VISION EXPERIMENT ON CHROMA SATURATION FOR COLOR QUALITY PREFERENCE

Yoshi Ohno[1], Mira Fein[2], and Cameron Miller[1]

[1] National Institute of Standards and Technology, Gaithersburg, MD USA
[2] Psychology Department, Oberlin College, Oberlin, Ohio, USA.
ohno@nist.gov

Abstract

Color Rendering Index (CRI) often does not correlate well with visual evaluation of color rendering of light sources at real illuminated scenes. The main reason is that CRI measures color fidelity, while general users judge color rendering based on their preference of object color appearance, thus there is a need for a color-preference based metric. Color preference is mainly affected by saturation of object chroma. To obtain data for such color preference evaluation, a series of vision experiments have been conducted using the NIST Spectrally Tunable Lighting Facility simulating an interior room, where 20 subjects viewed various fruits, vegetables, and their skin tones, under illumination of varied saturation levels at correlated color temperatures (CCT) of 2700 K, 3500 K, and 5000 K. The results of the experiment show that subjects' preference is consistently peaked at saturation level of $\Delta C^{*}_{ab} \approx 5$ at all CCT conditions and for all target objects. The results may be useful to develop a color preference metric.

Keywords: color rendering, color preference, chroma saturation, perception, vision experiment

1. Introduction

Color Rendering Index (CRI) often does not correlate well with perceived color rendering of illuminated scenes, especially with light-emitting diode (LED) sources, as summarized in reference (CIE 2007). There have been several proposals for alternative metrics (e.g, Hashimoto 2007, Rea 2008, Davis 2010, Smet 2010) but none of them has been adopted as a standard. The Color Quality Scale (CQS) developed by NIST (Davis 2010) has not been accepted as a standard mainly due to a difficulty defining a metric that evaluates combined effects of color fidelity and color preference. International Commission on Illumination (CIE) has now two separate TCs, one developing a new color fidelity metric (TC1-90) and another writing a report on existing metrics for color quality other than fidelity (TC1-91). It is considered that an improved color fidelity metric alone will not solve the problem of correlation with perception, and there is a need for a preference-based metric to evaluate the color quality of light sources in real applications as perceived by general users.

To develop such a preference-based metric, visual evaluation data are needed. The main reason for the discrepancy between the CRI score and the perception is that the CRI measures color fidelity, only one aspect of color quality, while general users judge color rendering based on their preference of object color appearance. The discrepancy mainly occurs when the chroma of objects is enhanced by lighting, e.g., by narrowband white LED sources or hybrid sources (combination of broadband and narrow-band red). It is experienced that chroma-enhanced sources are generally preferred but if the chroma saturation is excessive, the objects appear unnatural and the preference will decrease. Such quantitative evaluation data on the preferred level of chroma saturation have not been available. To obtain such data, a series of vision experiments have been conducted using the NIST Spectrally Tunable Lighting Facility (STLF) (Miller 2009) simulating an interior room, where 20 subjects evaluated color appearance of variety of fruits, vegetables, and their skin tones, under lights with varied chroma saturation levels. The results of the experiment are presented and a way toward developing a color preference metric based on these results is discussed.

2. Experimental settings with NIST STLF

The NIST STLF, as shown in Fig. 1, was used, which has 25 channels of LED spectra (from 405 nm to 650 nm peak) and can control spectral distribution, correlated color temperature (CCT), Duv[1], and illuminance, independently, illuminating real-room size cubicles (each 2.5 m x 2.5 m x 2.4 m). There are two cubicles side by side, independently controlled, and the walls of different colors and surface textures can be replaced easily. For this experiment, only one cubicle with off-white (achromatic) walls was used (the right side in Fig. 1). The facility can produce up to about 300 lx to 800 lx of white light illumination on the table, depending on the spectrum of light.

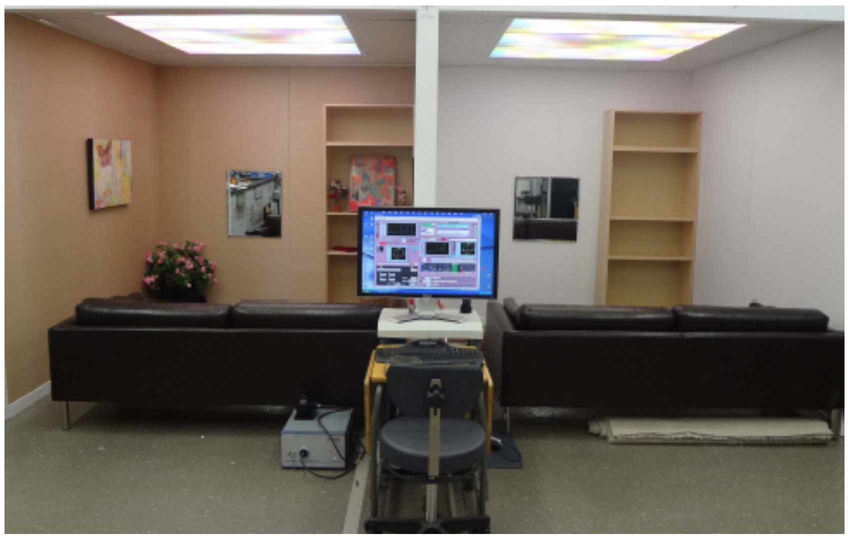

Figure 1 – View of the two cubicles of NIST Spectrally Tunable Lighting Facility.

