Experimental validation of colour rendition specification criteria based on ANSI/IES TM-30-18

Michael P Royer[1]
Minchen Wei[2]
Andrea Wilkerson[1]
Sarah Safranek[1]

[1]Pacific Northwest National Laboratory
620 SW 5th Avenue, Suite 810
Portland, OR 97204
michael.royer@pnnl.gov

[2]Hong Kong Polytechnic University

This is an archival copy of an article published in *Lighting Research & Technology*. Please cite as:

Royer M, Wilkerson A, Wei M, Safranak S. Experimental validation of color rendition specification criteria based on ANSI/IES TM-30-18. *Lighting Research & Technology* 2019; Online before print. DOI: 10.1177/1477153519857625

Abstract

An experiment was conducted to examine colour rendition specification criteria. Twenty-five participants each evaluated 90 lighting scenes in a room filled with objects. The lighting scenes included nine chromaticity groups, each with 10 systematically-varied colour rendition conditions designed to meet or not meet previously proposed colour preference specification criteria using ANSI/IES TM-30-18 R_f, $R_{cs,h1}$, and R_g. The colour rendition conditions did not meet the criterion for none, one, two, or all three of these measures. Participants, who chromatically adapted to each chromaticity group, rated the objects' colour appearance on eight-point scales for saturated-dull, normal-shifted, and like-dislike (preference), as well as a binary for acceptable or unacceptable. The findings corroborate past work, but also indicate that colour preference criteria could be adjusted slightly to improve performance, with Tier A having $R_f \geq 78$, $R_g \geq 95$, and -1% ≤ $R_{cs,h1}$ ≤ 15%, Tier B having $R_f \geq 74$, $R_g \geq 92$, and -7% ≤ $R_{cs,h1}$ ≤ 19%, and Tier C having $R_f \geq 70$, $R_g \geq 89$, -12% ≤ $R_{cs,h1}$ ≤ 23%. A companion regression analysis shows models based on R_f, R_g, and $R_{cs,h1}$ were superior in predicting colour preference compared to those using other measures of colour rendition.

1 Introduction

Standardized methods for evaluating light source colour rendition have progressed rapidly in the past five years, beginning with the publication of Illuminating Engineering Society (IES) TM-30-15,[1] followed by the adoption of Commission Internationale de l'Eclairage (CIE) 224:2017,[2] and culminating with the issuance of American National Standard Institute (ANSI) and IES TM-30-18[3]—which revises TM-30-15. CIE 224:2017 and ANSI/IES TM-30-18 share a calculation framework (colour samples, colour vision model, reference illuminant scheme); ANSI/IES TM-30-18 provides a comprehensive system of output measures, whereas CIE 224:2017 specifies only colour fidelity calculations. Importantly, neither document provides guidance on how to specify colour rendition using the included measures. That is, no value or combination of values is identified that leads to any desired quality of the light. To fill this void, numerous research efforts have investigated the accuracy of the new measures and their relationships to subjective qualities of the lit environment.[4-13]

While considerable effort has been made to understand correlations between subjective evaluations of colour quality (i.e., colour preference, colour naturalness, colour vividness, colour acceptability), little attention has been given to establishing new colour rendition specification criteria, which are the primary mechanism by which colour quality is evaluated in practical applications by entities such as the IES, United States Environmental Protection Agency (EPA), the DesignLights Consortium (DLC), the state of California, International Organization for Standardization (ISO), or the Chartered Institution of Building Services Engineers (CIBSE). Establishing colour rendition specification criteria requires detailed consideration of their purpose and level of stringency, among many other factors that influence the type of measures included and their threshold values. These factors are discussed in a separate article.[14]

The experiment described herein was explicitly developed as a test of previously-proposed colour preference specification criteria based on TM-30, which resulted from two prior psychophysical experiments.[4, 5] These criteria were:

- Tier A: $R_f \geq 78$, $R_g \geq 100$, $-1\% \leq R_{cs,h1} \leq 15\%$
- Tier B: $R_f \geq 78$, $R_g \geq 98$, $-7\% \leq R_{cs,h1} \leq 15\%$

Note that these values were converted from IES TM-30-15 to IES TM-30-18, which changed the R_f criterion from 75 to 78. A third tier was developed based on a benchmarking exercise, given that the initial two tiers excluded a majority of products currently being sold. These criteria were:

- Tier C: $R_f \geq 70$, $R_g \geq 88$, $-12\% \leq R_{cs,h1} \leq 15\%$

Many experiments have examined colour preference,[4, 5, 7-11, 15-50] although the focus has traditionally been on developing a colour preference metric rather than specification criteria. This includes many experiments that have examined colour preference at different chromaticities.[20, 22-24, 35, 43-45, 47-50] While consistency in preferred attributes has emerged, such as the important roles of chroma enhancement and red hues,[4, 5, 10, 11, 15, 16, 18, 20, 22-24, 27, 38, 39, 41, 46-54] there has been variation in findings, with many models or metrics developed to fit specific datasets. Chromaticity has been a significant factor in some experiments but not others; experimental methods, particularly chromatic adaptation and visible transitions, may contribute to the differences. Colour preference may not be the most critical consideration in all lighting applications, but it is important because environmental satisfaction has been linked to wellbeing and performance.[55, 56]

For this experiment, ten colour rendition conditions were created that failed none, one, two, or all three criteria at various tiers. SPDs with these 10 colour rendition characteristics were created at nine different chromaticities. These chromaticities were chosen to further investigate a prior result that indicated a possible interactive effect of CCT and D_{uv} on colour preference at low CCTs.[4] Compared to this previous study, the objects being illuminated were changed in order to explore another aspect of the stimulus and provide a more robust test of the previously-developed specification criteria. Details of the methods used are provided in the next section, with additional information and justification in past articles.[4, 5]

In addition to evaluating the proposed criteria, existing criteria based on CIE 13.3-1995 were also analysed. The performance of previously developed colour preference metrics, including the Colour Quality Scale (CQS) Q_a,[41] Memory Colour Rendering Index R_m,[33, 57] the Gamut Volume Index (GVI),[58] Class A Colour,[59] the Feelings of Contrast Index (FCI),[60, 61] and a recently proposed, unnamed extension of CIE 13.3-1995,[62] were also examined.

2 Methods

2.1 Apparatus and test room

2.1.1 Experimental rooms and lighting equipment

Aside from a room with large windows that was used for welcoming and completing informed consent, the participants spent their time in two distinct rooms, the *experiment room* and the *adaptation* room, which were separated by a dark corridor. The experiment room, shown in Figure 1, was 3.7 m by 5.5 m, with a 3 m ceiling height, and the achromatic adaptation room, holding only a table and chairs, was 2 m by 3.5 m by 2.5 m. Seven ETC Source Four Series 2 Lustr luminaires in the experiment room and three ETC D22 Lustr+ luminaires in the adaptation room provided light. Both the rooms and lighting systems were the same as described by Royer *et al*.[4]

Figure 1. Photographs of the inside of the experiment room.

2.1.2 Objects for evaluation

As with the preceding experiments, objects in the experiment room were selected to provide a reasonable distribution within all three dimensions of the colour volume (hue, chroma, and lightness). Spectral reflectance functions were measured for the objects (up to six measurements for polychromatic objects) and room surfaces using a factory calibrated Minolta CM-600d spectrophotometer (SN: 21011777). Most of the objects were different from those used in the prior two experiments. Previous results indicated red objects—the Coke box in particular—as being influential to the observers' responses, but the Coke box was also one of the largest items among the packaged consumer goods. New, larger packaged items were chosen to determine if this presence, or the colour alone, influence the prior results. The unfamiliar textiles were also replaced with clothing featuring local sports teams, for which the colour had a greater chance of being familiar to the study participants. New arrangements of natural foods and a large plant were also added. The spectral reflectance functions are provided in a supplemental file.

2.2 Lighting scenes

Ninety experimental SPDs were created, with ten systematically-varied colour rendition conditions (Table 1) in each of the nine chromaticity groups (A through I), delineated based on CCT and D_{uv} (Table 2). The range of D_{uv} values was extended compared to the prior study, based on results of a recently published study showing a greater preference for more negative D_{uv} values.[8] Measurements of record for each SPD (including illuminance) were taken in-situ using a calibrated Minolta CL-500A illuminance spectrophotometer (SN: 100020008) immediately prior to the start of subjective data collection. Prior calibration of this system indicated high levels of repeatability over time. Nine adaptation SPDs were created in the same manner. Data for all SPDs is provided in a supplemental file. More details on the stimulus characterization are provided in Appendix A.

Colour rendition conditions were specified using values for the Fidelity Index (R_f), Gamut Index (R_g), and red Local Chroma Shift ($R_{cs,h1}$), with additional consideration given to matching gamut shape.[63] Figure 2 shows

Table 1. Characteristics of the 10 color rendition conditions. Shaded cells indicate values that do not meet the Tier A specification. The numbers and order of the color rendition conditions does not have any meaning.

		Color Rendition Condition									
		1	2	3	4	5	6	7	8	9	10
R_f	**Target**	**72**	**74**	**82**	**82**	**82**	**82**	**82**	**91**	**70**	**84**
	Min	71	72	80	81	80	81	80	88	69	83
	Mean	72	75	81	84	82	82	81	90	71	84
	Max	74	77	83	85	85	85	82	92	74	86
R_g	**Target**	**84**	**100**	**95**	**101**	**101**	**101**	**101**	**101**	**120**	**112**
	Min	81	96*	93	98	99	99	101	100	119	110
	Mean	83	98	95	100	100	101	102	102	121	112
	Max	86	100	96	104	102	103	103	103	124	114
$R_{cs,h1}$	**Target**	**-19%**	**3%**	**3%**	**-4%**	**-11%**	**3%**	**9%**	**3%**	**20%**	**3%**
	Min	-22%	3%	1%	-6%	-12%	3%	8%	1%	19%	2%
	Mean	-20%	4%	2%	-4%	-11%	4%	9%	3%	20%	3%
	Max	-18%	5%	3%	-2%	-10%	4%	10%	3%	23%	5%

*Conditions A2, D2, and G2 fell below the criterion of $R_g \geq 98$, therefore failing two criteria instead of one. The other iterations of color rendition condition 2 only failed the R_f criterion for Tier A. It was not possible to meet the threshold with the luminaires used.

Table 2. Characteristics of the nine chromaticity groups.

		Chromaticity Groups								
		A	**B**	**C**	**D**	**E**	**F**	**G**	**H**	**I**
CCT (K)	**Target**	**2700**	**2700**	**2700**	**3100**	**3100**	**3100**	**3500**	**3500**	**3500**
	Min	2683	2681	2686	3084	3091	3076	3474	3481	3476
	Mean	2700	2699	2701	3102	3109	3096	3497	3501	3497
	Max	2712	2715	2714	3116	3117	3115	3524	3523	3517
D_{uv}	**Target**	**0.000**	**-0.007**	**-0.014**	**0.000**	**-0.007**	**-0.014**	**0.000**	**-0.007**	**-0.014**
	Min	-0.0002	-0.0074	-0.0140	-0.0005	-0.0073	-0.0142	-0.0005	-0.0074	-0.0143
	Mean	0.0000	-0.0069	-0.0138	0.0000	-0.0069	-0.0138	0.0000	-0.0071	-0.0139
	Max	0.0006	-0.0066	-0.0136	0.0006	-0.0066	-0.0132	0.0006	-0.0065	-0.0135
E (lux)	**Target**	**308**	**308**	**308**	**308**	**308**	**308**	**308**	**308**	**308**
	Min	305	305	306	305	306	307	306	305	306
	Mean	308	308	308	308	307	308	308	307	307
	Max	311	310	310	311	310	309	310	311	310

ANSI/IES TM-30-18 Colour Vector Graphics (CVGs) for the ten colour rendition conditions of chromaticity group E. The SPDs for the nine variants of each colour rendition condition were as similar as possible, as described in Appendix A.

The vertical illuminance at 1 m above the floor at the centre of the back wall—the main calibration point and the centre of the natural foods table—was 308 lx ± 3 lx across all 90 lighting scenes. The illuminance distribution throughout the room was consistent between lighting scenes, as illustrated in Appendix A, but was not perfectly uniform across all objects. Vertical illuminance at the objects ranged from approximately 210 to 315 lx. Horizontal illuminance on the central table was approximately 330 lx.

