Evaluation of Light Measurement Instruments

Bruce Leigh Myers, Ph.D.

Keywords: Bland-Altman plot, Tukey mean difference plot, light measurement, correlated color temperature, standardized viewing, color rendering index

The International Standards Organization (ISO) defines standards to ensure that viewing conditions are consistent when evaluating printed samples through ISO 3665 (2009), Graphic technology and photography – viewing conditions. The need for this standard stems from the necessity for human visual assessment as the key arbiter of the quality of complex images, and the tendency for various lighting conditions to shift the appearance of a color, specifically in relation to other adjacent colors.

Among the conditions specified by ISO 3665 (2009) are Correlated Color Temperature (CCT) and Color Rendering Index (CRI). Berns (2000) describes CCT as "Temperature, usually expressed in kelvins, of a blackbody radiator that most closely resembles the color of a stimulus of equal brightness" (p. 4). Color rendering is described by Field (2004) at "…the degree to which a test illuminant (e.g., fluorescent light) renders colors similar in appearance to their appearance under a reference daylight illuminant of the same color temperature" (p. 4). A method for determining CRI is defined by the Commission Internationale de l'Éclairage (CIE) for a given light source. Field (2004) states: "The optimal CRI (that for daylight, or for such continuous sources as tungsten lamps) is given as 100" (p. 4). CRI is expressed as CRI Ra, with Ra representing the international standard for CRI as defined by CIE 13.3-1995.

Both CCT and CRI Ra are quantifiable by a range of instruments, including traceable Spectroradiometers specifically designed for the purpose, general-use Spectrophotometers that can read CCT and CRI Ra, and handheld instruments designed for photographic applications that measure CCT.

The present study seeks to compare readings from a traceable Spectroradiometer with those from various other meters across a range of seven viewing booths, some of which are known to be out of specification. The goal is to ascertain how much variance can be expected when using these varied meters when compared to a

Rochester Institute of Technology

traceable benchmark instrument. For the purposes of this study, the benchmark instrument is referred to as the reference instrument, and the other measurement devices are the test instruments. The instruments used in this study are detailed on Table 1.

Meter	Metrics Measured	
Reference Instrument		
SpectriLight ILT950 Spectroradiometer International Light Technologies, NIST Traceable/ISO17025 Accredited Calibration	CCT	CRI Ra
Test Instruments		
X-Rite i1 Pro 2 Spectrophotometer	CCT	CRI Ra
Minolta Color Meter IIIF	CCT	
LUXI and Cine Meter II App / Mobile Phone Based Solution	CCT	
Gossen SixtiColor (1957)	CCT	

Table 1. *Measurement Instruments*

To determine the variability of the seven viewing booths used, at least 10 readings were taken with the SpectriLight ILT 950 in each booth, as illustrated in Table 2. As ISO 3665 (2009) mandates a CRI Ra value over 90, it is noted that booths 1 and 2 are in compliance, booths 3,4, and 5 are nearly out of specification for that metric, and booths 6 and 7 are well out of compliance. This level of variance represents a range of variability that provides a means to compare the instruments tested across dissimilar viewing conditions.

Booth #	*n*	CCT Mean	CCT *SD*	CRI Ra Mean	CRI Ra *SD*
1	30	4844.73	18.92	94.58	0.17
2	30	4816.03	10.47	96.28	0.06
3	10	4843.30	33.67	92.58	0.46
4	10	5024.80	11.23	91.29	0.14
5	10	4751.80	75.20	90.32	1.30
6	10	4589.20	32.55	83.18	0.55
7	10	4589.50	39.67	82.29	0.44

Table 2. *Mean and Standard Deviation of CCT and CRI Ra of Light Booths as measured with ILT950*

For a test instrument, measurement technique cannot be considered accurate unless measurements of a particular variable by the test instrument agrees closely with a reference instrument across all applied instances. A graphical approach to analyzing the comparison of a test and reference method that addresses these concerns as advanced by Bland and Altman (1986) and is referred to as the Bland-Altman (B-A) plot, and alternatively known as the Tukey Mean-Difference plot. Bland and Altman are credited with popularizing the use of this technique, and in the words of Earthman (2015): "They did not invent the method, but they advocated its application to the comparison of medical devices, laboratory tests, and other clinical techniques to ascertain bias in one method compared with another" (p. 794).

A B-A plot illustrates the mean difference between the two methods on the x-axis, and difference between paired readings of the two methods on the y-axis includes calculations of limits of agreement (LOA) when applicable (typically described as mean difference +/- 1.96 standard deviation to represent 95% confidence).