The light source unit of the STLF has very large heat sinks which are cooled by forced air and the temperature of the heat sink is only about 27 °C when these spectra at ~300 lx are produced, while the room temperature is kept at 25 °C ±1 °C. The STLF needs only about 15 minutes to stabilize, after which the chromaticity is stable to within ± 0.0005 in (*u', v'*) for four hours, and reproduces the set chromaticity to within ± 0.001 in (*u', v'*) over one month. The STLF can change the light spectrum instantly so that there was no switching time between two lights presented sequentially and stable immediately, which allows easy comparison of a pair of lights presented sequentially.

For the experiment, the STLF was set for RGBA spectra with peaks around 460 nm, 530 nm, 590 nm, and 635 nm. Such a combination of narrowband peaks was needed to create increased chroma saturation. (Smooth broadband spectra cannot achieve it). The chroma saturation was varied by changing the red/amber ratio. Nine different levels of chroma saturation as shown in Fig. 2 (left) were prepared at each of 3 different CCTs; 2700 K, 3500 K, and 5000 K. The chroma was set to CIELAB chroma differences, ΔC^{*}_{ab}= -16 to +16 from the neutral condition (the chroma of the reference illuminant of CRI) at intervals of 4 ΔC^{*}_{ab} units, measured on the red CQS sample (Davis & Ohno 2010). The chroma for the green CQS sample had slightly smaller changes of chroma than the red sample, while there were very small changes in chroma in yellow and blue regions with changes of the amber/red ratio. So the overall gamut area also increased or decreased as the chroma saturation was changed.

In addition, another condition was set for a negative Duv[1] (see also Ohno 2013), Duv = -0.015, at 3500 K as shown in Fig. 2 (right). All other conditions were at Duv=0.000. This condition was

[1] The signed distance from the Planckian locus on CIE u', 2/3 v' coordinates (ANSI 2011)

added, considering that the Duv level might affect the preference for chroma saturation, based on the results of the previous experiment on Duv preference (Ohno & Fein, 2014), in which the Duv level of -0.015 on the average was found most preferred.

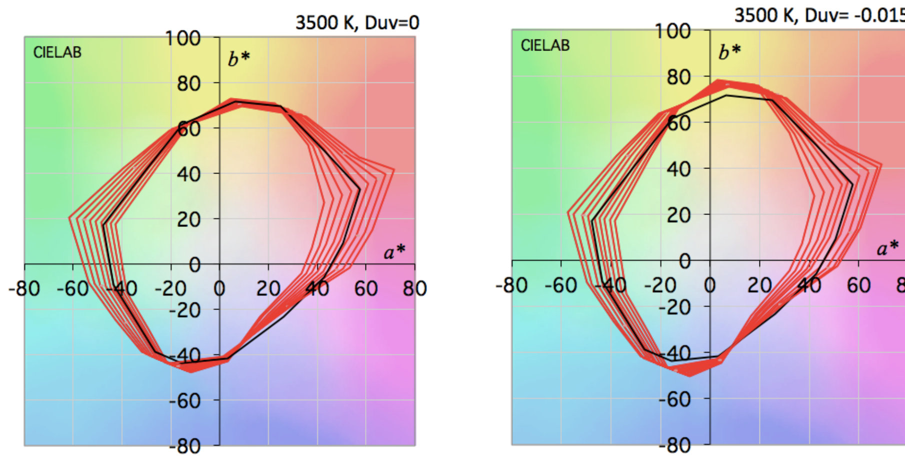

Figure 2 – The CIELAB (a*, b*) plots of the 15 CQS samples for the nine different saturation levels used for the experiment for 3500 K, Duv=0 (left) and Duv= -0.015 (right). The black line is for the reference illuminant in CRI (Planckian radiation at the same CCT).

At each of the four CCT/Duv conditions, the chromaticity on the STLF was kept constant, to within ± 0.0003 in (u', v'), while the chroma saturation was varied at 9 different levels; thus, there was no issue of chromatic adaptation for observations within each CCT/Duv condition. Chromatic adaptation period was needed only when the CCT/Duv condition was changed.

