Analysis of Color Rendition Specification Criteria

Michael P. Royer

Pacific Northwest National Laboratory, 620 SW 5th Ave Suite 810, Portland, OR, USA 97202

Royer, M. "Analysis of color rendition specification criteria." *Light-Emitting Devices, Materials, and Applications*. Vol. 10940. International Society for Optics and Photonics, 2019.

Analysis of Color Rendition Specification Criteria

Michael P. Royer*[a]

[a]Pacific Northwest National Laboratory, 620 SW 5th Ave Suite 810, Portland, OR, USA 97202

ABSTRACT

Methods for evaluating light source color rendition have recently undergone substantial changes. This article explores the ability of color rendition specification criteria to capture preferences for lighting color quality. For a compilation of five recent psychophysical studies on perceptions of colors, recently proposed specification criteria using ANSI/IES TM-30-18 substantially outperformed all currently used specification criteria in identifying preferred lighting conditions. To understand the consequences of changing color rendition specification criteria, the performance of a set of 484 commercially-available SPDs was evaluated.

Keywords: Color rendition, specification criteria, color preference, spectral efficiency, TM-30, color fidelity, CRI

1. INTRODUCTION

Measures (objective quantities) and metrics (subjective quantities) of color rendition are calculation methods that return a quantitative value describing how a light source will influence the appearance of colored objects. Dozens of measures and metrics of color rendition have been proposed over the past 70 years as the variety of light sources—specifically their spectral power distributions (SPDs)—has grown. These calculation tools vary in what they intend to quantify, as well as in the underlying framework that is used to make assumptions about object colors, viewing conditions, and psychological processes.

While many measures and metrics of color rendition have been proposed by researchers, only a small number are recommended for use by an authoritative lighting organization. In 1965, the Commission Internationale de L'Eclairage first adopted the *Method of Measuring and Specifying Colour Rendering Properties of Light Sources*; this document was substantially revised in 1974 and reaffirmed in 1995 [1]. It specifies the calculation of the General Color Rendering Index R_a (colloquially, CRI), as well as 14 Special Color Rendering Indices, R_i). Although widely used in practice and referenced in other specification and standard documents, this method has long been criticized for both the inaccuracies of its outdated color science and the limited extent of information that it provides, with a lone focus on color fidelity (the average similarity of a test light source to a reference illuminant) [2-12].

In 2015, the Illuminating Engineering Society (IES) published TM-30-15, *IES Method for Evaluating Light Source Color Rendition* [13]. The TM-30 method uses a common calculation framework (color samples, color space, reference illuminant scheme) to determine a wide range of measures that quantify different aspects of color rendition, including average color fidelity [5] and gamut area [14], as well as 16 values each for local (*i.e.,* hue specific) chroma shift [15], local hue shift, and local color fidelity. Rather than weighting various color shifts or attempting to derive a single value that quantifies preferred color rendition—or any other subjective quality—TM-30 provides a library of tools that can be used to specify appropriate color rendition across many different lighting applications. In essence, the TM-30 measures function as an alphabet that can be used to create many different words that describe desired color rendition qualities.

Subsequently, the CIE published 224:2017, *Colour Fidelity Index for Accurate Scientific Use* [8], adopting the calculation framework of IES TM-30-15 with minor changes, but only including one of the derived measures, the Fidelity Index (R_f). Most recently, the IES published TM-30-18 [7], incorporating the CIE's changes to R_f and providing additional supporting material, such as recommended formats for specification sheets. TM-30-18 is approved by the American National Standards Institute (ANSI), making it the only method of evaluating light source color rendition with such a designation.

ANSI/IES TM-30-18 defines numerous color rendition measures, but it intentionally avoids describing how they should be used in a specification system, which is the role of an IES Recommended Practice. This article demonstrates the benefits of a specification system based on ANSI/IES TM-30-18 and explores how conversion to such a system will influence light source development.

