number output, Q_a, the CQS has individual scores (Q_1 to Q_{15}) for each saturated test color sample, which are available for expert users. It is possible to add more samples (like skin colors as was done in the CRI), which would not be used in the calculation of Q_a.

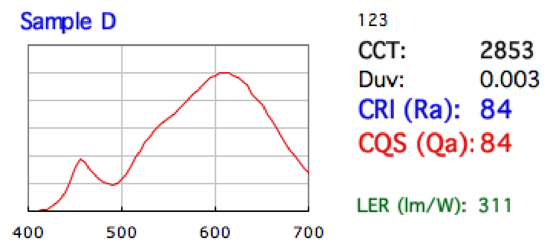
The CQS also introduced a 0 to 100 scale, as the CRI is often confusing because it produces negative scores. The CQS gives a score of Q_a=0 for lights with no color rendering (e.g., low pressure sodium lamp). This change affects only the scores for very low color quality sources, mostly sources with R_a lower than 30.

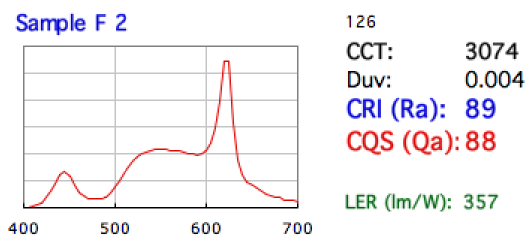
[6] The CCT factor and the new chromatic adaptation formula resulted in reduced scores for low CCT lamps and high Duv lamps, though slightly. This increases score deviation from CRI for some existing lamp products. If the current version of CQS is not acceptable by lamp manufacturers, these factors and formula can be removed or replaced, without losing the other functionality of the CQS.

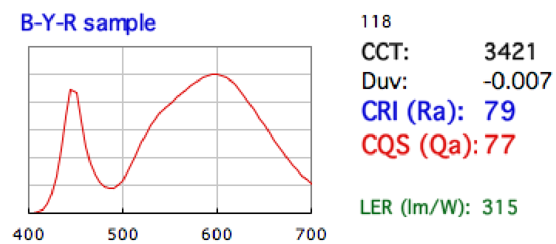
More examples of comparison of CRI and CQS

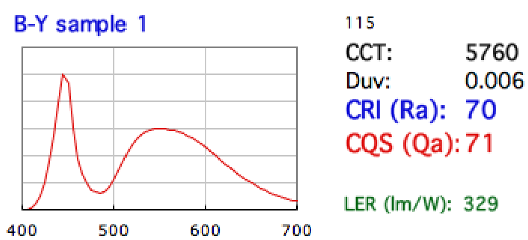
The CQS provides scores consistent with the CRI for most recent phosphor type LED products and traditional discharge lamps.