The measured spectra for these four CCT/Duv conditions are shown in Fig. 3. The spectral distributions of the lights were measured on the center of the table in the cubicle, using an array type spectroradiometer with a small integrating sphere input for cosine response, calibrated with a standard lamp traceable to NIST spectral irradiance scale (NIST, 2011).

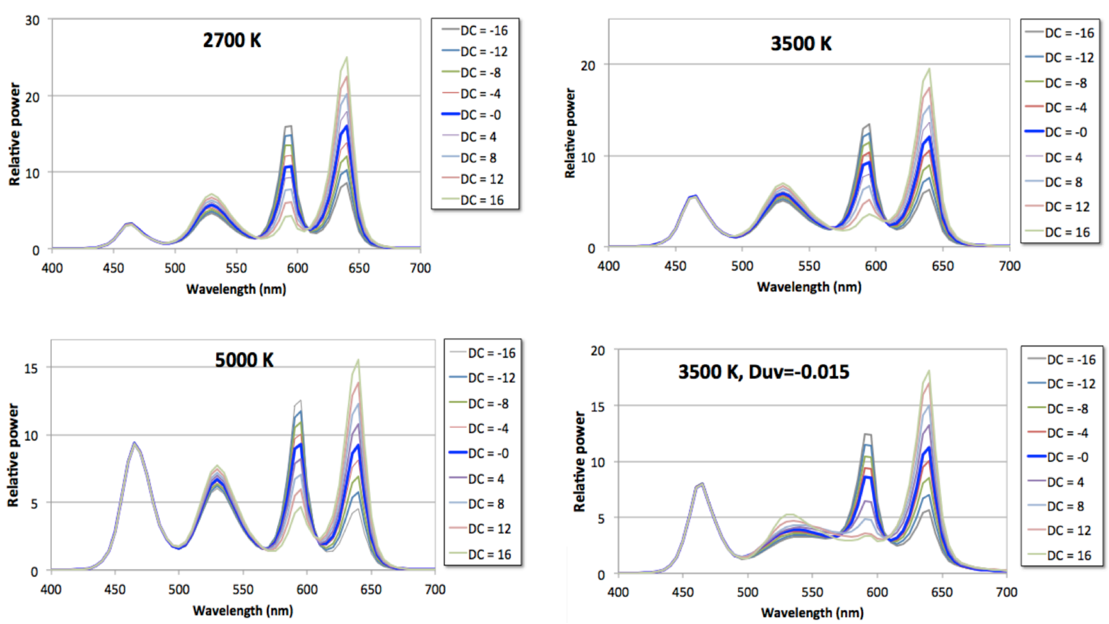

Figure 3 – The spectral distributions set at NIST STLF for the four CCT/Duv conditions for the experiment. "DC" in the figures represents ΔC^{*}_{ab}.

The spectroradiometer measured spectra and illuminance on the coffee table (see Fig. 5) from the 2π solid angle from the entire room including reflections from the walls and other objects as well as from the light source itself. The estimated expanded uncertainties (k=2) of measurements varied depending on spectra, but in all cases, they were within 0.0012 in u' and v', 0.0009 in Duv, 24 K in CCT at 2700 K and 92 K at 6500 K. The repeatability of the spectroradiometer was 0.0002 in u' and v'.

Fig. 4 shows the plots of the CRI R_a and R_9 values as well as the CQS Q_a values, calculated from the spectral distribution data in Fig. 3, as a function of chroma saturation ΔC^*_{ab}, for all the four (CCT, Duv) conditions.