The procedure for comparison is to first calculate the difference between the two methods as a new variable, and then to conduct a one sample *t*-test on this result to examine a potential systematic bias. When examining CRI Ra, 95% LOAs are calculated to visually analyze how far apart measurements are likely for most applications; these LOAs are determined by multiplying the standard deviation of mean difference by 1.96, and then adding/subtracting the resulting value from the mean difference.

Comparison of the ILT950 with the X-Rite i1 Pro 2

For this analysis, three different i1 Pro2's were used with the results averaged to compare to the ILT950. One hundred ten readings for each instrument were taken across the viewing seven booths. CCT and CRI Ra metrics were evaluated.

Correlated Color Temperature

A one-sample t-test was conducted to ascertain if the mean difference is significantly different than zero. The mean difference between the ILT950 and the average of the i1Pro 2 instruments (M =138.67, SD = 72.45) was significantly higher than zero $t(109) = 20.03$, $p < .001$.

A visual evaluation of the B-A plot shown in Figure 1 suggests that variation is dependent upon the magnitude of measurement. Readings nearer to 5,000 kelvins appear to have less variance than those further below, and readings above the 5,000 kelvins mark appear to be nearer to zero as shown by the dashed line in the plot. Across a range of viewing conditions, however, the results suggest that these units should not be used interchangeably. It is noted that the bias is not consistent, therefore subtracting a constant value from the i1Pro 2 readings to match the

ILT950 would not be recommended. Further, the presence of proportional bias is not clear, the variability recorded is evidently inconsistent.

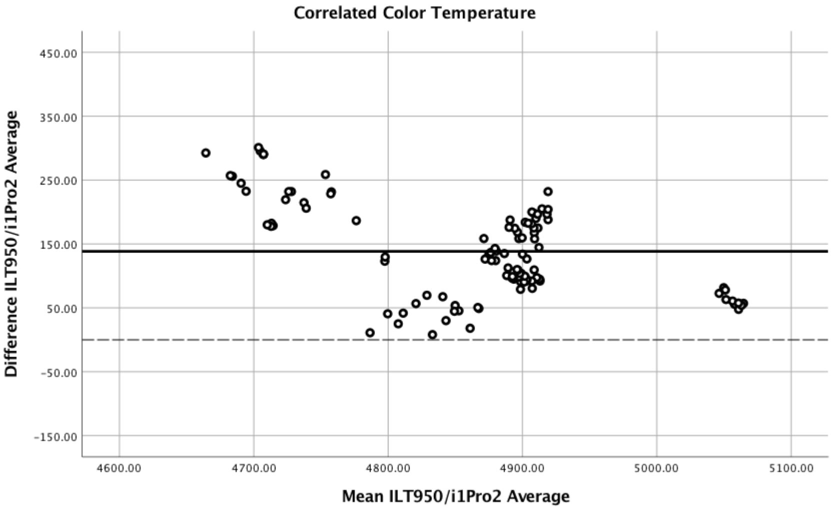

Figure 1. *Bland-Altman Plot comparing CCT readings from ILT950 with the average of the X-Rite i1Pro 2 across the range of viewing booths. The bold solid horizontal line indicates the mean difference, the dashed line indicates zero.*

Color Rendering Index

Turning to an analysis of the CRI Ra values obtained in the comparison of the reference ILT950 and the average of the i1 Pro 2's, a one-sample t-test was used to determine systematic bias. In this instance, the mean difference (M = 1.10, SD = 0.84) was higher than zero $t(109) = 13.73$, $p < .001$. This suggests that the test i1 Pro 2 consistently measures CRI Ra higher than the reference ILT950.

The B-A plot, as shown in Figure 2, indicates that variation is dependent upon the magnitude of the measurement, with variation decreasing as the values approach the ideal 100. Inconsistency in the apparent variability, together with the observation that several readings are beyond the 95% LOA, suggest that these devices should not be used interchangeably.