1.1 Color rendition specification criteria

Color rendition specification criteria define the accepted range of values for a given color rendition measure(s) or metric(s)—hereafter always referred to as measures. Specification criteria are set by a variety of people and organizations, including lighting specifiers, energy efficiency programs (*e.g.*, ENERGY STAR™, the DesignLights Consortium), government (*e.g.*, California Title 20), and professional societies (*e.g.*, IES) for a variety of purposes. Table 1 provides a summary of existing color rendition specification criteria, both in active use and recently proposed by researchers, as well as recommendations from the IES.

The same pairing of measures and criteria exists in other aspects of lighting. For example, illuminance is a defined measure and minimum average illuminance criteria are set for a variety of applications by different organizations. Correlated color temperature (CCT) and distance from the Planckian locus (D_{uv}) are defined measures used by ANSI/NEMA C78.377-2017 [16] to establish nominal classifications. Beyond lighting, age and height are common quantities used to establish criteria for admission, and defined measures like miles per hour are used by agencies to establish speed limits. To best meet a desired outcome, both the measure(s) and criteria need to be appropriate.

In many cases, there are different criteria that use the same measure(s). This arises because there are several factors that influence the criteria that are set, allowing the person or organization to arrive at different solutions. Some of these include:

- Intent: Although it has rarely been stated in the past, a key element of setting color rendition criteria is deciding what characteristic of color appearance is to be addressed. This could include subjective qualities such as acceptability, naturalness, vividness, or preference, as well as more objective qualities like color fidelity or metameric uncertainty.
- Minimum qualification versus highest quality: Some color rendition criteria focus on minimum standards as a counter for an opposing performance aspect, such as luminous efficacy, whereas others are intended to permit only the light sources deemed to have the most appropriate color rendition performance for the intended purpose.
- Simplicity versus complexity: Using a single measure can be easier for users to understand and remember, but a more complex multi-measure approach can be more informative and more effective in some cases.
- Flexibility versus prescription: More lenient criteria may allow a greater variety of products, enabling a diversity of capabilities but perhaps requiring more individual discretion. In contrast, more stringent criteria can be more predictable, but may limit innovation and increase cost.

The latter three tradeoffs relate to a balance between false positives (products that meet the criteria but that are not appropriate) and false negatives (products that do not meet the criteria but that would be appropriate). In certain situations, one may be more acceptable than the other.

The history of institutional color rendition criteria is not well-documented, but perhaps began with several utility energy efficiency programs in the 1990s [17], with $R_a \geq 80$ being a common criterion. These programs were predecessors to ENERGY STAR, which adopted $R_a \geq 80$ in 2001 as part of CFLs version 2.0 [18]. Interestingly, this was more than 25 years after the initial publication of CIE 13. With fewer lamp types and thus fewer choices in color rendition, establishing a minimum CIE R_a threshold was perhaps not as pressing, and individual lighting specifiers could more easily distinguish between products. The $R_a \geq 80$ criterion was not based on experimental evidence, but rather on a combination of technology capabilities, manufacturing tolerances, and common practices, according to recollections of those involved or knowledgeable. Eventually, $R_a \geq 80$ became a de facto standard of lighting practice and was instrumental in the development of LED technology. It became known as a delineator between acceptable and unacceptable (or liked and disliked) lighting, even though no credible experiments ever supported that idea—and many contradicted it.

As Table 1 attests, there has been some movement away from $R_a \geq 80$, particularly in the past few years. With only one standardized method for evaluating color rendition prior to 2015, the options to improve the specified level of color quality amounted to increasing the R_a requirement (*e.g.*, to $R_a \geq 90$) or supplementing it with a requirement for R_9—a recognition of the specific importance of reds. Despite no changes in the measures being used, implementing more stringent color rendition criteria (*e.g.*, California Title 24 JA8 [19]) drew considerable pushback from manufacturers of lighting equipment and others. The concerns varied from the new values being too restrictive, to the energy efficiency consequences, to the new criteria not being inherently more correlated with lighting preferences than existing criteria.

Table 1. Color rendition specifications and recommended practices.