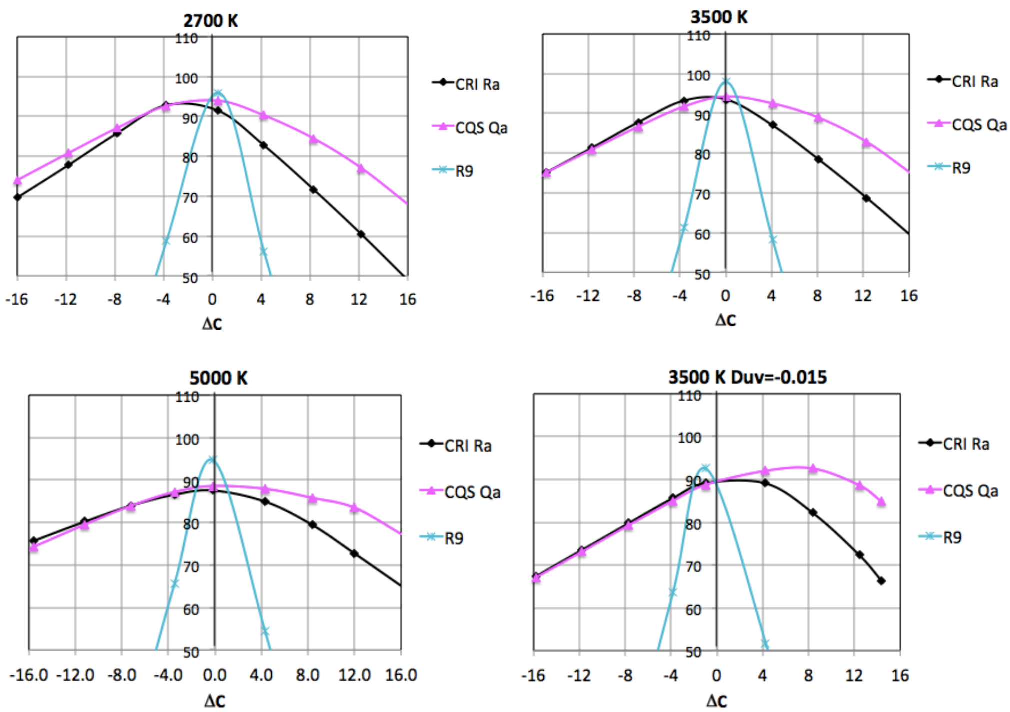

Figure 4 – The CRI R_a, R_9 values and CQS Q_a values of the spectra at the four CCT/Duv conditions used in the experiment.

The STLF cubicle for the experiment was prepared as an interior room as shown in Fig. 5, with a couch (not shown in the photo), a coffee table, a bookshelf with books, some artificial flowers, paintings on the walls. A mirror was also placed on the wall against the couch, to allow evaluation of skin tone of the face of the subject. On the table there were two plates of real fruits and vegetables; apples, oranges, bananas, strawberries, peppers, lettuce, tomatoes, red cabbage, and grapes. These fruits and vegetables were replaced every few days to keep them fresh during a few weeks of the experiment period. Pictures of these fruits and vegetables were taken, and when these were replaced, those having as similar colors and sizes as possible were purchased and used throughout the experiments.

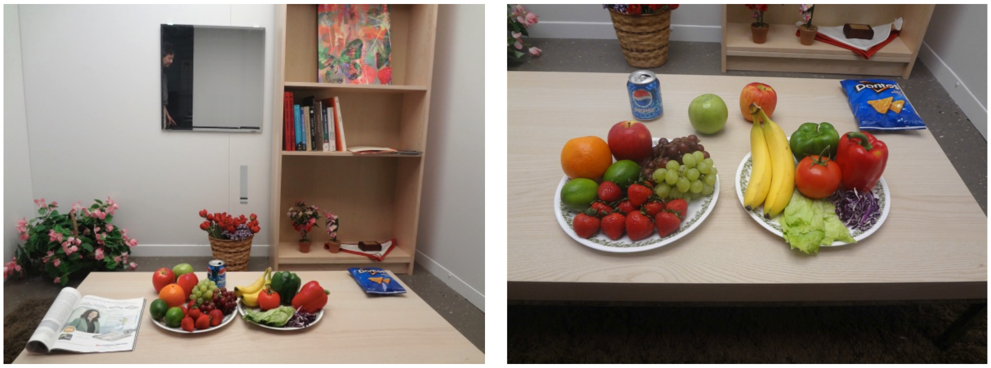

Figure 5 – The set up of the STLF cubicle with objects on the table used in the experiments.

In addition to these settings, the experiments were also done for red samples only and for green samples only. For the green samples, the chroma saturation levels were set for steps of 4 ΔC^{*}_{ab} units for the CQS green sample. Figure 6 shows the photo of the red samples only and the CIELAB (a*, b*) plots of the lights prepared, and Fig. 7 shows those for the green samples only.

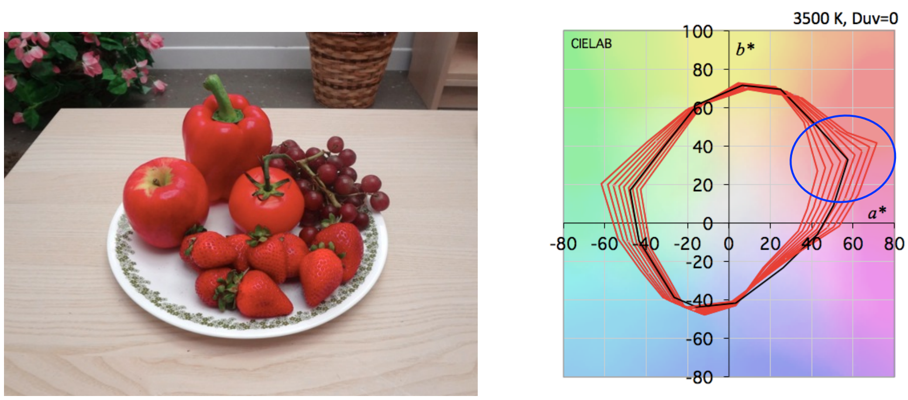

Figure 6 – Photo of the red samples only and the CIELAB plots of the chroma saturation setting of lights at STLF.