Neither the ANSI/IES TM-30-18 values nor the SPDs themselves form the independent variable because the participants were viewing the interaction of the SPDs and the objects.[53] Custom measures were calculated based on the spectral reflectance measurements of the experimental objects (with no other changes to the calculation framework) and compared to standard ANSI/IES TM-30-18 calculations. Details, including demonstrated correlations (Figure A6), are provided in Appendix A. The custom and standard ANSI/IES TM-30-18 measures are a close match, except for measures in hue-angle bin 1 (e.g., $R_{cs,h1}$); this is due to an atypical representation of objects of this hue and does not extend to the adjacent hue-angle bins.

2.3 Participants

Twenty-five people participated in the experiment: 10 males, 14 females, and 1 unidentified. None of their professions was related to lighting. Ages of the participants ranged from 20 to 70 years, with a mean of 34 years. Before participating, each person completed a colour vision test (Ishihara's Test for Colour Deficiency, 24 plates). All results were normal.

2.4 Participant ratings (dependent measures)

For each lighting scene, participants completed a paper response form that had three semantic differential rating questions, each with an eight-point scale, and one choice question (Appendix B). The first two

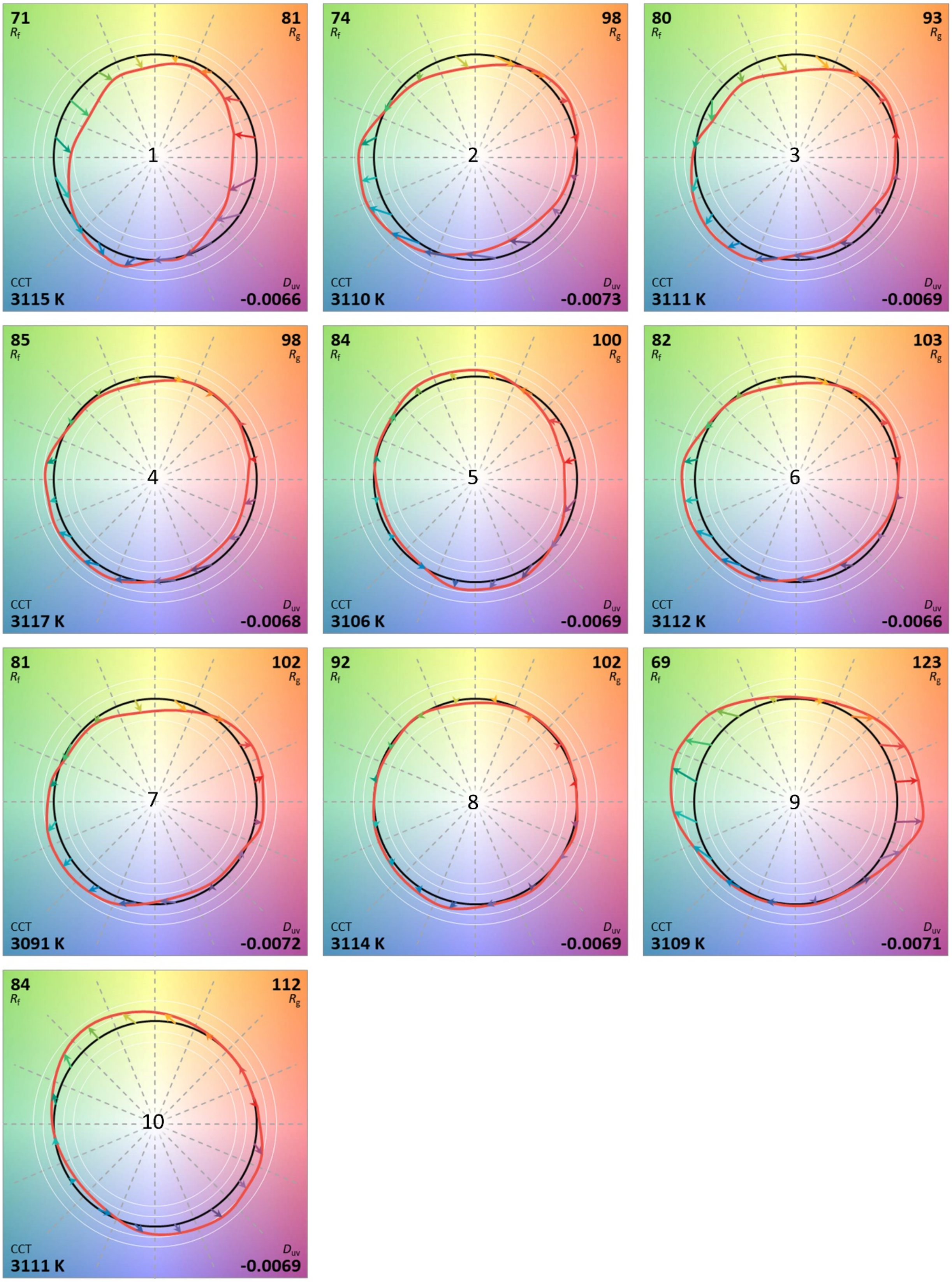

Figure 2. ANSI/IES TM-30-18 Color Vector Graphics (CVGs) representing the 10 color rendition conditions. These specific CVGs are for chromaticity group E.

semantic differential questions requested participants to circle a response, from 1 to 8, indicating whether they felt the lighting made the colour of objects appear normal (1) or shifted (8), and whether they felt the lighting made the colour of objects appear saturated (1) or dull (8). The third semantic differential question asked whether their overall opinion was that they liked (1) or disliked (8) the way the lighting made the objects appear, constituting a rating of preference. The fourth question required participants to choose whether they found the scene to be acceptable or unacceptable, with the results helpful in providing context to otherwise arbitrarily scaled responses.

To conclude the experiment, the participants completed a questionnaire to describe their experience, which provided insight into which objects or colours were the most influential in determining their judgments.

2.5 Procedure

Each subject's participation was split into two sessions, on different days, to minimize fatigue. One participant completed the experiment alone, six as part of a pair, and the remaining 18 as part of a triplet. Participants were instructed not to communicate with each other.

Upon arrival for the first session, participants completed informed consent and a demographic questionnaire, then had their colour vision checked. They were then led into a dark high-bay space housing the experiment room, which was pre-set to colour rendition condition 8 of whichever chromaticity group was randomly selected to be viewed first. Once in the experiment room, instructions were provided for several minutes, then the participants viewed conditions that demonstrated the range that would be experienced (1, 10, and 8) to provide a form of response anchoring. After concluding the instructions, the participants completed two practice trials, conditions 5 and 6, to familiarize themselves with the procedures.

Next, the participants were escorted through a vestibule and into the adaptation room. Doors on either side of the vestibule prevented the participants from ever being able to simultaneously view the lighting in the experimental and adaptation rooms. The lighting in the adaptation room was always set to the highest R_f possible with chromaticity matching the group of conditions to be subsequently viewed in the experimental room (all colour rendition conditions in each chromaticity group were viewed consecutively, in random order). After brief instructions, the participants performed a numerical verification task for three minutes while sitting at the table. The horizontal illuminance on the table was approximately 230 lx. The three-minute period was used to ensure chromatic adaptation to the same chromaticity viewed in the next rotation through the experimental room,[64-66] although long-term colour contrast adaptation artefacts cannot be completely accounted for in a short-term laboratory experiment

After completing the (non-scored) adaptation task, the participants were escorted back to the experimental room, where the lighting was set to the first scene of a randomly-selected chromaticity group. The first trial in a block, randomly chosen from the 10 colour rendition conditions, was an unidentified practice trial intended to provide further chromatic adaptation and be a randomized precursor to the first recorded response. The light from the researcher's computer monitor was altered using the F.lux program. No other light sources were visible to the participants at any time.

Participants viewed each lighting scene for at least 30 seconds before completing their response form, and then stepping out of the room and providing their form to the experimenter. Participants never saw transitions between lighting scenes. After viewing all 10 conditions in a chromaticity group, the participants were escorted back to the adaptation room, which was set to match the chromaticity of the upcoming chromaticity group viewed in the experiment room.

Four or five rotations between the experimental and adaptation rooms were made on the first day of participation, with the remainder made on the second day. On the second day, the task in the adaptation room was changed to a number search. There were no other differences. There was 48 hours of separation between the first and second day for all participants.

After the final experimental trial, the researcher set the lighting scene back to colour rendition condition 8 of the final chromaticity group, and the participants entered the room to complete the concluding summary questionnaire. Each experimental session required a total of about 90 minutes, for a total of 180 minutes for each participant.

3 Results

Means and standard deviations for each of the four responses for each of the 90 lighting scenes are provided in the supplemental data file. Corroborating previous experimental data, the results indicate correlation among the outcome measures. Mean ratings for preference and normalness (r^2 = 0.68), preference and acceptability (r^2 = 0.89) and normalness and acceptability (r^2 = 0.71) exhibited linear correlations. Mean ratings for preference (r^2 = 0.69), normalness (r^2 = 0.70), and acceptability (r^2 = 0.69) exhibited quadratic relationships with mean ratings for saturation; preference and acceptability peaked at higher levels of rated saturation than normalness. For clarity, some of the analysis focuses only on colour preference ratings. The results for normalness and acceptability follow similar trends.

Figure 3 shows mean preference ratings with 95% confidence intervals for all 90 scenes, as well as the mean preference rating for each of the 10 colour rendition conditions. In only one case were mean colour preference ratings significantly different within a condition (i.e., across chromaticity groups), but there were

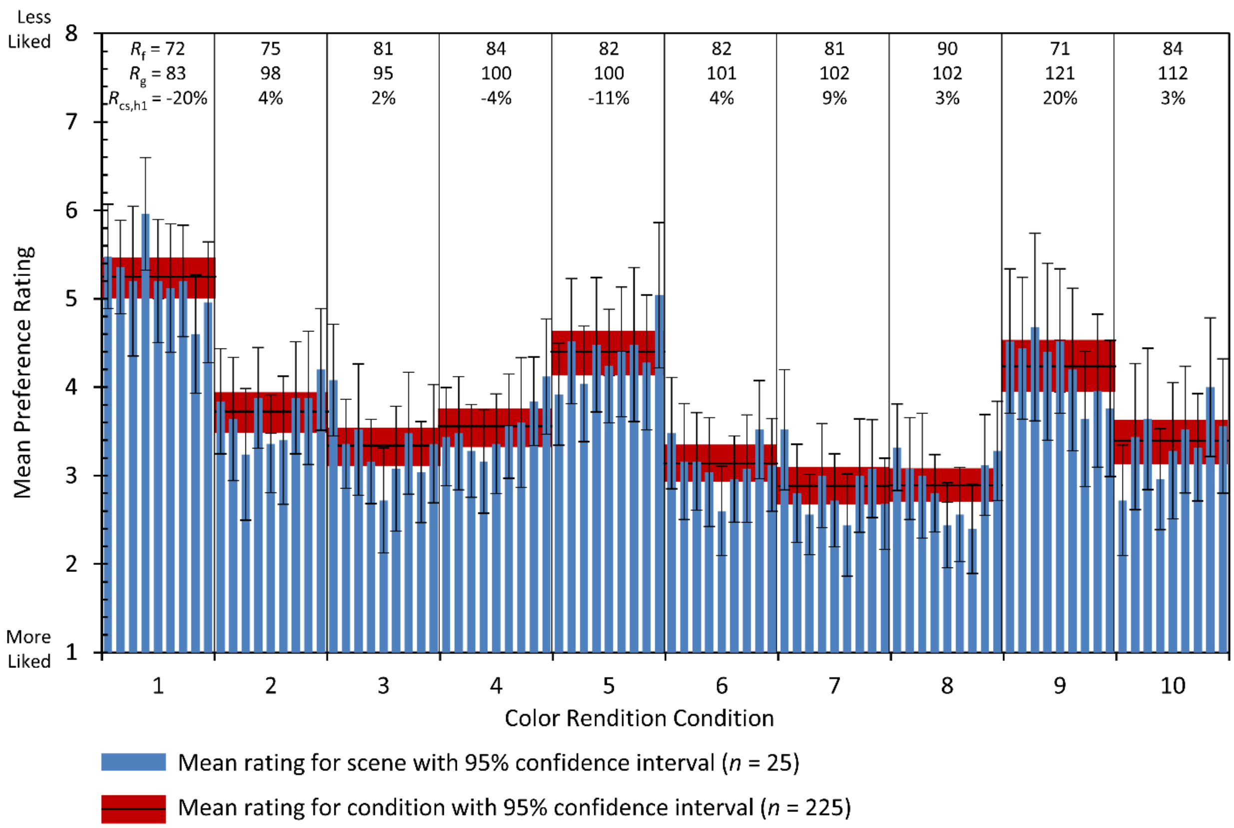

Figure 3. Mean preference rating for each of the 90 scenes grouped by color rendition condition.