Type	Name	Criteria
Voluntary (Energy Efficiency Rebate)	DesignLights Consortium Qualified Products List, Technical Requirements V4.4, Indoor Luminaires	$R_a \geq 80$
Voluntary (Energy Efficiency Rebate)	ENERGY STAR Certified Light Bulbs V2.0	$R_a \geq 80$, $R_9 \geq 0$
Voluntary (Building Certification)	WELL Building Standard V1	$R_a \geq 80$, $R_9 \geq 50$
Voluntary (Building Certification)	WELL Building Standard V2	
	All Spaces Except Circulation	$R_a \geq 90$ OR $R_a \geq 80$, $R_9 \geq 50$, OR $R_f \geq 78$, $R_g \geq 98$, $-1\% \leq R_{cs,h1} \leq 15\%$
	Circulation Areas	$R_a \geq 80$ OR $R_f \geq 78$, $R_g \geq 95$, $-7\% \leq \mathrm{Rcs,h1} \leq 15\%$
Mandatory (for sale in state)	California Appliance Efficiency Regulations (Title 20)	$R_a \geq 82$
Mandatory (residential new constr)	California Building Efficiency Standards (Title 24 JA8)	$R_a \geq 90$, $R_9 \geq 50$
Mandatory (military medical facilities)	U.S. DOD UFC 4-510-01: Design Military Medical Facilities	$R_f \geq 80$*, $R_g \geq 97$, $R_g \leq 110$, $-9\% \leq R_{cs,h1} \leq 9\%$ (with exceptions)
Proposal	Class A [20-22]	$R_a \geq 80$, $80 \leq \mathrm{GAI} \leq 100$
Proposal	Royer et al. *Color Preference* [10, 23, 24]	
	Tier A (Best)	$R_f \geq 78$, $R_g \geq 95$, $-1\% \leq R_{cs,h1} \leq 15\%$
	Tier B (Good)	$R_f \geq 74$, $R_g \geq 92$, $-7\% \leq R_{cs,h1} \leq 19\%$
	Tier C (Acceptable)	$R_f \geq 70$, $R_g \geq 89$, $-12\% \leq R_{cs,h1} \leq 23\%$
Recommendation	IES Lighting Handbook, 10th Ed.	
	General Interior	$R_a \geq 80$
	Color Appraisal	$R_a \geq 85$
	Color Matching & Reproduction	$R_a \geq 90$
American National Standard Recommended Practice	ANSI/IES RP-1-12: Office Lighting	
	General	$R_a \geq 80$
	Color Matching/Discrimination	$R_a \geq 90$
American National Standard Recommended Practice	ANSI/IES RP-3-13: Educational Facilities	
	General	$R_a \geq 80$
	Color Discrimination	$R_a \geq 90$
Recommended Practice	IES RP-4-13: Lbiraries	
	Meeting Rooms	$R_a \geq 85$
Recommended Practice	IES RP-7-01: Industrial	
	Important	$R_a \geq 70$
	Critical	$R_a \geq 85$
American National Standard Rec. Practice	ANSI/IES RP-28-16: for Seniors and the Low Vision Population	$R_a \geq 80$

*IES TM-30-15 value. Equivalent to $R_f = 82$ for IES TM-30-18. Other values are equivalent for 2018 and 2015 versions.

1.2 Development and implementation of new color rendition criteria

Now that new measures of color rendition have been standardized and are being used by lighting manufacturers, specifiers, and photometric laboratories, there is an important opportunity to revisit color rendition specification criteria. Simply applying old values (*e.g.*, 80) to a new measure (*e.g.*, R_f) would be inappropriate. The knowledge required to identify and achieve consensus on specification criteria for new color rendition measures can be gained in several ways. The most straightforward approach is benchmarking, which is used to translate criteria from one measure(s) to another using a set of qualified products. That is, new qualifying ranges for a new measure(s) can be set based on the properties of products that qualified based on previous criteria. While rapid, this method relies heavily on the assumption that the previous criteria were effective at qualifying appropriate light sources and disqualifying those that were not. It relies on existing products, which may not be a good representation of future products.

A second approach to establishing new color rendition specification criteria is experimentation, where responses from human participants can be gathered for questions about subjective aspects of color appearance. Collecting objective data on physical responses is also possible, but to date has not been done in the context of color rendition. Experimental apparatuses can range from small scale viewing booths to large scale field experiments, with tradeoffs for each. Experiments present the opportunity to investigate novel light sources but rely on the assumption that the effects of color rendition in a controlled environment, usually with short-duration exposures, are applicable to other situations.