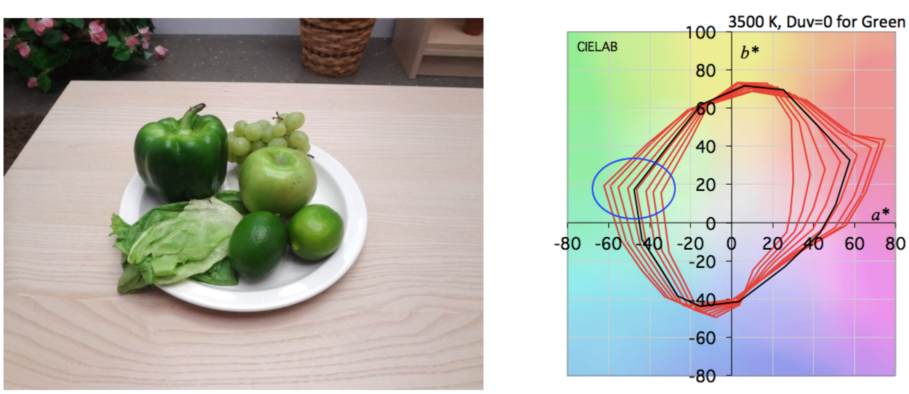

Figure 7 – The photo of the green samples only and the CIELAB plots of the chroma saturation setting of lights at STLF.

3. Experimental Procedures

Total 20 subjects having normal color vision were recruited for the experiments. They were 10 males and 10 females with their ages from 20 to 59, consisting of 7 Caucasians, 5 Hispanics, 4 Asians, and 4 dark-skin persons (African or Indian origin). The subjects were volunteers working at NIST, who are not experts on color, including eight summer students. Each subject was first tested for normal color vision using Ishihara Test, and instructions were given for the experiment including a trial comparison of light pairs.

Experiments were conducted for the four CCT/Duv conditions described in Section 2. The subject was adapted to the first CCT/Duv condition for at least ten minutes before the experiment started (this was achieved when instructions were given), and at for least three minutes when CCT or Duv was changed. The experiments were repeated for 3500 K condition only (due to limited time available) for all viewing conditions and for each subject, to evaluate reproducibility of results.

For each CCT/Duv condition, experiments were conducted on four viewing targets: 1) the mixed fruits/vegetables on the table and the entire room, 2) their skin tone (face in a mirror and hands), 3) red fruits/vegetables only, and 4) green fruits/vegetables only, as shown in Figs. 5, 6 and 7.

For the above CCT/Duv conditions and viewing targets, the experiments were grouped into three sessions, and conducted for 20 experiment-runs for each subject as listed in Table 1, and conducted in this order.

Table 1 – Conditions for the 16 experiment-runs for each subject

Session	Run	CCT [K]	Duv	Viewing target
1	1	3500	0	Mixed fruits and vegetables
	2	3500	0	Skin tone
	3	2700	0	Mixed fruits and vegetables
	4	2700	0	Skin tone
	5	5000	0	Mixed fruits and vegetables
	6	5000	0	Skin tone
	7	3500	-0.015	Mixed fruits and vegetables
	8	3500	-0.015	Skin tone
2	9	3500	0	Red fruits and vegetables only
	10	3500	0	Green fruits and vegetables only
	11	2700	0	Red fruits and vegetables only
	12	2700	0	Green fruits and vegetables only
	13	5000	0	Red fruits and vegetables only
	14	5000	0	Green fruits and vegetables only
	15	3500	-0.015	Red fruits and vegetables only
	16	3500	-0.015	Green fruits and vegetables only
		<Repeat>		
3	17	3500	0	Mixed fruits and vegetables
	18	3500	0	Skin tone
	19	3500	0	Red fruits and vegetables only
	20	3500	0	Green fruits and vegetables only

In each experiment-run, 8 pairs of lights were prepared from the 9 different saturation levels as shown in Fig. 2, with one light of each pair being always the reference light (neutral saturation). Each pair of light was presented sequentially. Fig. 8 shows an example of such a pair. In this example, light "A" is the reference (ΔC^*_{ab}=0), and light "B" is at ΔC^*_{ab}=8. When each light was presented, the operator said "A" or "B" and asked "which light looks better, A or B?," and the pair was repeated two or more times as necessary, till subject answered. Each light was presented for about three seconds before switching to next light. The order of the reference light, coming first or second, was changed for different pairs and the subject did not know about the existence of reference light and when it was presented.