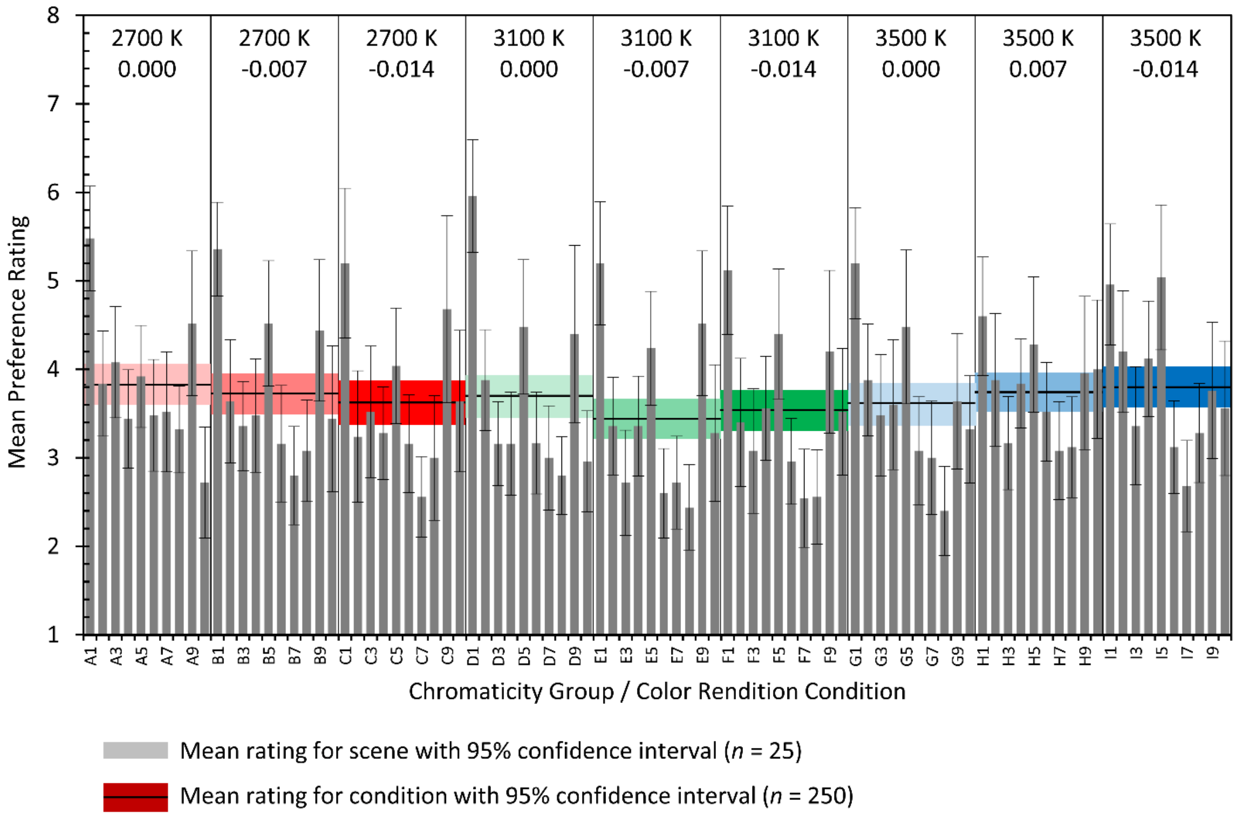

Figure 4. Mean preference rating for each of the 90 scenes grouped by chromaticity group.

substantial differences between the 10 conditions. This contrasts with the results for chromaticity (Figure 4), where the mean colour preference ratings within each group varied substantially, but there was little difference between the mean preference ratings of the nine chromaticity groups. Subsequent statistical analyses address these observations.

The results are summarized in Table 3, averaged over participant and colour rendition condition and arranged in rank order for colour preference. Table 3 also indicates groups of products, based on colour preference, using the Tukey method for multiple paired comparisons with $\alpha = 0.05$. Colour rendition conditions in different groups exhibit a statistically significant difference in mean colour preference rating. Each condition is also identified based on the specification tier that it met. The adjusted criteria tiers are discussed in the next section.

3.1 Analysis of variance and regression

It is possible to treat the colour rendition condition and the chromaticity group as categorical variables to complete analysis of variance (ANOVA) testing. For each of the four perceptions evaluated (normalness, saturation, preference, acceptability), ANOVA was conducted using chromaticity group (nine levels), colour rendition condition (ten levels), and participant (25 levels, random). For each of the four perceptions, participant and colour rendition conditions reached statistical significance ($\alpha < 0.05$) with $p < 0.001$. Chromaticity group reached statistical significance only for the mean saturation rating ($p < 0.001$). Additional tests were carried out by replacing chromaticity group with CCT (3 levels) and D_{uv} (3 levels). Both were statistically significant for mean saturation ($p = 0.021$ and $p = 0.001$, respectively), and CCT was significant ($p = 0.034$) for mean preference rating. The 3500 K conditions were rated as less saturated than the other CCT groups, and the 3100 K conditions were rated as more preferred than the other CCT groups. D_{uv} of 0.000 was

Table 3. Mean ratings for each color rendition condition, ordered based on mean preference rating. Groups the share a bar are not significantly different according to the Tukey Method for Multiple Comparisons. Shaded cells indicate failure against the Tier A specification

Color Rendition Condition	**7**	**8**	**6**	**3**	**10**	**4**	**2**	**9**	**5**	**1**
***A priori* Criteria Tier**	**A**	**A**	**A**	**C**	**A**	**B**	**C**	**-**	**C**	**-**
Adjusted Criteria Tier	**A**	**A**	**A**	**A**	**A**	**B**	**B**	**C**	**C**	**-**
Mean R_f	81	90	82	81	84	84	75	71	82	72
Mean R_g	102	102	101	95	112	100	98	121	100	83
Mean $R_{cs,h1}$	9%	3%	4%	2%	3%	-4%	4%	20%	-11%	-20%
Mean Preference Rating	2.88	2.89	3.14	3.33	3.38	3.54	3.70	4.24	4.38	5.23
Percent Acceptable	95%	97%	93%	92%	88%	88%	81%	65%	69%	50%
Mean Normalness Rating	3.19	2.85	3.06	3.00	3.72	3.12	3.42	4.94	4.02	4.74
Mean Saturation Rating	2.92	3.46	3.76	3.98	2.67	4.13	4.29	1.64	4.38	5.92

Tukey Comparisons (Color Preference)

rated as less saturated than the alternatives.[i] A third set of models including an interaction term for CCT and D_{uv} found that effect to not be statistically significant for any of the dependent measures, contrasting prior work.

Regression analysis can be used with R_f, R_g, and $R_{cs,h1}$ (third order) as continuous predictors replacing colour rendition condition and CCT and D_{uv}, plus their interaction, replacing chromaticity group. This model returned similar results, as illustrated in Table 4. The overall coefficient of determination for these models is relatively low ($r^2 <= 0.50$). The experiment was designed to counterbalance order effects, which only occurs with mean data. When ANOVA tests were performed with mean data, colour rendition condition maintained its significance, but only D_{uv} level for saturation was statistically significant among the variables associated with chromaticity. No variables related to chromaticity were statistically significant when modelled as continuous predictors in linear regression (Table 4). With mean data, the coefficient of determination for each model was greater than or equal to 0.83.

Additional linear regression models were fitted to the mean data using various measures from TM-30, alone or in combination. These models do not include chromaticity as a predictor, as it provided little increase in predictive power. Figure 5 illustrates the performance for colour preference, based on the coefficient of determination (r^2). With $r^2 = 0.74$, the best single predictor was $R_{cs,h1}$ (using a third-order polynomial fit). Adding R_f to the model resulted in the best performance of a metric pair, with $r^2 = 0.81$, and adding R_g to the latter improved the fit a small amount ($r^2 = 0.82$)—adding chromaticity improves the correlation to $r^2 = 0.83$. Colour preference was consistent across chromaticities, as partially illustrated by $R_{cs,h1}$ in Figure 6 ($r^2 \geq 0.68$ for all nine groups).

Mean saturation ratings were linearly correlated with $R_{cs,h16}$ ($r^2 = 0.89$)—and to a lesser extent R_g ($r^2 = 0.86$) and $R_{cs,h1}$ ($r^2 = 0.76$). When R_g is combined with $R_{cs,h16}$ or $R_{cs,h1}$, r^2 increases to 0.93 and 0.92, respectively. The difference between $R_{cs,h1}$ and $R_{cs,h16}$ is perhaps due to the design of the colour rendition conditions to vary

[i] This occurred despite the gamut shape differences for condition 4. The D_{uv} = 0.000 iterations of condition 4 were rated more saturated, counter to the overall trend.

Table 4. ***p*** **values for predictors in linear regression models using individual or mean data.** Participant was a categorical predictor but cannot be assigned as random in regression models. Red shading indicates statistically significant predictors.

	p value							
	Individual Data				Mean Data (per Scene)			
	Norm	**Sat**	**Pref**	**Accept**	**Norm**	**Sat**	**Pref**	**Accept**
Participant	<0.001	<0.001	<0.001	<0.001	*(Not Applicable)*			
R_f	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001
R_g	<0.001	<0.001	0.004	0.001	<0.001	<0.001	0.016	0.002
$R_{cs,h1}$	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001	<0.001
$R_{cs,h1}^2$	<0.001	0.001	<0.001	<0.001	<0.001	0.008	0.001	<0.001
$R_{cs,h1}^3$	0.534	0.015	0.011	0.398	0.521	0.053	0.033	0.288
CCT	0.289	0.705	0.697	0.648	0.275	0.760	0.734	0.662
D_{uv}	0.959	0.026	0.106	0.456	0.957	0.076	0.170	0.418
CCT* D_{uv}	0.901	0.058	0.139	0.508	0.898	0.129	0.207	0.463
r^2	0.35	0.51	0.37	0.31	0.83	0.94	0.83	0.86

other quantities while holding $R_{cs,h1}$ constant. As was found previously,[4] and in line with the ANOVA results, there are differences in rated saturation based on chromaticity group (Figure 7), with D_{uv} being the statistically significant differentiator. More than for other rated attributes, it is speculated that the observers may have considered saturation of the light, rather than the appearance of only the objects.

3.2 Influential objects and hues

As with others that have reported on equivalent questions,[4, 5, 15, 38-40] red was the hue most frequently indicated (72%) as being in the top three most influential colours (Figure 8). Compared to prior responses

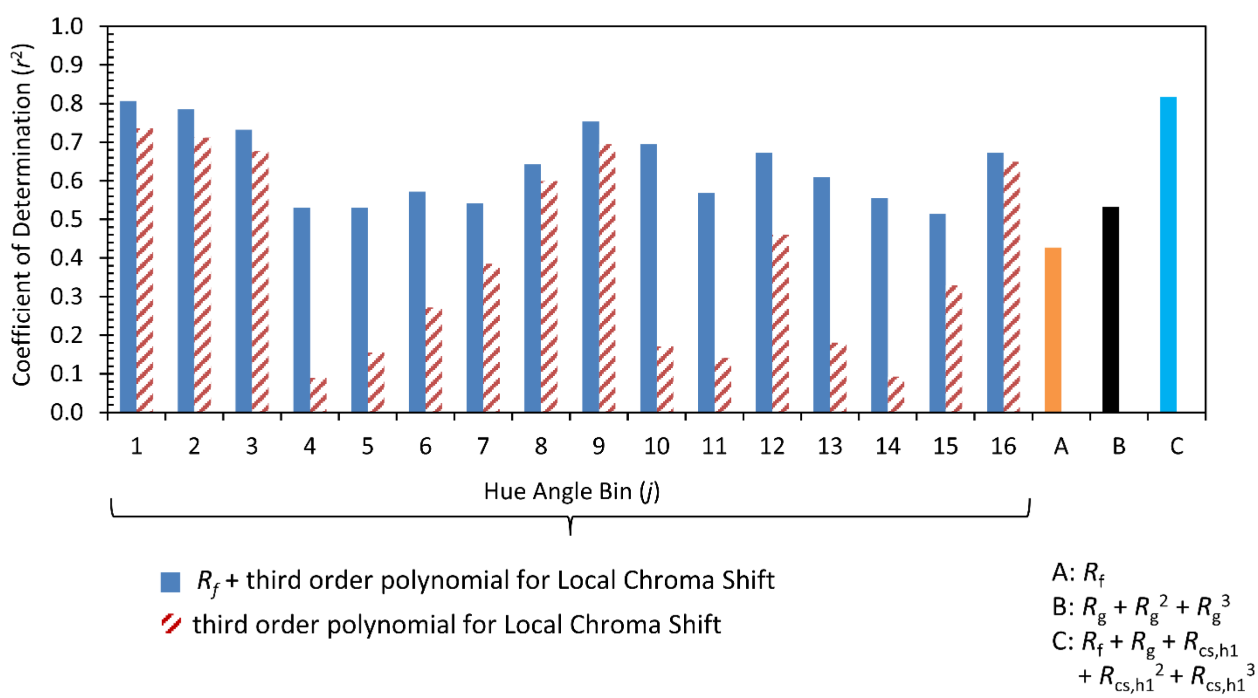

Figure 5. Regression models for mean color preference using ANSI/IES TM-30-18 measures.