An alternative to these two approaches is relying on experience. This is the most subjective approach to establishing new criteria but can fill an important void, providing information about a variety of applications under real-world viewing conditions when other methods are not applicable or have not yet been pursued.

There are undoubtedly technical and emotional concerns associated with widespread conversion to new methods and criteria for evaluating light source color rendition. Such an undertaking requires substantial planning and communication, as both user expectations and product literature must be revised. With most updates to specification criteria, there are likely to be products that met previous criteria but do not meet updated criteria. Engineering developments may also be required. All these issues can be addressed with forethought, and there are many alternatives to a hard changeover. For example, the change can be preceded by required reporting of new measures, with a planned phase-in period for qualification criteria. The phase-in can potentially include a period with dual paths to qualification, through either the new method or the old.

2. SPECIFICATION CRITERIA PERFORMANCE COMPARISON

This analysis focuses on color preference, exploring the relative performance of currently-used and recently proposed color rendition specification criteria by applying them to existing experimental data. Although color preference may not be the most important consideration in all lighting applications, environmental satisfaction has been linked to wellbeing and performance [25, 26]. Rather than a combined meta-analysis, this investigation considers each set of experimental data individually to preserve differences that arise from variations in experimental conditions.

2.1 SPDs

To evaluate the performance of the specification criteria, it was first necessary to identify experimental data that could generate reliable results. More than 50 articles describing psychophysical experiments on perception of subjective aspects of color rendition were reviewed [10, 20, 21, 23, 24, 27-80]. The following inclusion criteria were applied to select studies most appropriate for analyzing further for this article:

- Controlled chromaticity and/or allowed for chromatic adaptation.
- Used typical illuminance levels for an architectural interior (arbitrarily 200-650 lux)
- Presented a wide variety of SPDs (going beyond commercially available products), including variations in gamut shape.
- Appropriately counterbalanced order effects.
- Presented objects beyond color swatches.

- Used interval-based response forms to gather subjective evaluations of color preference.
- Made SPDs available, either within the article or through personal communication.

Approximately 90% of the studies reviewed were excluded because they failed to present a wide variety of SPDs; in most cases, eight or fewer were used, which is generally insufficient to vary color rendition across all its dimensions. Many also failed to properly account for chromatic adaptation, frequently relying on side-by-side booths or sequential presentation of light sources having different chromaticities—note that matching CCTs is insufficient to control for chromatic adaptation. Some also used only an X-Rite ColorChecker card, which provides no context for the colors and can exhibit greater-than-average shifts due to the highly saturated colors [81], or as few as one object (*e.g.*, skin). Many of these studies were in accordance with accepted practices of their time, but lighting technology, color rendition measures, and experimental protocols have all changed in recent years. It is still possible to test specification criteria against these past results, but the individual contexts are more important for interpretation and they could not be provided in this article. For example, average measures tend to perform better when gamut shape is held constant than when it is varied [5].

Five studies were chosen for analyzing the performance of color rendition criteria, all of which were conducted in 2015 or later. Three of the five were a series of related experiments conducted by Royer *et al.* at Pacific Northwest National Laboratory [10, 23, 24]; they were analyzed as one group because they shared the same methodology. In total for these three studies, 166 unique SPDs were presented to participants in a full-scale room filled with a wide variety of objects. The fourth included study was conducted by Zhang *et al.* at Zhejiang University [71]. It featured 164 SPDs divided among four CCT groups. Real objects and color samples were presented in a viewing booth. The fifth included study was conducted by Esposito and Houser at Penn State University [72]. It featured 24 SPDs at 3500 K on the Planckian locus, using a viewing booth that held 12 real objects. Additional details for all five studies are shown in Figure 1.