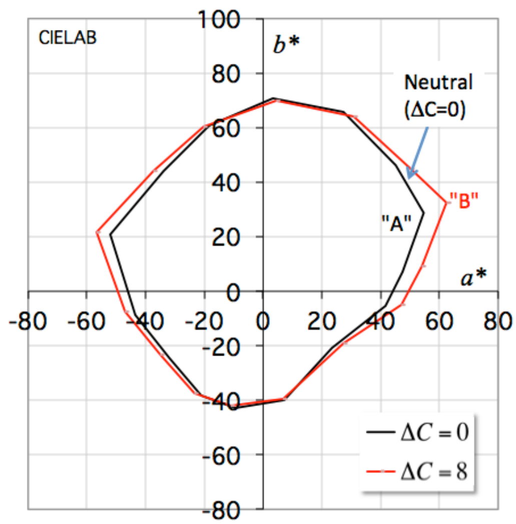

Figure 8 – An example of a pair of lights in the experiment.

The subject's answer was a forced choice, but in addition, the subject was also instructed to report if the choice was difficult, and such information was recorded together with the answers A or B. The order of pairs of light for different saturation levels were set in a random manner as shown in Table 2.

Table 2 – Light pairs for the eight comparisons in an experiment-run

Pair	ΔC^*_{ab} of 1st light "A"	ΔC^*_{ab} of 2nd light "B"
1	-8	0
2	8	0
3	0	16
4	-12	0
5	-4	0
6	4	0
7	0	12
8	0	-16

Each experiment-run took about five minutes on the average, and all sessions took about two and half hours on the average for each subject. The three sessions for each subjects were scheduled in one or two days depending on their schedules. Experiments were done for one or two subjects a day. All experiments for 20 subjects were carried out over a period of three weeks.

4. Results

The subjects' response data were first analyzed as shown in Table 3. In that table, 0 as the subject response means that the subject chose the reference light as better, and 1 means that the subject chose the light with the ΔC^*_{ab} value of the column as better. For example, subject 1 responded that the reference was better than all de-saturated lights, and the saturated lights ΔC^*_{ab}=4, 8, 12, were better than the reference, but the saturated lights ΔC^*_{ab}=16 was not as good as the reference. The last two rows show the average values of all subjects at each ΔC^*_{ab} level for this CCT/Duv condition and the percentages that subjects chose the light with each ΔC^*_{ab} level better than the reference (neutral). The percentage value 50 % is given for $\Delta C^*_{ab}=0$ (comparing the same light, which was not included in the experiment), as it is supposed to be even.

Table 3. Example of raw data for Run 1 (3500 K, mixed fruits and vegetables)

	Chroma saturation ΔC^*_{ab} and subject response								
Subject	-16	-12	-6	-4	0	4	8	12	16
1	0	0	0	0		1	1	1	0
2	0	0	0	0		1	1	1	0
3	0	0	0	0		1	1	1	1
4	0	0	0	1		1	1	1	1
…	…	…	…	…		…	…	…	…
20	0	0	0	0		1	1	1	1
Average	0.05	0.05	0	0.15		0.95	1	0.8	0.6
Percentage	5 %	5 %	0 %	15 %	50 %	95 %	100 %	80 %	60 %

The analysis shown in Table 3 was done for all CCT/Duv conditions and for all viewing targets. Figure 9 shows all the results for mixed fruits/vegetables and skin tone. The solid lines show the percentage values as in the bottom of Table 3, which show subjects' preference of each ΔC^*_{ab} light over the neutral, and the thin dashed lines are the percentages that subjects answered as "difficult to choose". Figure 10 shows all the results for red fruits/vegetables only and green fruits/vegetables only.