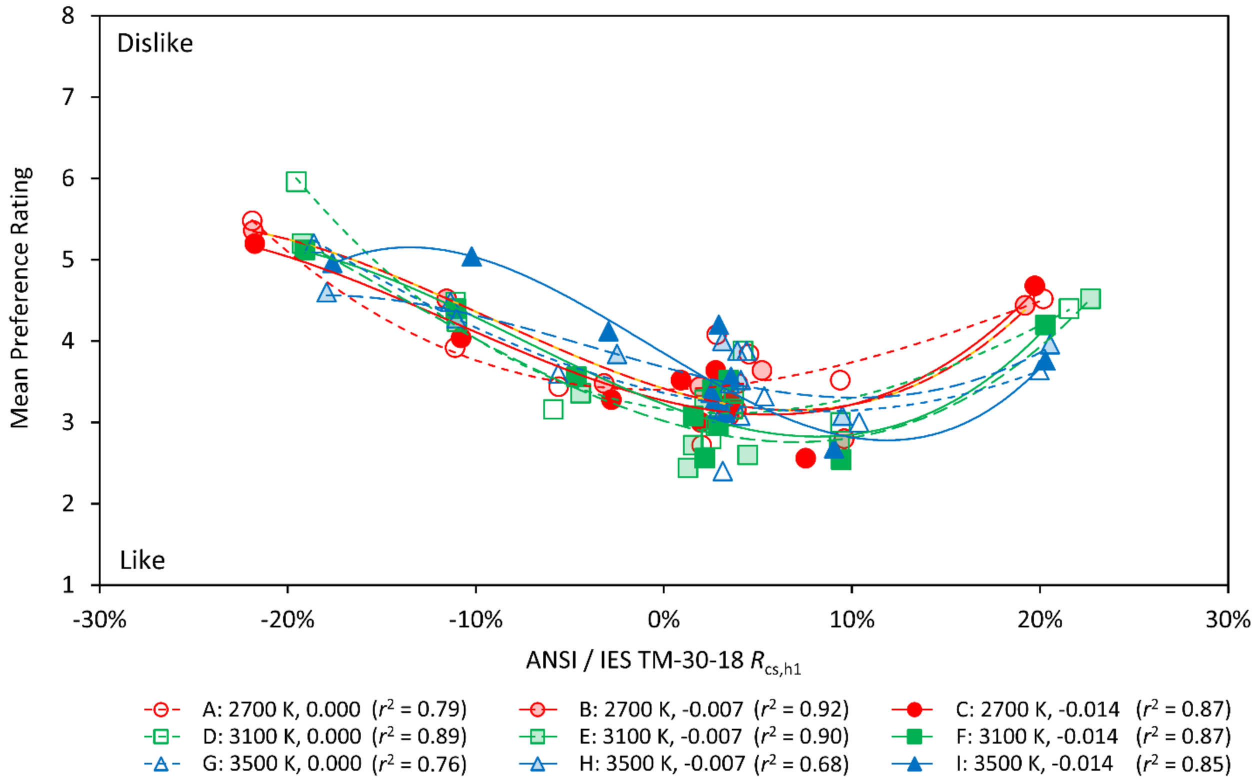

Figure 6. Mean preference rating versus ANSI/IES TM-30-18 $R_{cs,h1}$ with differentiation for each chromaticity group.

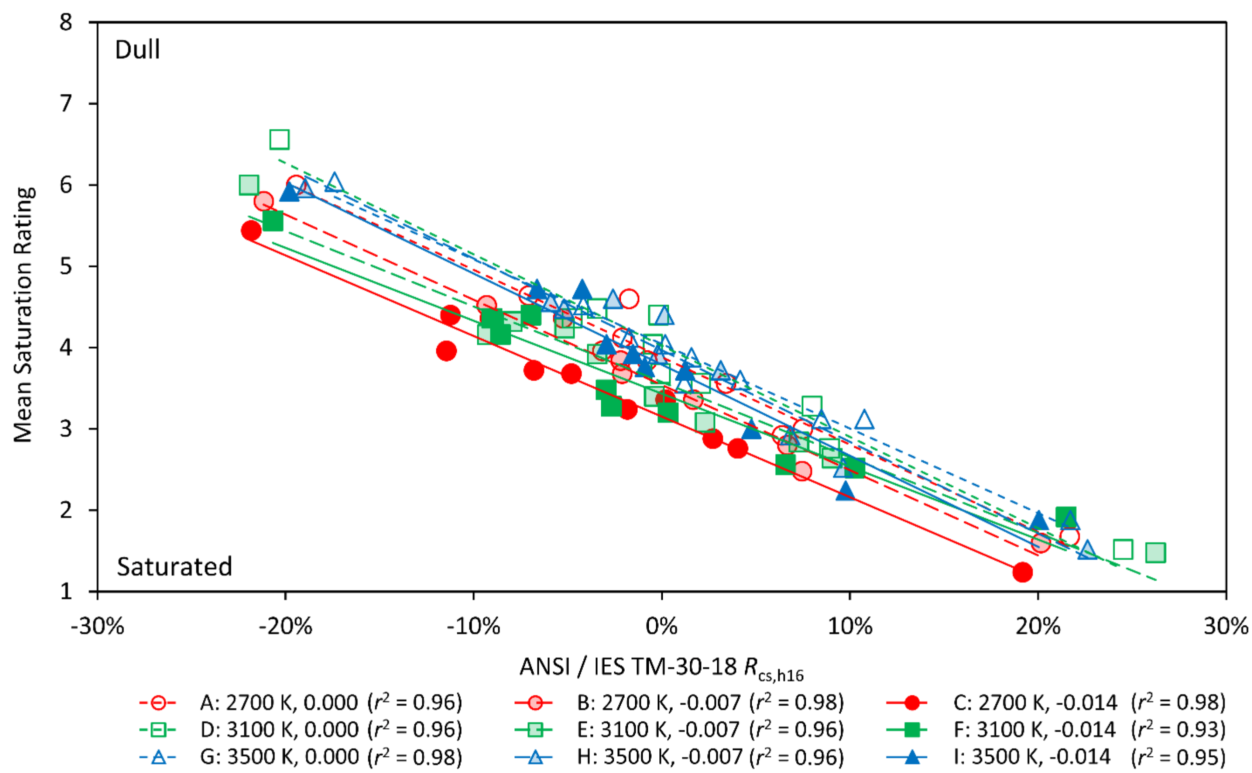

Figure 7. Mean saturating rating versus ANSI/IES TM-30-18 $R_{cs,h16}$ with differentiation for each chromaticity group.

from this line of experiments, the percentages for red were lower, with yellow gaining the most—likely due to the change in size and prevalence of different objects. Likewise, the selection of influential objects was more diverse, with the strawberries (8), green pepper (6), purple coat (6), and flower photo (6) receiving the most identifications among the three most influential objects. All others had fewer than four, including the Coke box (3). This change in results juxtaposes the similarity in models of subjective qualities, perhaps illustrating less of an effect of specific objects. This idea is bolstered by the similarity of findings from two other studies that used substantially different viewing arrangements.[10, 11]

This version of the final questionnaire included a new question regarding the aspects of influential objects that were important. The results are provided in Figure 9. Based on previous responses on influential object groups, it was expected that object familiarity would be a key factor; however, it was ranked fourth, behind the colour of the objects, the amount they varied, and their contrast with adjacent colours. Personal preferences and room position received few selections (participants were instructed to choose one or more). Overall, these results continue to support the inclusion of red-specific Local Chroma Shift in specification criteria and indicate reasonably broad applicability of the recommendations.

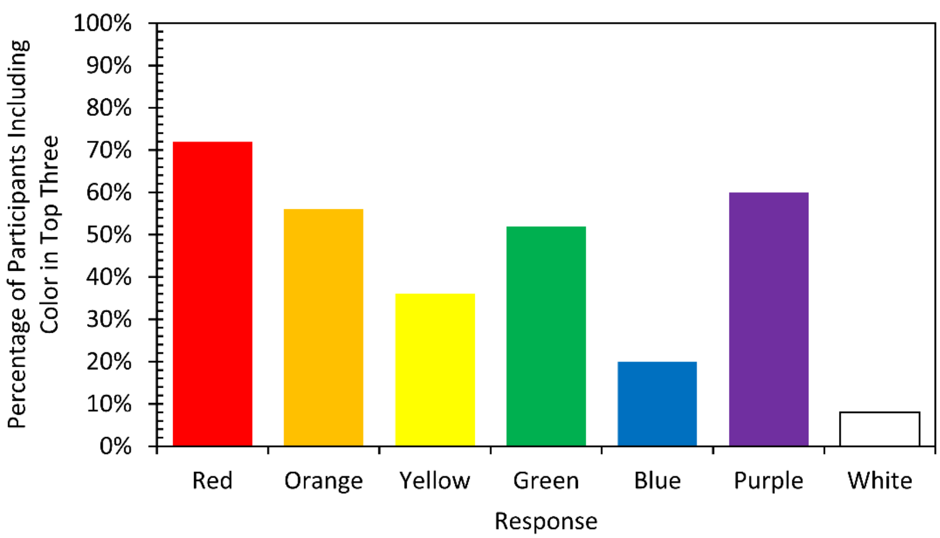

Figure 9. Percentage of participants ranking each color among the top three most influential.

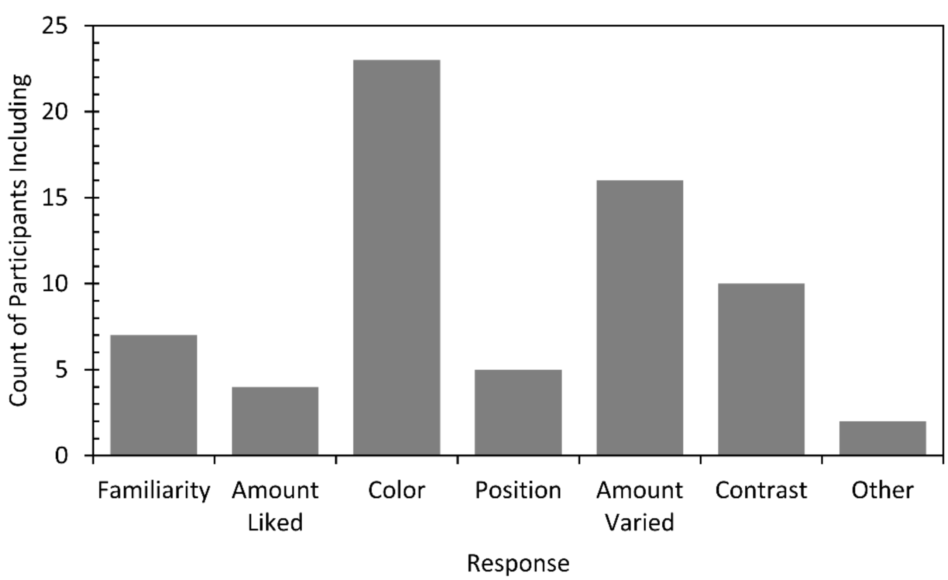

Figure 9. Count of participants choosing each characteristic as contributing to the color ranking.

4 Discussion

4.1 The role of chromaticity

The results of this experiment, where chromaticity was only a significant factor for rated saturation, contrast with the previous experiment, which indicated statistical significance for at least one variable related to chromaticity (CCT, D_{uv}, or CCT*D_{uv}) for all four studied outcome measures. However, this contrast requires context: the trends did not change, just the level of statistical significance. There were several differences between this experiment and its predecessor that may have shifted chromaticity effects enough to alter the statistical significance. These include the number of chromaticity groups, the range of D_{uv}, the range of CCT, the types of colour rendition conditions, the objects in the experimental space, greater consistency in colour rendition conditions between chromaticity groups, and pure randomization instead of an initial replication at 3500 K. In summary, the trends agree with past findings[4, 8] and some research on perceptions of whiteness for light,[67] but the overwhelmingly dominant factor for the appearance of objects is colour rendition. Note that this experiment did not address the appearance of the light itself. The findings sharply contrast the concept of an overall CCT preference, which may occur only when chromatic adaptation is not sufficiently addressed. Additional review of this topic, with a focus on methodology, is warranted.