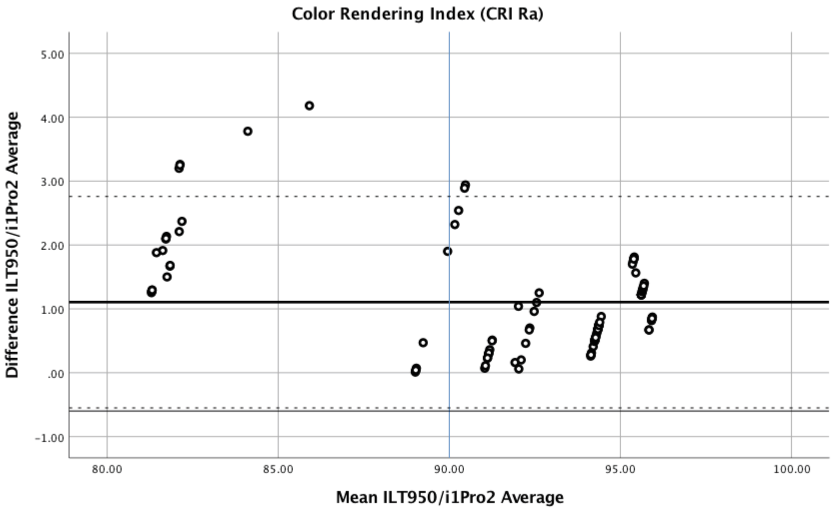

Figure 2. *Bland-Altman Plot comparing CRI Ra readings from ILT950 with the average of the X-Rite i1Pro 2 across the range of viewing booths. The bold solid horizontal line indicates the mean difference, the dashed lines indicate the 95% limits of agreement.*

Comparison of the ILT950 with the Konica Minolta Color Meter IIIF (CMIIIF)

The CMIIIF is a hand held color meter designed for photographic and cinema graphic applications. It is A battery-operated, simple and easy to use instrument, and measures CCT but not CRI Ra.

Correlated Color Temperature

In comparison with the Spectroradiometer ILT950, the CMIIIF exhibited systematic bias (M = 400.51, SD = 436.10) with a mean difference greater than zero t(69) 7.68, p < .001. An examination of the B-A plot, once again the observed variation is dependent upon the magnitude of measurement. As the kelvins exceed 5,000, the difference tends to increase, suggesting that no constant value subtracted from the test instrument values across the range of measured conditions would allow these devices to be used as substitutes for each other in practical applications across a variety of viewing conditions.

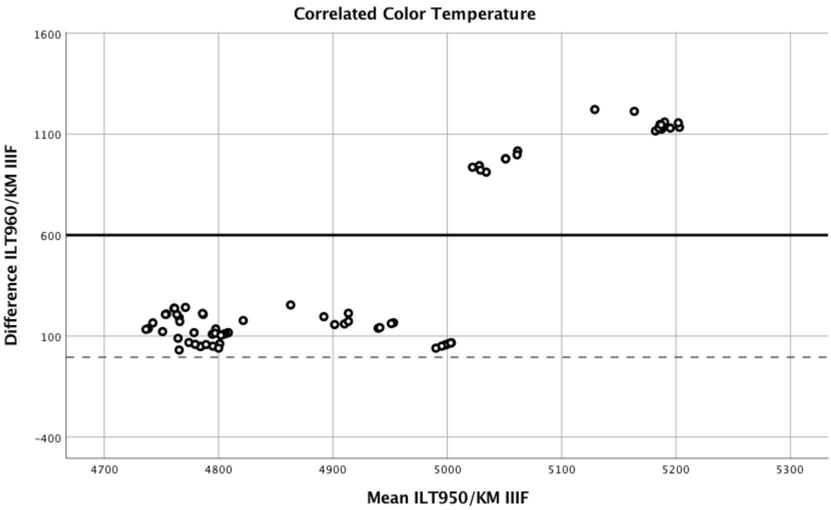

Figure 3. *Bland-Altman Plot comparing CCT readings from ILT950 with the Konica Minolta CMIIIF across the range of viewing booths. The bold solid horizontal line indicates the mean difference, the dashed line indicates zero.*

Comparison of the ILT950 with the LUXI For All and Cine Meter II

The LUXI For All is an attachment for mobile devices, including smart phones and tablets that feature a front-facing camera. The attachment acts as a diffuser, allowing the camera to be used as a light meter. The LUXI For All comes with an app that is functional as an exposure meter, but when combined with the Cine Meter II iOS app, which has the ability to read CCT among other attributes, the combination was used to compare the benchmark ILT950 Spectroradiometer. Like the Konica Minolta Color Meter IIIF, this solution is designed for photographic and cinema graphic applications. It is hoped that as a low-cost alternative to the other solutions examined, the LUXI and Cine Meter II would be an interesting test in the present context.