All five of these studies included subjective evaluations of naturalness/normalness, vividness/saturation, and preference, with Likert scales featuring six (Esposito and Houser), seven (Zhang *et al.*), or eight (Royer *et al.*) values. Two of the three studies from Royer *et al.* also asked a binary question about acceptability. All studies found that the TM-30 Fidelity Index (R_f), Gamut Index (R_g), and Local Chroma Shift for nominally red objects ($R_{cs,h1}$ or $R_{cs,h16}$) were important factors for predicting subjective evaluations, outperforming all other measures of color rendition that were fit to the data. Only the studies from Royer *et al.* explicitly suggested specification criteria (Table 1), although Esposito and Houser observed a preferred range ($R_g \geq 100$, $R_{cs,h16} \geq 0\%$). All three sets of experiments included regression modelling of the subjective evaluations.

2.2 Rank order comparisons with experimental SPDs

Each experimental SPD was tested against each set of criteria, generating a pass or fail result. Within each set of experimental data, the SPDs were arranged in rank order based on the participants' mean preference ratings, allowing for a visual analysis of the effectiveness of each set of specification criteria by color coding the pass (green) or fail (red) outcomes. An ideal situation would have all the passed SPDs in one block and all the failed SPDs in a second block, as illustrated in Figure 2. The point of transition between the blocks can vary based on how stringent the criteria are or on the types of SPDs shown in the specific experiment. For this reason, each set of SPDs was considered individually The key indicator of performance is the homogeneity of the blocks.

Figures 3-5 demonstrate the performance of a variety of color rendition criteria for SPDs from the three datasets. Although this is not a numerical or statistical analysis, the results are clear: the recently proposed Color Preference criteria, based on ANSI/IES TM-30-18, are superior for all three datasets, with fewer false positives and false negatives. An expected follow-up analysis will explore numerical optimization of these criteria and additional alternatives based on proposed measures of color rendition.

The Color Preference criteria were developed and refined over a series of three experiments, which form the complete dataset shown in Figure 5. The initial experiment [10] helped to identify the most significant factors influencing these perceptions: $R_{cs,h16}$ or $R_{cs,h1}$ along with R_f and R_g for preference. Initial criteria for the most preferred products were suggested (not shown in Table 1), focusing only on the top tier (A). The second experiment [24] refined the criteria, adding a second tier of performance (B). These two sets of criteria were included as a path