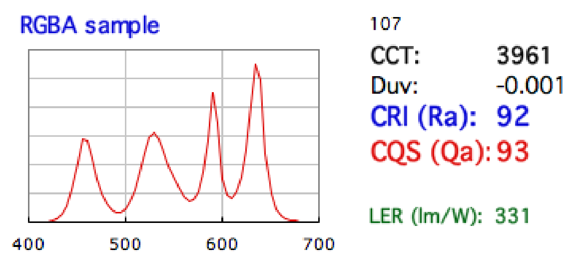


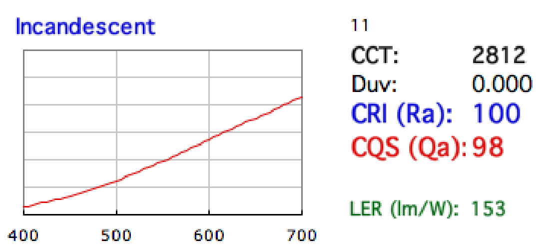


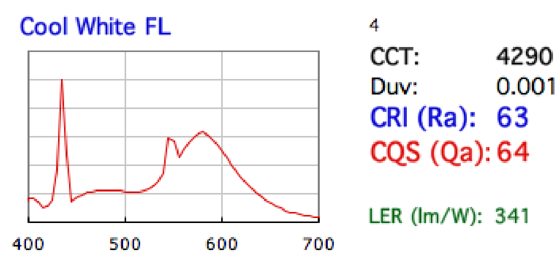


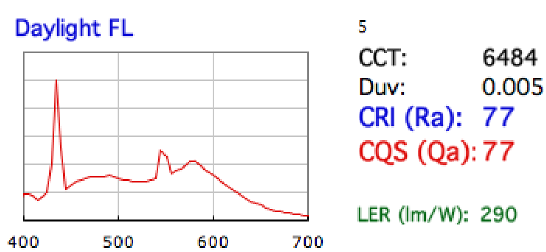


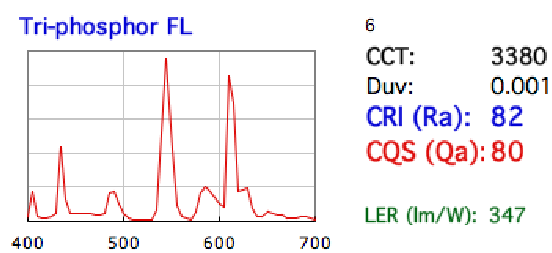


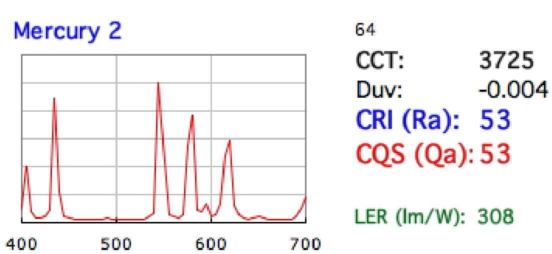


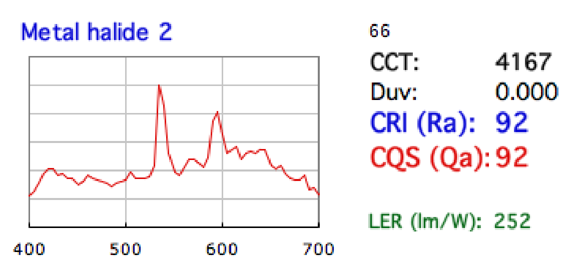












References

[1] CIE 13.3:1995, “Method of measuring and specifying colour rendering properties of light sources,” (1995).

[2] N. Narendran and L. Deng, “Color rendering properties of LED sources,” *Proc. SPIE*, vol. 4776, pp. 61–67, 2002.

[3] P. Bodrogi, P. Csuti, P. Horv´ath, and J. Schanda, “Why does the CIE color rendering index fail for white RGB LED light sources?,” in *Proc. CIE Expert Symp. LED Light Sources: Phys. Meas. Visual Photobiol. Assess.*, 2004, pp. 24–27

[4] Y. Ohno, “Spectral design considerations for white LED color rendering,” *Opt. Eng.*, vol. 44, no. 11, pp. 111302-1–111302-9, Nov. 2005.

[5] M. S. Rea and J. P. Freyssinier-Nova, “Color rendering: A tale of two metrics,” *Color Res. Appl.*, vol. 33, no. 3, pp. 192–202, Jun. 2008.

[6] A. Zukauskas, R. Vaicekauskas, F. Ivanauskas, H. Vaitkevicius, P. Vitta, M. Shur, Statistical Approach to Color Quality of Solid-State Lamps, IEEE Journal of Selected Topics in Quantum Electronics, Vol. 15. No. 6, 2009.

[7] W. Davis and Y. Ohno, “Approaches to color rendering measurement,” *J. Modern Optics*, Vol. 56, No. 13, 1412-1419 (2009).

[8] CIE 177:2007, “Colour rendering of white LED sources, 2007.

[9] W. Davis and Y. Ohno, “Toward and improved color rendering metrics,” *Proc. SPIE*, vol. 5941, pp. 59411G-1–59411G-8, 2005.

[10] Davis, W. and Ohno, Y., “Development of a Color Quality Scale (CQS),” *Sixth International Lighting Research Symposium on Light and Color*, Orlando, FL (2006).

[11] W. Davis and Y. Ohno, “The Color Quality Scale”, *Optical Engineering*, 49(3): 033602 (2010).

[12] D. B. Judd, “A flattery index for artificial illuminants,” *Illum. Eng.*, vol. 62, pp. 593–598, Oct. 1967.

[13] W. A. Thornton, “A validation of the color-preference index,” *J. Illum. Eng. Soc.*, vol. 4, pp. 48–52, Oct. 1974.

[14] W. A. Thornton, “Color-discrimination index,” *J. Opt. Soc. Amer.*, vol. 62, no. 2, pp. 191–194, Feb. 1972.

[15] K. Hashimoto and Y. Nayatani, “Visual clarity and feeling of contrast,” *Color Res. Appl.*, vol. 19, no. 3, pp. 171–185, Jun. 1994.

[16] Davis, W. & Ohno, Y., “Studies on the effect of illuminance on color rendering,” *Proc. CIE 2009: Light and Lighting Conference*, Budapest, Hungary (2009).

[17] H. Xu, “Color-rendering capacity of illumination,” *J. Opt. Soc. Amer.*, vol. 73, no. 12, pp. 1709–1713, Dec. 1983.

[18] Miller, C.C., Ohno, Y., Davis, W., Zong, Y., and Dowling, K., “NIST spectrally tunable lighting facility for lighting vision science experiments,” CIE peer-reviewed Paper publication from CIE 2009 Light and Lighting Conference, Budapest, 2009.

[19] Y. Ohno and W. Davis, “Visual Evaluation Experiment on Chroma Enhancement Effects in Color Rendering of Light Source”, submitted to CR&A (being reviewed).

[20] Report by Japanese national committee of TC1-69, submitted to the TC on Jan. 15, 2010. Also, private communication from Philips in USA.

[21] CIE135/2:1999, Colour rendering, TC 1-33 closing remarks, CIE Collection in vision and colour and in physical measurement of light and radiation (1999).