It is important to note that the Cine Meter II app recommends that the solution be adjusted to comply with another device of known accuracy. It is therefore implied that this solution is not intended to be a primary meter, but rather a low-cost supplemental solution to another instrument. In the current study, the LUXI and Cine Meter II were adjusted to comply with the ILT950 for a reading near 5,000 kelvins.

Correlated Color Temperature

A one-sample *t*-test was performed to examine if the mean difference between the two solutions was significantly different than zero across the range light booths in the present study. The mean difference between the ILT950 and the LUXI with the

Cine Meter II app (M =168.03, SD = 112.62) was significantly higher than zero $t(69) = 12.48$, $p < .001$.

A visual evaluation of the B-A plot in Figure 4 suggests lower variance near 5,000 kelvins; perhaps this is not surprising as this is where the devices were correlated. Across the range of light booths examined, the results indicated that the agreement between these two instruments is not reliable, with inconsistent biases noted.

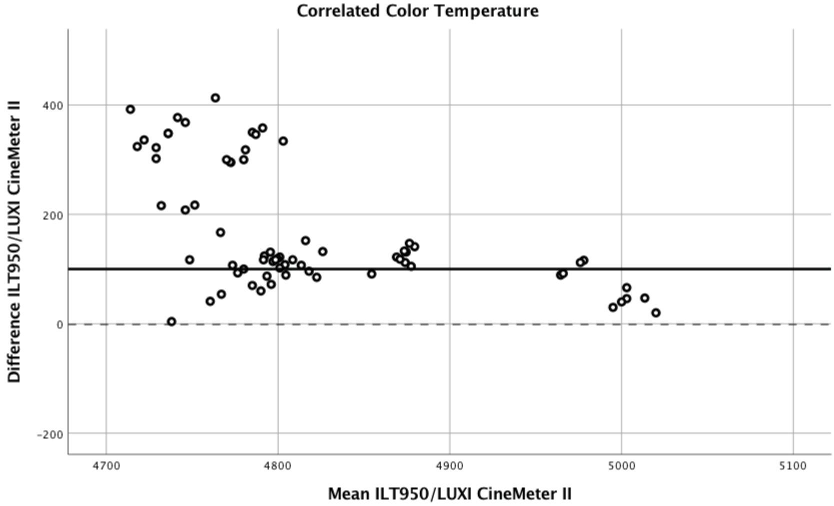

Figure 4. *Bland-Altman Plot comparing CCT readings from ILT950 with the LUXI For All with Cine Meter II app across the range of viewing booths. The bold solid horizontal line indicates the mean difference, the dashed line indicates zero.*

Comparison of the ILT950 with the Gossen Sixticolor

Finally, a Gossen Sixticolor meter was tested and compared to the reference ILT 950. Manufactured in the 1950's, the Sixticolor is another photographic light color meter. This particular meter uses a selenium photo cell and "match-needle" metering: the ambient light powers the movement of the needle, therefore the device requires no batteries. Further, this device uses only red and blue filters in the optics, and being entirely analog does not output readings with the precision of the other devices tested. Nonetheless, it is hoped that a comparison here would be of interest.

Correlated Color Temperature

When compared to the ILT950, the Sixticolor test instrument exhibited systematic bias (M = 532.86, SD = 188.37) with a mean difference greater than zero $t(69)$ 23.67, $p < .001$. In an examination of the B-A plot in Figure 5, the observed variation is evidently dependent upon the magnitude of measurement. In instances where the recorded kelvins were less than 4,800 the difference tends to increase, suggesting that no constant across the range of measured conditions would allow these devices to be used in substitute for each other.