4.2 Adjustments to colour rendition specification criteria

The proposed colour rendition specification criteria generally performed well. However, they were designed to be conservative, with minimum thresholds only as low as examined in experiments. Three adjustments are recommended based on this experiment, where SPDs were designed to fall just outside of specific thresholds.

- Colour rendition condition 3 had a mean R_g value of 95, just below the previously proposed thresholds of 100 (Tier A) and 98 (Tier B). This condition ranked fourth for colour preference, with a rating not statistically different from the other four conditions in Tier A. This suggests a reduction in the R_g criterion for both tiers is warranted.

- Colour rendition condition 2 had a mean R_f value of 75, just below the previously proposed threshold of 78 for Tiers A and B. The mean colour preference rating of this condition was not statistically significantly different from the rating for the condition in Tier B or the two lowest-rated conditions in Tier A. This suggests a reduction in the R_f criterion for Tier B.

- Colour rendition condition 9 had a mean $R_{cs,h1}$ value of 20%, exceeding the previously proposed upper limit of 15%. This condition was grouped with colour rendition condition 5 according to statistical analysis. This suggests increasing the upper limit for $R_{cs,h1}$ for Tier C.

In addition to these observations, the proposed colour rendition criteria were recently applied to the results from two other independent experiments, Zhang *et al.*[11] and Esposito and Houser,[10] that examined colour preference over a wide range of colour rendition conditions. This analysis[14] supports the identified adjustments, as well as smoothing for even increments. The final proposed colour preference specification criteria are:

- Tier A: $R_f \geq 78$, $R_g \geq 95$, $-1\% \leq R_{cs,h1} \leq 15\%$
- Tier B: $R_f \geq 74$, $R_g \geq 92$, $-7\% \leq R_{cs,h1} \leq 19\%$
- Tier C: $R_f \geq 70$, $R_g \geq 89$, $-12\% \leq R_{cs,h1} \leq 23\%$

Future research could explore the refinement of these specification criteria by examining different SPDs, different viewing contexts, and/or different illuminance levels.

Colour preference regression models

While specification criteria are the most frequent way that colour rendition is incorporated into the lighting design process, research has typically focused on regression models to evaluate the performance of various metrics and measures. Models based on ANSI/IES TM-30-18 measures, including those using the same measures as found in the specification criteria, were compared against a variety of previously proposed metrics and recommended combinations (Table 5). The best-fit TM-30-based combination outperformed all other contenders, as others have also found.[11] Still, such models are not recommended for practical use; even though the terms stay the same, the coefficients can vary considerably from experiment to experiment, and it is difficult to establish a universal model.

Single-number, hue-averaged metrics (e.g., Q_a, R_m, GVI, FCI, R_f, R_g, R_a, and G_a) generally performed poorly, especially when examining only linear relationships. This occurred despite the limited number of colour rendition conditions and is exacerbated with a wider variety of conditions, as has been explored previously.[5] The performance of gamut area-based measures improved when considering non-linear relationships, reflecting the effects of oversaturation. Another factor contributing to poor performance is CCT-bias, which often occurs with measures not using a relative reference illuminant (*e.g.,* R_m, GVI, GAI, FCI).[68, 69] The best performing models from a prior review[70] generally performed poorly. Success with prior results likely arises because many older experiments did not consider gamut shape,[63] which was not well described or quantified at the time. It has been demonstrated that fixing gamut shape can greatly improve correlations for hue-averaged metrics, masking poor performance when a wide variety of colour rendition conditions is considered.[71]

The best alternative to the best-fit TM-30 model uses conceptually equivalent measures but relies on the CIE 13.3-1995 framework[62]—including the scant colour samples, obsolete colour space, and wrong chromatic adaptation transformation, all of which can lead to inaccurate results.[6, 9, 12, 13, 69, 71-74] While having a relatively small effect on correlations—due to the induced correlation between the measures themselves[69]—the inaccuracies of the CIE 13.3-1995 method have substantial consequences when trying to establish specification criteria. For example, the condition with the highest mean rating for colour preference had a R_f value of 81, but a CIE R_a value of 74. Considering a large set of theoretical SPDs[75, 76]—developed for a previous analysis—it is possible to meet Tier A of the adjusted TM-30 colour preference criteria but have a R_a value of 61. This number drops to 51 for Tier B. Thus, specifications based on a method that incorporates the CIE 13.3-1995 calculation components would require unusually low thresholds for R_a, which would end up qualifying non-preferred products, or higher thresholds that would not qualify some preferred products.

4.3 Limitations

The experimental room used in this work did not provide an identifiable application, which may influence how colour shifts are evaluated. Only one illuminance level was used; due to the Hunt effect,[62, 63] colour quality evaluations can change with illuminance.[7] The applicability of these short-term evaluations to long-term perceptions is unverified, although the general findings are supported by trends in colour psychology, photography, digital display, and other lighting systems (e.g., neodymium incandescent).

Table 5. Performance of various metrics and models for predicting mean color preference ratings. The measures are arranged in rank order.

Base Method	Measure Type(s)	Model Components	r^2
ANSI/IES TM-30-18	Fidelity, Gamut Area, Red Chroma	R_f, R_g, and $R_{cs,h1} + R_{cs,h1}^2 + R_{cs,h1}^3$	0.82
ANSI/IES TM-30-18 (Cust. CES)	Fidelity, Gamut Area, Red Chroma	R_f, R_g, and $R_{cs,h1} + R_{cs,h1}^2 + R_{cs,h1}^3$	0.82
ANSI/IES TM-30-18	Fidelity, Red Chroma	R_f and $R_{cs,h1} + R_{cs,h1}^2 + R_{cs,h1}^3$	0.81
ANSI/IES TM-30-18	Fidelity, Gamut Area, Red Chroma	R_f, R_g, $R_{cs,h1}$ and R_g*$R_{cs,h1}$	0.80
CIE 13.3-1995 (Expanded)	Fidelity, Gamut Area, Red Chroma	R_a, G_a, and $C_9 + C_9^2 + C_9^3$	0.78
CIE 13.3-1995 (Expanded)	Fidelity, Red Chroma	R_a and $C_9 + C_9^2 + C_9^3$	0.78
ANSI/IES TM-30-18	Red Chroma Shift	$R_{cs,h1} + R_{cs,h1}^2 + R_{cs,h1}^3$	0.74
CIE 13.3-1995 (Expanded)	Red Chroma Shift	$C_9 + C_9^2 + C_9^3$	0.73
CIE 13.3-1995 (Expanded)	Fidelity, Red Chroma	R_a and C_9	0.73
CIE 13.3-1995 (Expanded)	Fidelity, Gamut Area	R_a, and $G_a + G_a^2 + G_a^3$	0.68
CIE 13.3-1995 (Expanded)	Gamut Area	$G_a + G_a^2 + G_a^3$	0.68
Other	Weighted Color Fidelity	Q_a	0.55
ANSI/IES TM-30-18	Gamut Area	$R_g + R_g^2 + R_g^3$	0.53
CIE 13.3-1995 (Expanded)	Fidelity, Gamut Area	R_a and G_a	0.49
CIE 13.3-1995 (Expanded)	Fidelity, Gamut Area	R_a and FCI	0.48
ANSI/IES TM-30-18	Fidelity, Gamut Area	R_f and R_g	0.48
ANSI/IES TM-30-18	Average Color Fidelity	R_f	0.43
CIE 13.3-1995	Fidelity, Red Fidelity	R_a and R_9	0.36
CIE 13.3-1995 (Expanded)	Fidelity, Gamut Area	R_a and GAI	0.33
Other	Memory Color	R_m	0.29
CIE 13.3-1995 (Expanded)	Fidelity, Gamut Area	(R_a + GAI)/2	0.28
CIE 13.3-1995	Average Color Fidelity	R_a	0.21
Other	Gamut Area	FCI	0.14
Other	Gamut Volume	GVI	0.12
CIE 13.3-1995 (Expanded)	Gamut Area	G_a	0.11
ANSI/IES TM-30-18	Gamut Area	R_g	0.08

5 Conclusions

The primary focus of this experiment was to explore the performance of a previously proposed set of colour preference specification criteria. The secondary goal was to further explore the role of chromaticity on subjective evaluations of colour quality. A set of 90 lighting scenes with varied colour rendition and varied chromaticity were evaluated by 25 experiment participants, who provided ratings for normalness, saturation, preference, and acceptability. While there were substantial differences in ratings based on colour rendition, there were minimal differences based on chromaticity, with only D_{uv} level being a statistically significant factor for ratings of saturation. This supports the development of colour rendition specification criteria that are independent of chromaticity—within the range of nominally white light.

The colour preference specification criteria performed well, correctly identifying four of the top five rated colour rendition conditions. Based on the new data, the ranges of allowable characteristics in each tier were slightly expanded, so that Tier A identified the five highest-rated conditions, Tier B the next two, and Tier C the next two conditions. The new criteria, based on ANSI/IES TM-30-18 measures, are:

- Tier A: $R_f \geq 78$, $R_g \geq 95$, $-1\% \leq R_{cs,h1} \leq 15\%$
- Tier B: $R_f \geq 74$, $R_g \geq 92$, $-7\% \leq R_{cs,h1} \leq 19\%$
- Tier C: $R_f \geq 70$, $R_g \geq 89$, $-12\% \leq R_{cs,h1} \leq 23\%$.

These criteria have been validated with independent datasets. The included measures align with best-fit regression models, which demonstrated the improved performance of ANSI/IES TM-30-18 measures compared to other research proposals.

Acknowledgements

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This project was supported by Solid-State Lighting program of the U.S. Department of Energy. The contribution of Minchen Wei was supported by The Public Policy Research Funding Scheme from Policy Innovation and Co-ordination Office of the Hong Kong Special Administrative Region Government (Project Number: 2017.A6.091.18A).

Appendix A: Stimulus characterization

This appendix provides additional data regarding the characterization of the stimulus. Figures A1 and A2 provide the 90 SPDs, grouped by colour rendition conditions. With a few exceptions, the spectral features were equivalent across the nine chromaticity groups, with only the peak intensity of the channels varying. This is further illustrated in Figure A3, which provides CVGs for colour rendition condition 6 across the nine chromaticity groups. For one of the conditions, number 4, the gamut shape could not be maintained across all nine chromaticities; for chromaticities on the Planckian Locus, the gamut shape was similar to that shown in Figure A4a, and for the other chromaticity groups it was similar to that shown in Figure A4b. Table A1 documents the average colour rendering conditions across the three CCT levels and three D_{uv} levels. The similarity of these values helps justify collapsing colour rendition condition to a categorical variable for use in analysis of variance (ANOVA).

Table A1. Mean color rendition characteristics

	Nominal CCT (K)			Nominal D_{uv}		
	2700	**3100**	**3500**	**0.000**	**-0.007**	**-0.014**
R_f	80	80	81	80	80	80
R_g	102	101	101	101	102	102
$R_{cs,h1}$	1%	1%	1%	1%	1%	1%
$R_{cs,h5}$	-2%	-5%	-5%	-3%	-4%	-4%
$R_{cs,h9}$	5%	5%	5%	-3%	-4%	-4%
$R_{cs,h13}$	0%	-1%	-1%	0%	-1%	-1%
$R_{cs,h16}$	-1%	0%	1%	1%	0%	-1%
R_a	75	77	77	78	77	75
R_9	30	28	31	28	29	31
GAI	62	72	81	61	72	81
G_a	104	104	105	104	104	105
C9	103	103	104	103	103	104
R_m	88	89	90	88	89	90
CQS Q_a	82	82	82	81	82	83
FCI	106	109	110	114	108	103

Figure A5 documents the consistency of illuminance distribution using high dynamic range photographs of four randomly selected conditions.

Figure A6 addresses the correlation of custom measures of colour rendition based on the objects in the experimental room to standard ANSI/IES TM-30-18 calculations. The agreement is generally strong, but the custom $R_{cs,h1}$ values are consistently lower by approximately 5% compared to ANSI/IES TM-30-18 values. This behaviour was not seen for the prior object sets and is also not seen for the adjacent hue angle bins (1 and 16). This was explored further by examining the objects in hue angle bins 1 and 2 (Figure A7). The objects in bin 1 have different spectral reflectance characteristics and a different distribution in the *a*'-*b*' plane of CAM02-UCS than do the CES in the same hue-angle bin. Several of the experimental objects fell just across the adjacent hue-angle bin borders. The discrepancy in spectral reflectances does not exist for hue-angle bin 2. It is concluded that this anomaly is largely an artefact of the binning method and is not a cause for concern when analysing the data using standardized measures.