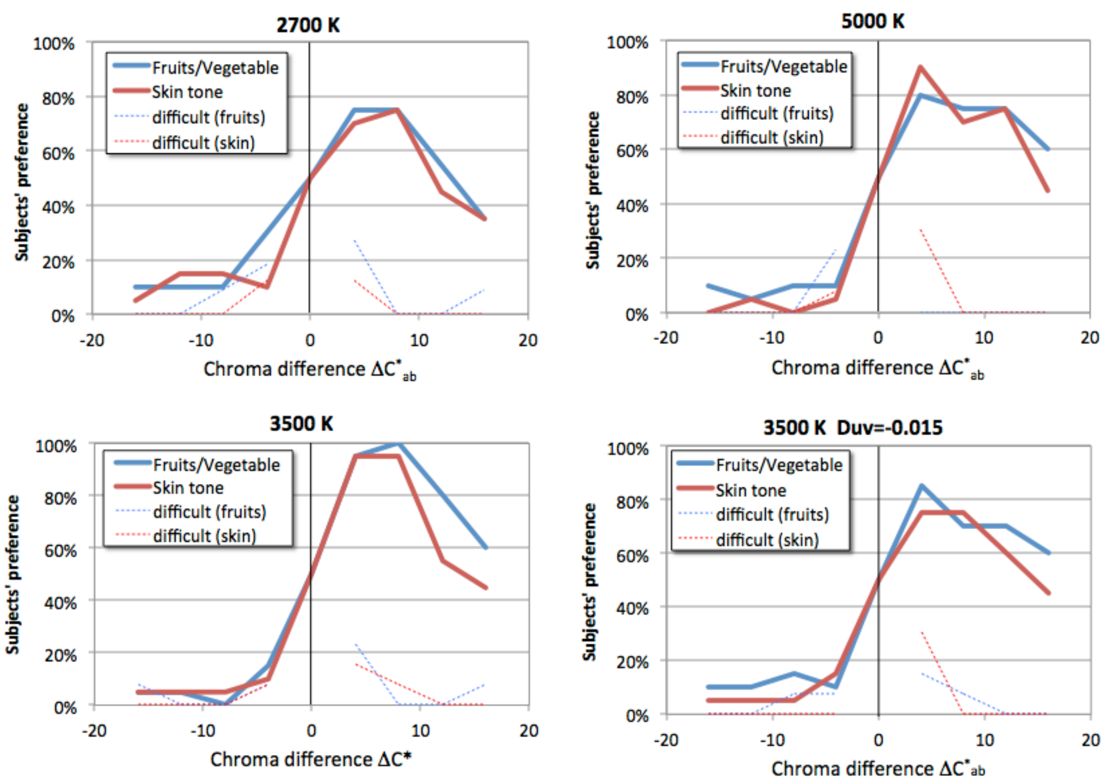

Figure 9 –Average results of all subjects for mixed fruits/vegetables and skin tone, shown in percentage of subjects' preference over the neutral.

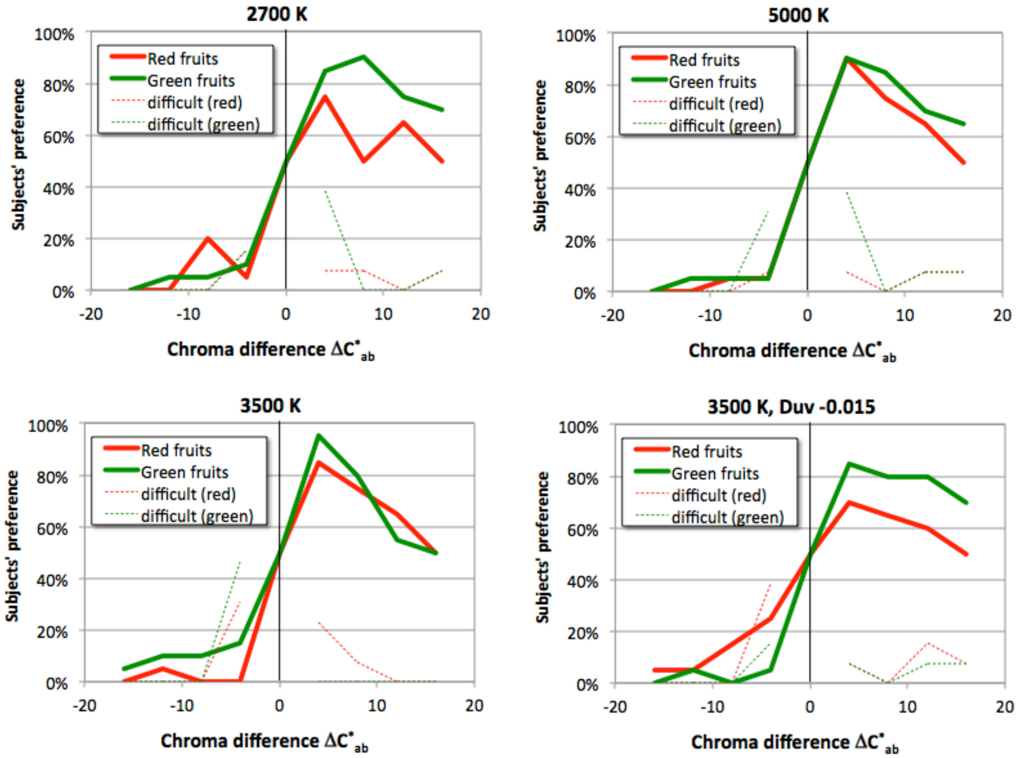

Figure 10 – The average results of all subjects for red and green fruits/vegetables only, shown in percentage of subjects' preference over the neutral.

Examining Figs. 9 and 10, all curves are surprisingly similar for all different CCT/Duv conditions and viewing targets. Subjects' preference consistently peaked (80 to 90 %) at a saturation level of $\Delta C^*_{ab} \approx 5$ and decreases slowly as light becomes more saturated ($\approx$ 50 % at ΔC^*_{ab}=16), while the percentages at all levels of de-saturation are very low (less than 20 %).

Figure 11 shows the averages of all CCT/Duv conditions for the four viewing targets, and Fig. 12 shows the grand average of all results. The error bars are the standard deviations of four points at each ΔC^*_{ab} in Fig. 11.