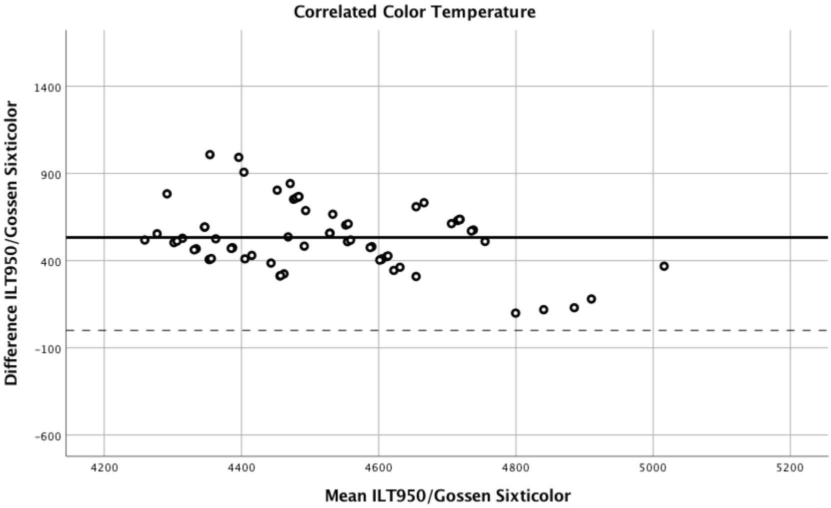

Figure 5. *Bland-Altman Plot comparing CCT readings from ILT950 with the Gossen Sixticolor across the range of viewing booths. The bold solid horizontal line indicates the mean difference, the dashed line indicates zero.*

Conclusions and Implications

As compliance to standards becomes increasingly important, graphic professionals look to ways to assure that they are adhering to the specifications published by the relevant standards bodies. While Spectroradiometers are designed for the purpose of measuring the attributes which comprise ISO 3665 2009 *Graphic technology and photography – viewing conditions* are available, these devices are not in widespread use. It is more likely that a multi-use Spectrophotometer, such as the X-Rite i1 Pro 2, be re-purposed for such applications. Alternatively, graphics professionals may look to instruments specifically designed for photographic applications to measure CCT, which is perhaps the most prominent metric of standardized color viewing.

The present study compared a Spectroradiometer as a reference instrument, to several test instruments, namely, the X-Rite i1 Pro 2 and three photographic light color meters, with a goal of determining if these devices can be used interchangeably.

When examining the i1 Pro 2, both CCT and CRI Ra were compared to the reference instrument, here, the ILT 950. In both instances, it is determined that the differences examined do not suggest that these devices can be substituted for each other.

The same was true in an examination of CCT as measured by the reference instrument when compared to three photographic light color meter solutions: the

Konica Minolta Color Meter IIIF, the LUXI For All with Cine Meter II App Mobile solution, and the Gossen Sixticolor.

In each instance of instrument comparison, the variation was dependent upon the magnitude of the readings. In looking at CCT, it is interesting to note that the readings for the test instruments were nearer to the reference instrument when the viewing condition was closer to 5,000 kelvins. Likewise, in examining CRI Ra with the i1 Pro 2 as compared to the reference ILT 950, the readings from the two devices appeared to be closer the nearer the viewing condition was to the ideal 100.

This suggests that, although at present these lower-cost test instruments cannot be empirically supported as substitutes for a traceable Spectroradiometer, the nearer the viewing condition is to the ISO standard the closer the test instruments are to the reference instrument. As discussed, in the case of the LUXI For All with the Cine Meter II app, the test solution was adjusted to match the reference meter in this condition, but this was not the case for the other solutions.

It is also important to recognize that the present analysis does not address clinical importance. For example, the differences recorded with the i1 Pro 2 in terms of CCT and CRI Ra may be within the acceptable tolerances of some facilities in terms of assuring the compliance of their viewing conditions, at least in a relative way. A user may choose to measure a viewing condition when new, and use the instrument to check those values over time. It is also relevant to note that the manufacturers of these devices make no claim on their devices being a substitute for a certified and traceable Spectroradiometer.

When absolute readings are required, however, the data obtained indicate that, for practitioners, the test instruments evaluated are no practical substitute for the reference instrument. Such users are therefore recommended to invest in a traceable Spectroradiometer or employ the services of a qualified and equipped vendor to provide such data to assure their compliance. Strict adherence to the viewing booth manufacturer's recommendation for usage hours and age of their light sources is also suggested.

References

Berns, R. S., Billmeyer, F. W., & Saltzman, M. (2000). *Billmeyer and Saltzman's principles of color technology* (3rd ed.). New York: Wiley.

Bland, J.M., Altman, D.G. (1986). Statistical Methods for Assessing Agreement Between Two Methods of clinical Measurement. *Lancet*, 307-310.

Earthman, C. P. (2015), Body Composition Tools for Assessment of Adult Malnutrition at the Bedside. *Journal of Parenteral and Enteral Nutrition*, 39: 787-822.

Field, G. G., & Graphic Arts Technical Foundation. (2004). *Color and its reproduction : fundamentals for the digital imaging and printing industry* (3rd ed.). Pittsburgh: GATF Press.

International Organization for Standardization. Standard ISO 3665. *Graphic technology and photography – viewing conditions* Geneve: ISO; 2009.