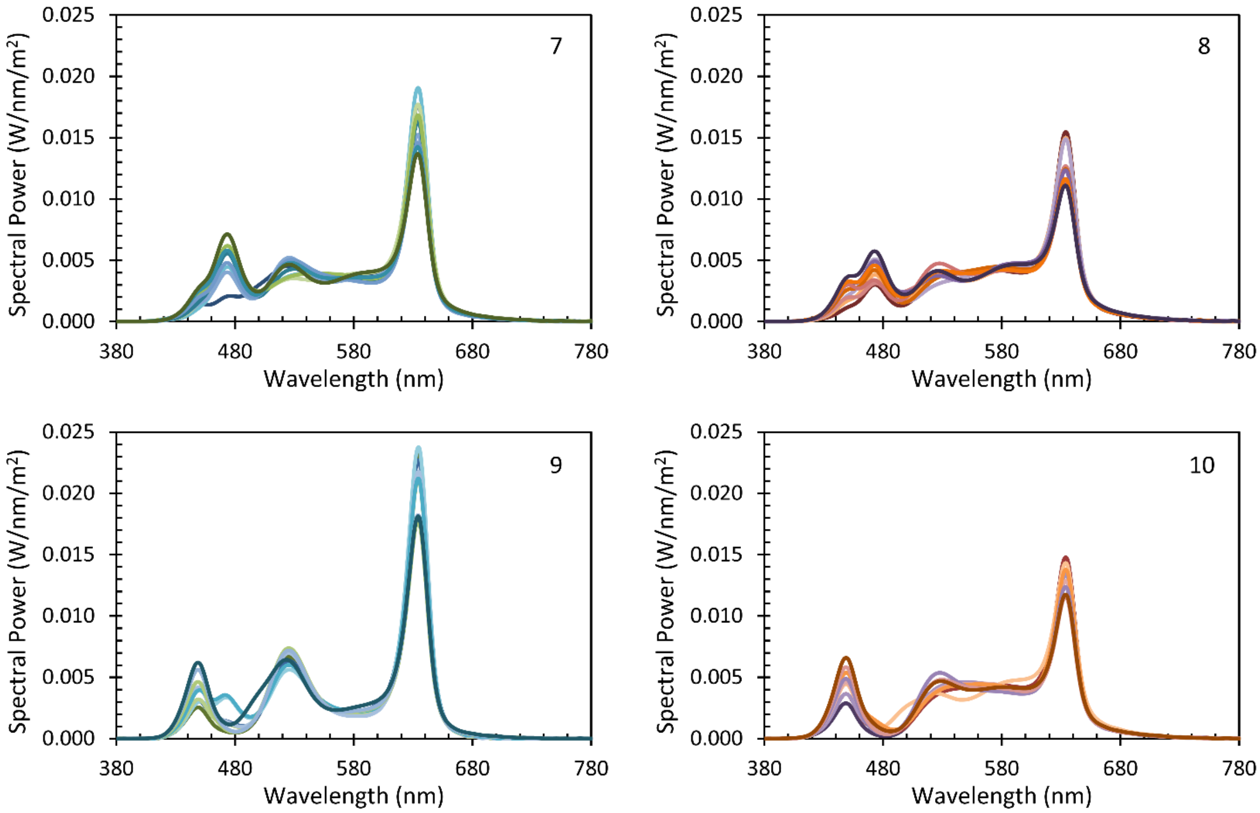

Figure A2. SPDs for color rendition conditions 7-10.

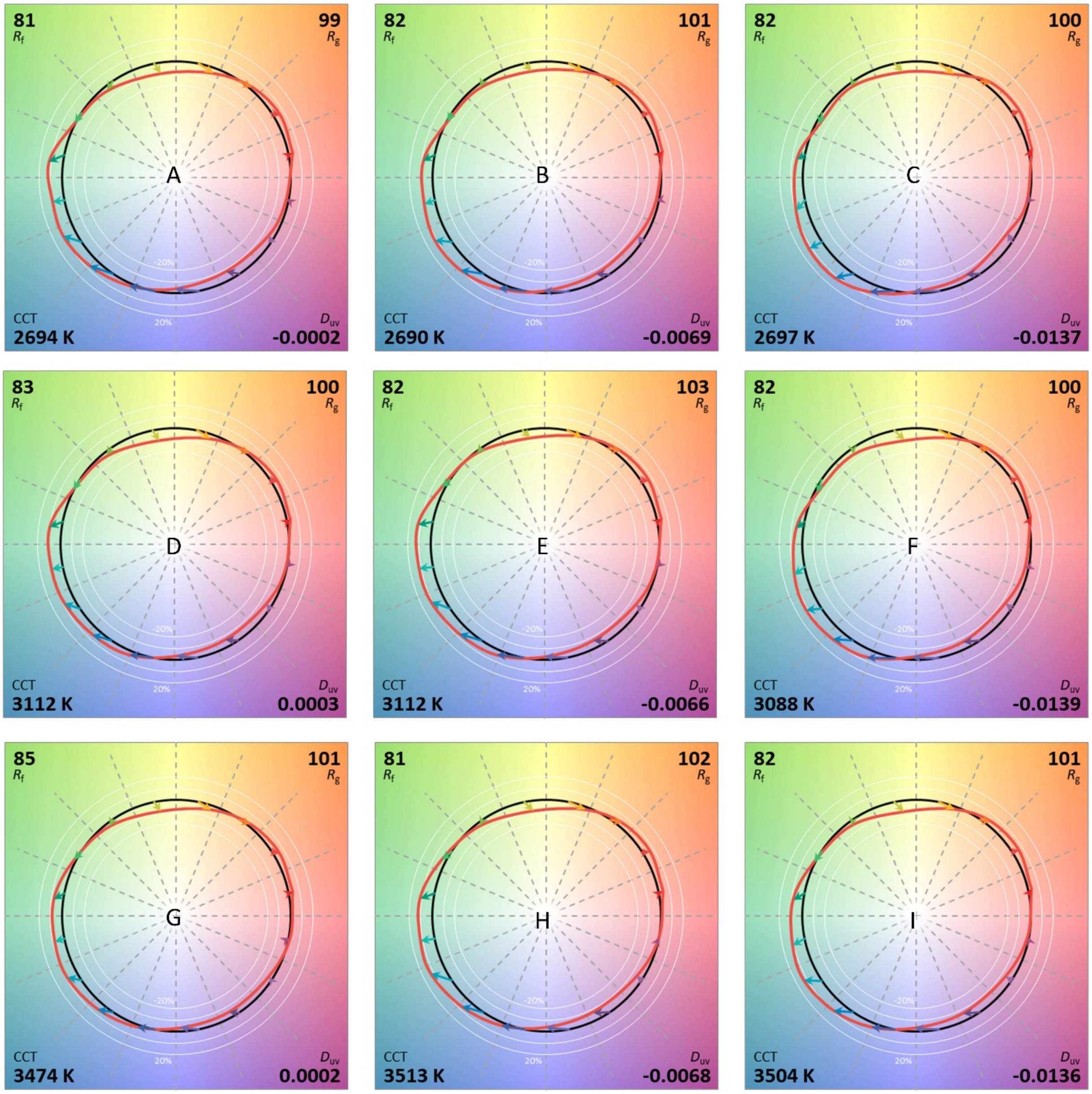

Figure A3. ANSI/IES TM-30-18 Color Vector Graphics for color rendition condition 6 in each of the nine chromaticity groups.

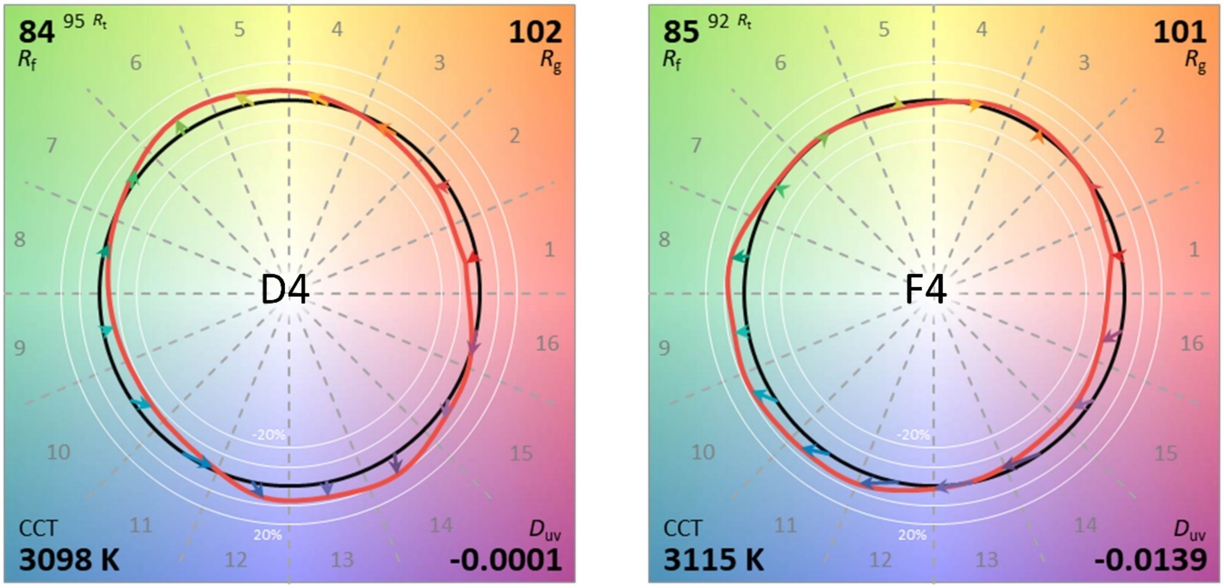

Figure A4. Illustration of deviation on gamut shape for example SPDs within color rendition condition 4.

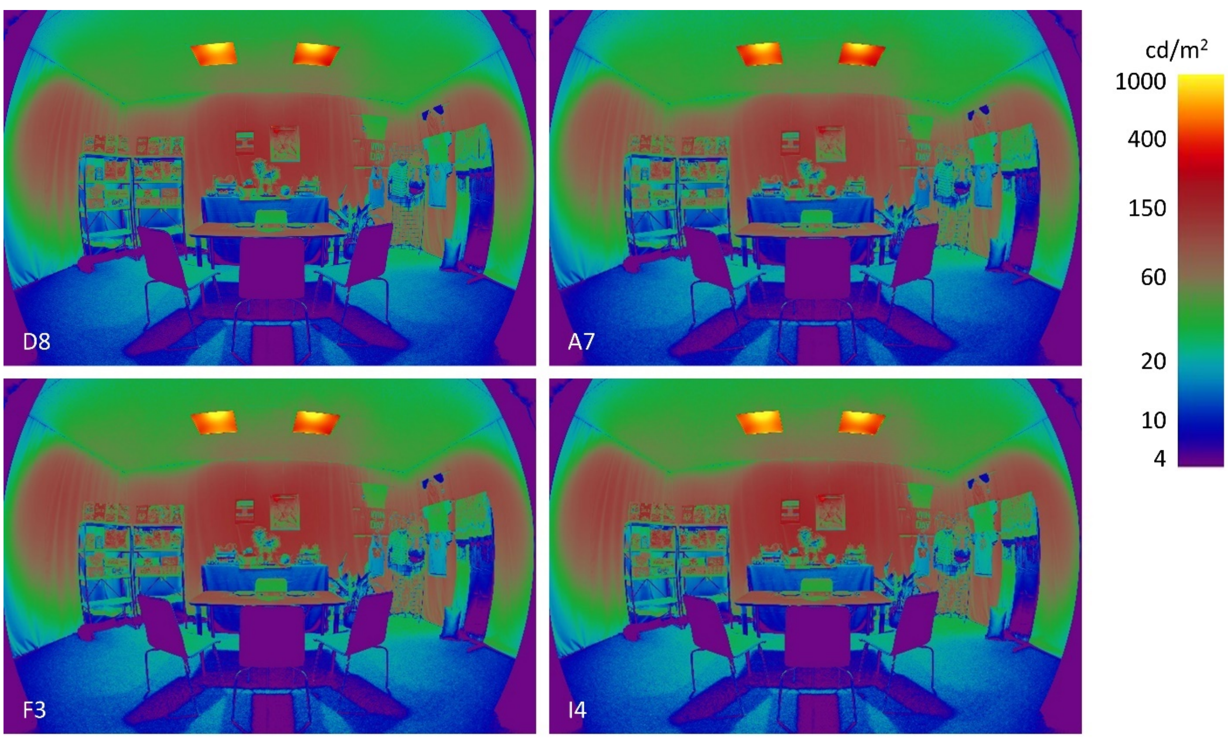

Figure A5. Illustration of consistency in luminance between scenes using HDR photography.