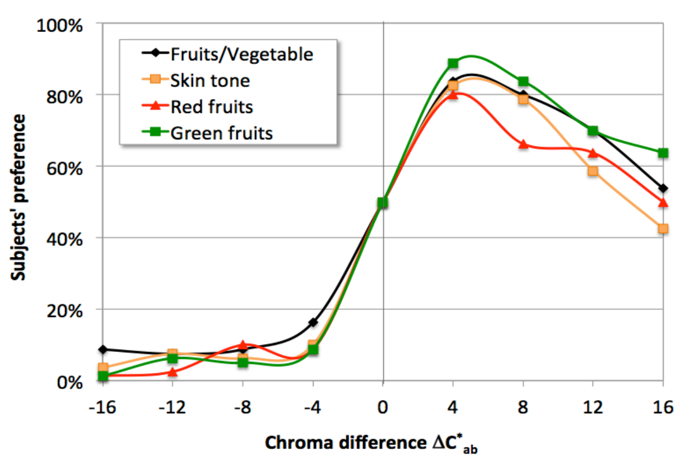

Figure 11 – the average results of all CCT/Duv conditions for the different viewing targets

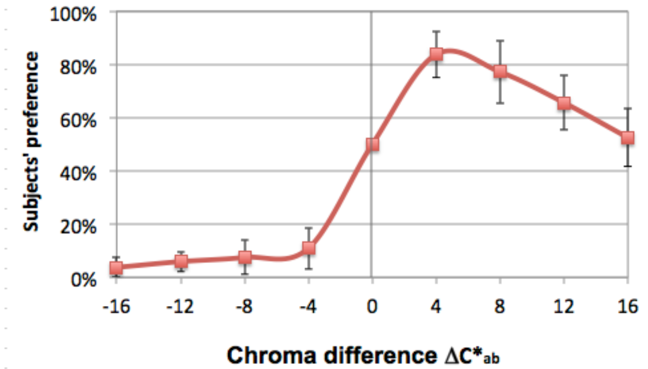

Figure 12 – Average results for all CCT/Duv conditions and all viewing targets

Discussions and conclusions

The results of the experiments indicate that lights having chroma saturation around $\Delta C^*_{ab} \approx 5$ from the neutral are most preferred and the preference slowly decreases when chroma saturates further, when viewing fruits/vegetables and skin tone. Based on the results, a color preference metric may be developed by using a reference illuminant that has chroma saturation of $\Delta C^*_{ab} \approx 5$ so that the metric would yield the highest score for such lights and score decreases when ΔC^*_{ab} deviates in either direction. In this experiment, however, the chroma was changed mostly in the red and green region. It is not clear whether the preference effects in yellow and blue regions (or other colors) are the same. It was difficult to produce such experimental lights for yellow and blue this time. It is hoped that such preference experiments can be conducted for yellow, blue, or even other colors in the future. Also, further studies are desired to verify applicability of these results in various real application conditions.

Acknowledgement

This research was conducted when Mira Fein stayed as a guest researcher at NIST under NIST-NSF Summer Undergraduate Research Fellowship Program.

References

ANSI 2011, ANSI_NEMA_ANSLG, C78.377-2011 Specifications for the Chromaticity of Solid State Lighting Products.

CIE 2007. CIE 177: 2007 Colour rendering of white LED light sources.

Davis, W. and Ohno, Y. 2010. “Color Quality Scale,” *Optical Engineering*, 033602-1 March 2010/Vol. 49_3.

Hashimoto, K., Yano, T., Shimizu, M. and Nayatani, Y. 2007. New method for specifying color-rendering properties of light sources based on feeling of contrast,” Color Res. Appl. 32, 361–371.

Ohno, Y. and Fein, M. 2014. Vision Experiment on Acceptable and Preferred White Light Chromaticity for Lighting, CIE x039:2014, pp. 192-199.

Ohno, Y. 2013, Practical Use and Calculation of CCT and Duv” *LEUKOS*, 10:1, 47-55, DOI: 10.1080/15502724.2014.839020.

Smet, K., Ryckaert, W., Pointer, M., Deconinck, G., Hanselaer, P. 2010. Memory colors and color quality evaluation of conventional and solid-state lamps. Opt. Expr. 18.

Rea, M., Freyssinier, J.P. 2008. Color rendering: a tale of two metrics. Color Res. Appl. 33.

Yoon, H. W. and Gibson, C. E. 2011. NIST Special Publication 250-89 Spectral Irradiance Calibration.

Miller, C., Ohno, Y., Davis, W., Zong, Y., and Dowling, K. “NIST spectrally tunable lighting facility for color rendering and lighting experiments,” in *Proc. CIE 2009: Light and Lighting Conference*. 5 pages (2009).