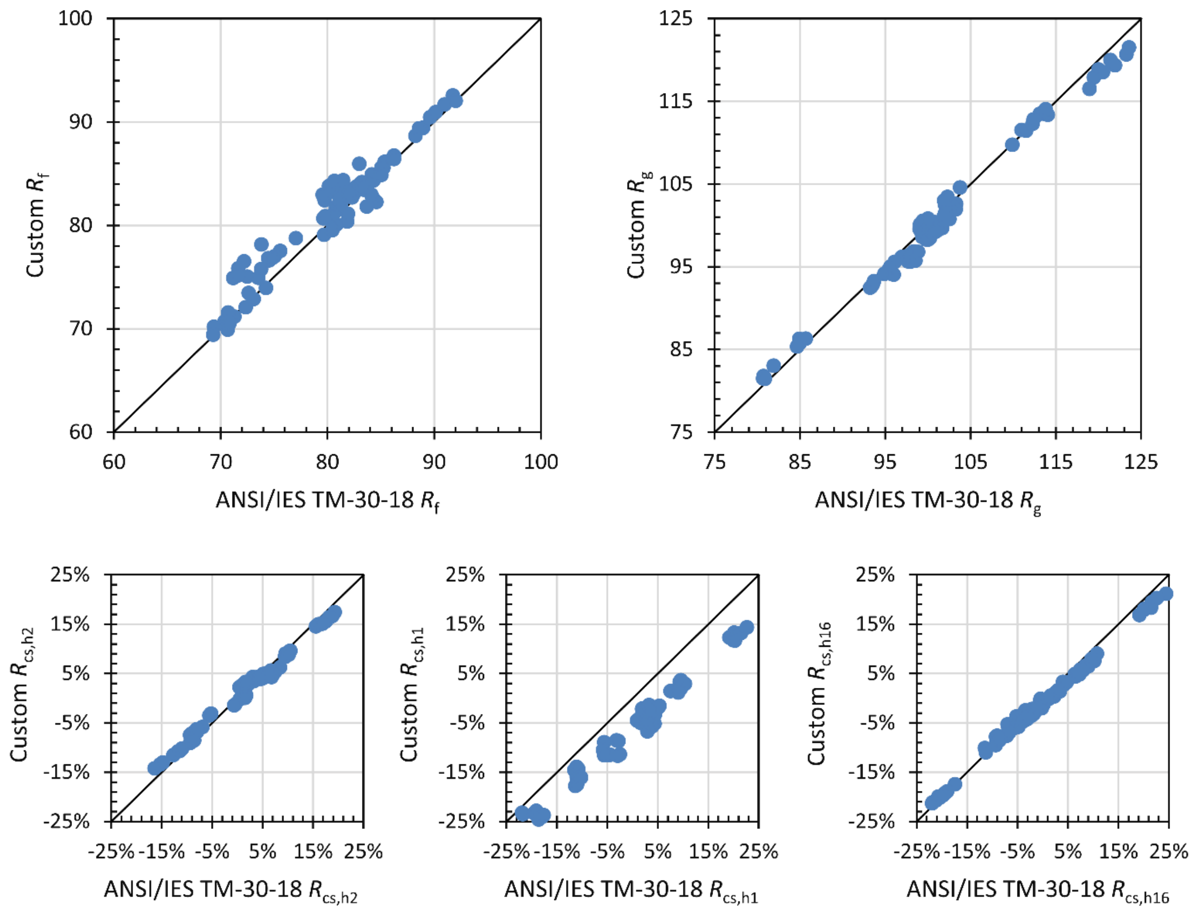

Figure A6. Comparison of standard ANSI/IES TM-30-18 measures of color rendition to customized measures using spectral reflectance measurements of the experimental objects instead of the standard 99 color evaluation samples.

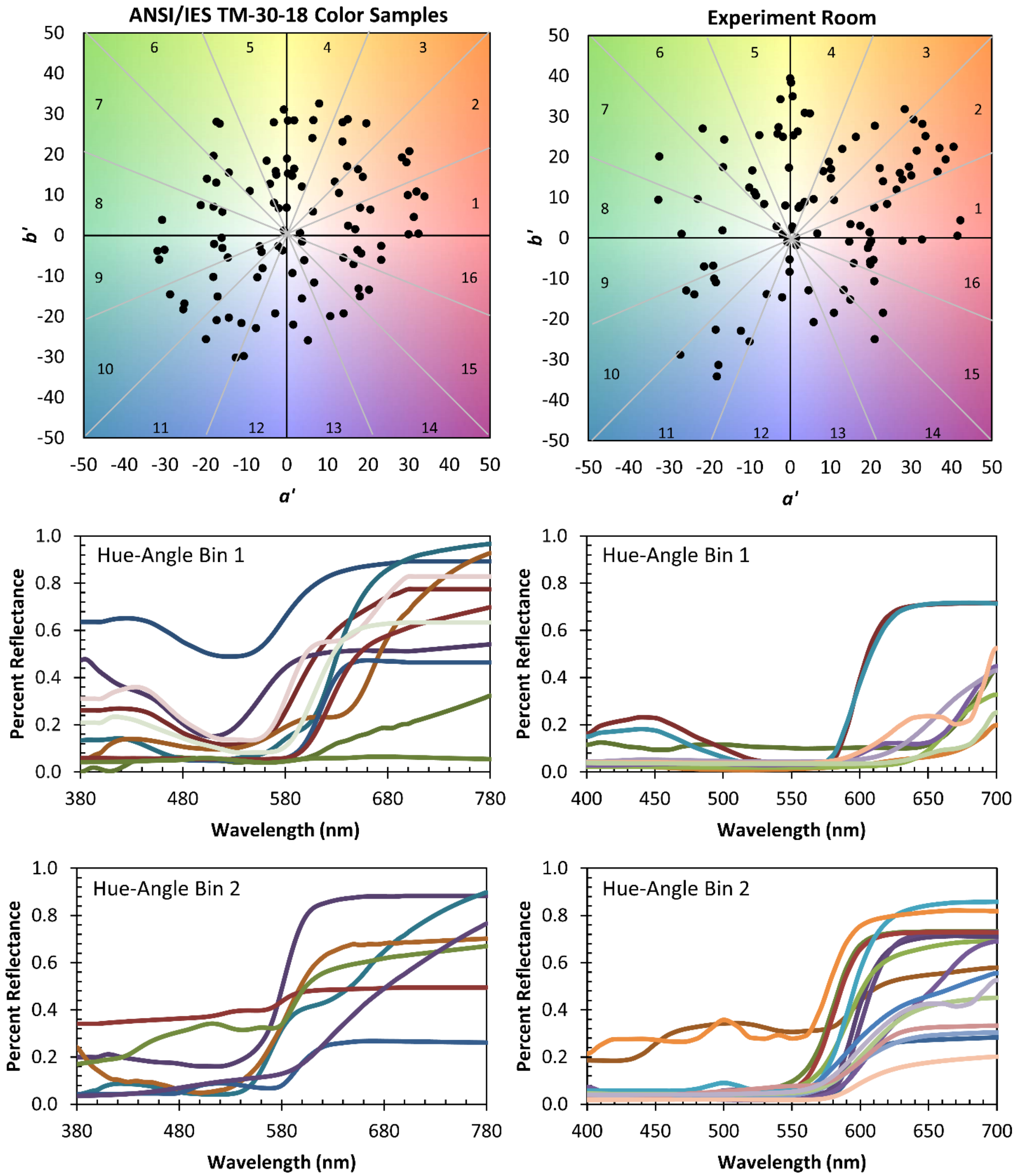

Figure A7. Comparison of the spectral reflectance functions for hue-angle-bins 1 and 2 for standardized color evaluation samples and measured spectral reflectance functions.

Appendix B: Response form

For the first three questions, please circle the number on the scale that most accurately characterizes your impression of the space, based on the words provided as anchors at either end of the scale. You may only circle one number, and you may not place your circle between numbers. For the fourth question, please circle one of the two choices.

1. This lighting makes the colors of the objects in the space appear: **Normal** 1 2 3 4 5 6 7 8 **Shifted**
2. This lighting makes the colors of the objects in the space appear: **Saturated** 1 2 3 4 5 6 7 8 **Dull**
3. My overall opinion of how the lighting makes the objects appear is: **Like** 1 2 3 4 5 6 7 8 **Dislike**
4. In general, the way this lighting makes the colors of the objects appear is: **Acceptable** **Unacceptable**

References

1. Illuminating Engineering Society of North America . IES-TM-30-15 Method for Evaluating Light Source Colour Rendition. New York, NY: IESNA, 2015. 26.
2. International Commision on Illumination. 224:2017 CIE 2017 Colour Fidelity Index for accurate scientific use. Vienna, Austria: CIE, 2017. 54.
3. Illuminating Engineering Society of North America . IES-TM-30-18 Method for Evaluating Light Source Colour Rendition. New York, NY: IESNA, 2018. 34.
4. Royer MP, Wilkerson A, Wei M. Human perceptions of colour rendition at different chromaticities. *Lighting Research and Technology* 2017; 50: 965-994. DOI: 10.1177/1477153517725974.
5. Royer MP, Wilkerson A, Wei M, Houser K, Davis R. Human perceptions of colour rendition vary with average fidelity, average gamut, and gamut shape. *Lighting Research and Technology* 2016; 49: 966-991. DOI: 10.1177/1477153516663615.
6. Wei M, Royer M, Huang H-P. Perceived colour fidelity under LEDs with similar Rf but different Ra. *Lighting Research and Technology* 2019; Online before print: DOI: 10.1177/1477153519825997.: 1477153519825997.
7. Wei M, Bao W, Huang H. Consideration of Light Level in Specifying Light Source Colour Rendition. *Leukos* 2018; Online Before Print. DOI: 10.1080/15502724.2018.1448992.
8. Wang Y, Wei M. Preference among light sources with different Duv but similar colour rendition: A pilot study. *Lighting Research and Technology* 2017; Online Before Print. DOI: 10.1177/1477153517712552.
9. Xu W, Wei M, Smet KAG, Lin Y. The prediction of perceived colour differences by colour fidelity metrics. *Lighting Research and Technology* 2016; 49: 805-817. DOI: 10.1177/1477153516653650.
10. Esposito T, Houser K. Models of colour quality over a wide range of spectral power distributions. *Lighting Research and Technology* 2018; Online Before Print. DOI: 10.1177/1477153518765953.
11. Zhang F, Xu H, Feng H. Toward a unified model for predicting colour quality of light sources. *Appl Opt* 2017; 56: 8186-8195. 2017/10/20. DOI: 10.1364/AO.56.008186.
12. Jost S, Cauwerts C, Avouac P. CIE 2017 colour fidelity index Rf: a better index to predict perceived colour difference? *J Opt Soc Am A Opt Image Sci Vis* 2018; 35: B202-B213. 2018/04/01. DOI: 10.1364/JOSAA.35.00B202.
13. Gu HT, Luo MR, Liu XY. Testing different colour rendering metrics using colour difference data. *Lighting Research and Technology* 2017; 49: 539-560. DOI: 10.1177/1477153516653649.
14. Royer M. *Analysis of Colour Rendition Specification Criteria*. San Francisco, CA: SPIE, 2019.
15. Rea MS, Freyssinier JP. Colour Rendering: Beyond Pride and Prejudice. *Colour Research and Application* 2010; 35: 401-409. DOI: 10.1002/col.20562.
16. Teunissen C, van der Heijden F, Poort SHM, de Beer E. Characterising user preference for white LED light sources with CIE colour rendering index combined with a relative gamut area index. *Lighting Research and Technology* 2016; 49: 461-480. DOI: 10.1177/1477153515624484.
17. Houser KW, Wei M, David A, Krames MR, Shen XS. Review of measures for light-source colour rendition and considerations for a two-measure system for characterizing colour rendition. *Optics Express* 2013; 21: 10393-10411. DOI: 10.1364/OE.21.010393.
18. Wei M, Houser KW, Allen GR, Beers WW. Colour Preference under LEDs with Diminished Yellow Emission. *Leukos* 2014; 10: 119-131. DOI: 10.1080/15502724.2013.865212.
19. Wei M, Houser K, David A, Krames M. Perceptual responses to LED illumination with colour rendering indices of 85 and 97. *Lighting Research and Technology* 2014; 47: 810-827. DOI: 10.1177/1477153514548089.
20. Ohno Y, Fein G, Miller C. Vision Experiment on chroma saturation for colour quality preference. In: *28th CIE Session* Manchester, UK, Jun 28 - Jul 4 2015, p.2124. Vienna (Austria): Commision Internationale de l'Eclairage.
21. Liu A, Tuzikas A, Zukauskas A, Vaicekauskas R, Vitta P, Shur M. Cultural Preferences to Colour Quality of Illumination of Different Artwork Objects Revealed by a Colour Rendition Engine. *IEEE Photonics Journal* 2013; 5: 6801010. DOI: 10.1109/jphot.2013.2276742.

22. Islam MS, Dangol R, Hyvarinen M, Bhusal P, Puolakka M, Halonen L. User preferences for LED lighting in terms of light spectrum. *Lighting Research and Technology* 2013; 45: 641-665. DOI: 10.1177/1477153513475913.
23. Szabo F, Keri R, Schanda J, Csuti P, Mihalyko-Orban E. A study of preferred colour rendering of light sources: Home lighting. *Lighting Research and Technology* 2014; 48: 103-125. DOI: 10.1177/1477153514555536.
24. Jost-Boissard S, Avouac P, Fontoynont M. Assessing the colour quality of LED sources: Naturalness, attractiveness, colourfulness and colour difference. *Lighting Research and Technology* 2014; 47: 769-794. DOI: 10.1177/1477153514555882.
25. Dikel EE, Burns GJ, Veitch JA, Mancini S, Newsham GR. Preferred Chromaticity of Colour-Tunable LED Lighting. *Leukos* 2013; 10: 101-115. DOI: 10.1080/15502724.2013.855614.
26. Royer MP, Houser KW and Wilkerson AM. Colour discrimination capability under highly structured spectra. *Colour Research and Application* 2012; 37: 441-449. DOI: 10.1002/col.20702.
27. Rea MS, Freyssinier-Nova JP. Colour rendering: A tale of two metrics. *Colour Research and Application* 2008; 33: 192-202. DOI: 10.1002/col.20399.
28. Szabó F, Csuti P, Schanda J. Colour preference under different illuminants—new approach of light source coloour quality. In: *Light and Lighting Conference with Special Emphasis on LEDs and Solid State Lighting* Budapest, Hungary, 27-29 May 2009 2009, pp.PWDAS-43. CIE.
29. Veitch JA, Whitehead LA, Mossman M, Pilditch TD. Chromaticity-Matched but Spectrally Different Light Source Effects on Simple and Complex Colour Judgments. *Colour Research and Application* 2014; 39: 263-274. DOI: 10.1002/col.21811.
30. Jost-Boissard S, Fontoynont M, Blanc-Gonnet J. Perceived lighting quality of LED sources for the presentation of fruit and vegetables. *Journal of Modern Optics* 2009; 56: 1420-1432. DOI: 10.1080/09500340903056550.
31. Lin Y, Wei M, Smet KAG, Tsukitani A, Bodrogi P, Khanh TQ. Colour preference varies with lighting application. *Lighting Research and Technology* 2015; 49: 316-328. DOI: 10.1177/1477153515611458.
32. Smet K, Lin Y, Nagy BV, Nemeth Z, Duque-Chica GL, Quintero JM, Chen HS, Luo RM, Safi M, Hanselaer P. Cross-cultural variation of memory colours of familiar objects. *Optics Express* 2014; 22: 32308-32328. DOI: 10.1364/OE.22.032308.
33. Smet K, Ryckaert WR, Pointer MR, Deconinck G and Hanselaer P. Memory colours and colour quality evaluation of conventional and solid-state lamps. *Optics Express* 2010; 18: 26229-26244. DOI: 10.1364/OE.18.026229.
34. Spaulding JM. Evaluation of Desirability Assessment Techniques for Tunable Solid State Lighting Applications. In: *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* 22-26 Oct 2012, pp.643-647.
35. Zukauskas A, Vaicekauskas R, Vitta P, Tuzikas A, Petrulis A, Shur M. Colour rendition engine. *Optics Express* 2012; 20: 5356-5367. DOI: 10.1364/OE.20.005356.
36. Zukauskas A, Vaicekauskas R, Shur MS. Colour-rendition properties of solid-state lamps. *Journal of Physics D-Applied Physics* 2010; 43: 354006. DOI:10.1088/0022-3727/43/35/354006.
37. Zukauskas A, Vaicekauskas R, Shur M. Solid-state lamps with optimized colour saturation ability. *Optics Express* 2010; 18: 2287-2295. DOI: 10.1364/OE.18.002287.
38. Wei MC, Houser KW. Systematic Changes in Gamut Size Affect Colour Preference. *Leukos* 2017; 13: 23-32. DOI: 10.1080/15502724.2016.1192402.
39. Wei M, Houser K, David A, Krames M. Effect of gamut shape on colour preference. *CIE 2016 Lighting Quality and Energy Efficiency*. Melbourne, Australia: Vienna (Austria): Commission Internationale de l'Eclairage., 2016, p. 32-41.
40. Wei M, Houser KW, David A, Krames MR. Colour gamut size and shape influence colour preference. *Lighting Research and Technology* 2016; 49: 992-1014. DOI: 10.1177/1477153516651472.
41. Davis W, Ohno Y. Colour quality scale. *Optical Engineering* 2010; 49: 033602. DOI: 10.1117/1.3360335.
42. Ohno Y, Davis W. Colour quality and spectra. *Photonics Spectra* 2008; 42: 71-72.

43. Liu Q, Huang Z, Pointer MR, Luo MR, Xiao K, Westland S. Evaluating colour preference of lighting with an empty light booth. *Lighting Research and Technology* 2017: 1477153517727330.
44. Huang Z, Liu Q, Westland S, Pointer MR, Luo MR, Xiao K. Light dominates colour preference when correlated colour temperature differs. *Lighting Research and Technology* 2017: 1477153517713542.
45. Huang Z, Liu Q, Pointer M, Luo M, Wu B, Liu A. White lighting and colour preference, Part 1: Correlation analysis and metrics validation. *Lighting Research and Technology*; 0: 1477153518824789. DOI: 10.1177/1477153518824789.
46. Tang X, Teunissen C. The appreciation of LED-based white light sources by Dutch and Chinese people in three application areas. *Lighting Research and Technology* 2018; Online Before Print. DOI: 10.1177/1477153517754130.
47. Khanh TQ, Bodrogi P, Vinh QT, Stojanovic D. Colour preference, naturalness, vividness and colour quality metrics, Part 1: Experiments in a room. *Lighting Research and Technology* 2017; 49: 697-713. DOI: 10.1177/1477153516643359.
48. Khanh TQ, Bodrogi P, Vinh QT, Stojanovic D. Colour preference, naturalness, vividness and colour quality metrics, Part 2: Experiments in a viewing booth and analysis of the combined dataset. *Lighting Research and Technology* 2017; 49: 714-726. DOI: 10.1177/1477153516643570.
49. Khanh TQ, Bodrogi P. Colour preference, naturalness, vividness and colour quality metrics, Part 3: Experiments with makeup products and analysis of the complete warm white dataset. *Lighting Research and Technology* 2018; 50: 218-236. DOI: 10.1177/1477153516669558.
50. Khanh TQ, Bodrogi P, Vinh QT, Guo X, Anh TT. Colour preference, naturalness, vividness and colour quality metrics, Part 4: Experiments with still life arrangements at different correlated colour temperatures. *Lighting Research and Technology* 2018; 50: 862-879. DOI: 10.1177/1477153517700705.
51. Jerome CW. The Flattery Index. *Journal of the Illuminating Engineering Society* 1973; 2: 351-354. DOI: 10.1080/00994480.1973.10747727.
52. Jerome CW. Flattery vs Colour Rendition. *Journal of the Illuminating Engineering Society* 1972; 1: 208-211. DOI: 10.1080/00994480.1972.10732210.
53. Judd DB. Flattery Index for Artificial Illuminants. *IES Transactions* 1967: 593-598.
54. Thornton WA. A Validation of the Colour-Preference Index. *Journal of the Illuminating Engineering Society* 1974; 4: 48-52. DOI: 10.1080/00994480.1974.10732288.
55. Veitch JA, Newsham GR, Boyce PR, Jones CC. Lighting appraisal, well-being and performance in open-plan offices: A linked mechanisms approach. *Lighting Research and Technology* 2008; 40: 133-151. DOI: 10.1177/1477153507086279.
56. Veitch JA, Stokkermans MGM, Newsham GR. Linking Lighting Appraisals to Work Behaviors. *Environment and Behavior* 2013; 45: 198-214. DOI: 10.1177/0013916511420560.
57. Smet K, Ryckaert WR, Pointer MR, Deconinck G, Hanselaer P. Optimization of colour quality of LED lighting with reference to memory colours. *Lighting Research and Technology* 2012; 44: 7-15. DOI: 10.1177/1477153511432250.
58. Liu Q, Huang Z, Xiao K, Pointer MR, Westland S, Luo MR. Gamut Volume Index: a colour preference metric based on meta-analysis and optimized colour samples. *Optics Express* 2017; 25: 16378-16391. 2017/08/10. DOI: 10.1364/OE.25.016378.
59. Freyssinier JP, Rea M. The Class A Colour Designation for Light Sources. In: *DOE Solid-State Lighting RandD Workshop* Long Beach, CA, January 29-31, 2013 2013.
60. Hashimoto K, Yano T, Shimizu M, Nayatani Y. New method for specifying colour-rendering properties of light sources based on feeling of contrast. *Colour Research and Application* 2007; 32: 361-371. DOI: 10.1002/col.20338.
61. Hashimoto K, Nayatani Y. Visual clarity and feeling of contrast. *Colour Research and Application* 1994; 19: 171-185.
62. Global Lighting Association. Application of CIE 13.3-1995 with Associated CRI-based Colour Rendition. http://www.globallightingassociation.org/library2018. Accessed March 2019.

63. Royer MP, Houser KW, David A. Chroma Shift and Gamut Shape: Going Beyond Average Colour Fidelity and Gamut Area. *Leukos* 2018; 14: 149-165. DOI: 10.1080/15502724.2017.1372203.
64. Fairchild MD, Reniff L. Time course of chromatic adaptation for colour-appearance judgments. *J Opt Soc Am A Opt Image Sci Vis* 1995; 12: 824-833. DOI: 10.1364/josaa.12.000824.
65. Rinner O, Gegenfurtner KR. Time course of chromatic adaptation for colour appearance and discrimination. *Vision Research* 2000; 40: 1813-1826.
66. Shevell SK. The time course of chromatic adaptation. *Colour Research and Application* 2001; 26: S170-S173.
67. Rea MS, Freyssinier JP. White lighting. *Colour Research and Application* 2013; 38: 82-92. DOI: 10.1002/col.20738.
68. Royer MP. What Is the Reference? An Examination of Alternatives to the Reference Sources Used in IES TM-30-15. *Leukos* 2016; 13: 71-89. DOI: 10.1080/15502724.2016.1255146.
69. Royer MP. Comparing Measures of Gamut Area. *Leukos* 2018; 15: 29-53. DOI: 10.1080/15502724.2018.1500485.
70. Smet K, Ryckaert WR, Pointer MR, Deconinck G, Hanselaer P. Correlation between colour quality metric predictions and visual appreciation of light sources. *Optics Express* 2011; 19: 8151-8166. DOI: 10.1364/OE.19.008151.
71. Royer MP. Comparing Measures of Average Colour Fidelity. *Leukos* 2017; 14: 69-85. DOI: 10.1080/15502724.2017.1389283.
72. David A. Colour Fidelity of Light Sources Evaluated over Large Sets of Reflectance Samples. *Leukos* 2013; 10: 59-75. DOI: 10.1080/15502724.2013.844654.
73. Smet KAG, David A, Whitehead L. Why Colour Space Uniformity and Sample Set Spectral Uniformity Are Essential for Colour Rendering Measures. *Leukos* 2015; 12: 39-50. DOI: 10.1080/15502724.2015.1091356.
74. David A, Fini PT, Houser KW, Ohno Y, Royer MP, Smet KA, Wei M, Whitehead L. Development of the IES method for evaluating the colour rendition of light sources. *Optics Express* 2015; 23: 15888-15906. 2015/07/21. DOI: 10.1364/OE.23.015888.
75. Royer M. Spectral power distributions. In: figshare. 2019. DOI: 10.6084/m9.figshare.7704566.v1
76. Royer M. Colour Rendition Characteristics. In: figshare. 2019. DOI: 10.6084/m9.figshare.7704521.v1