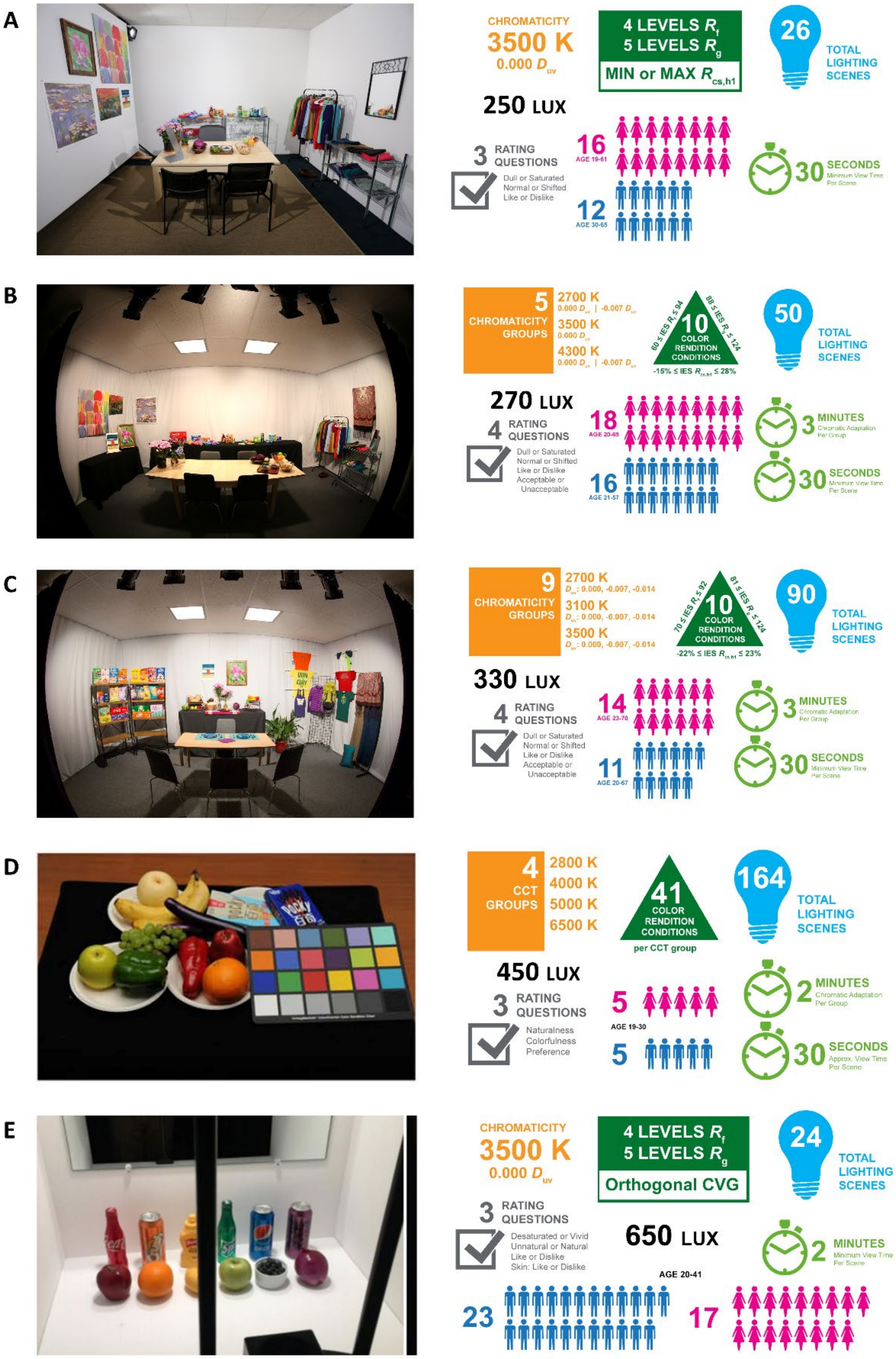

Figure 1. Summary of experiments used for comparing color rendition specification criteria. A [10], B [24], C [23], D [71], E [72].

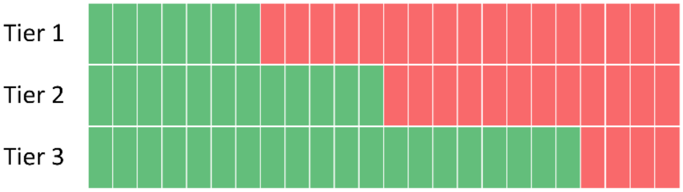

Figure 2. Ideal color rendition specification criteria performance. Each bar represents one SPD. The SPDs are arranged in rank order from most preferred (left) to least preferred (right).

to qualification for the Version 2 pilot of the Well Building Standard [82] (Table 1). These tiers relied on observed groupings of products based on preference ratings, which are closely correlated ($r^2 = 0.85$) with percent acceptability (Figure 6). Percent acceptability offers a more understandable value compared to the Likert scale ratings. A third tier (C) was later developed by combining a benchmarking exercise, experimental data, and judgement; the criterion for red rendition allows most products with $R_a \geq 0$ and $R_9 \geq 0$ to qualify, but it offers more flexibility (*i.e.,* a lower threshold) for average color fidelity (R_f). This lowest tier generally has lower acceptability ratings. Finally, the third experiment [23] was used to test and refine the criteria, resulting in those shown in Table 1 and used to color the datapoints in Figure 5. Final refinement also included adjustments to provide even increments in R_f (4 points), R_g (3 points), and $R_{cs,h1}$ (4% upper limit). The tiered system allows for the stringency of the criteria to be adjusted; it is fully nested, so that qualifying for a higher tier automatically indicates qualification for lower tiers.

It must be acknowledged that the Color Preference criteria have the advantage of being developed to fit one of the three datasets evaluated for this article. At the same time, they are the newest and least vetted with experience. They were also adjusted for practicality instead of being perfectly optimized for the experimental results. Importantly, the performance of this criteria set for the other two datasets (Figure 3, Figure 5) remains the strongest of the options evaluated. This is an important validation relying on data that was independent of that used in the development of the criteria.

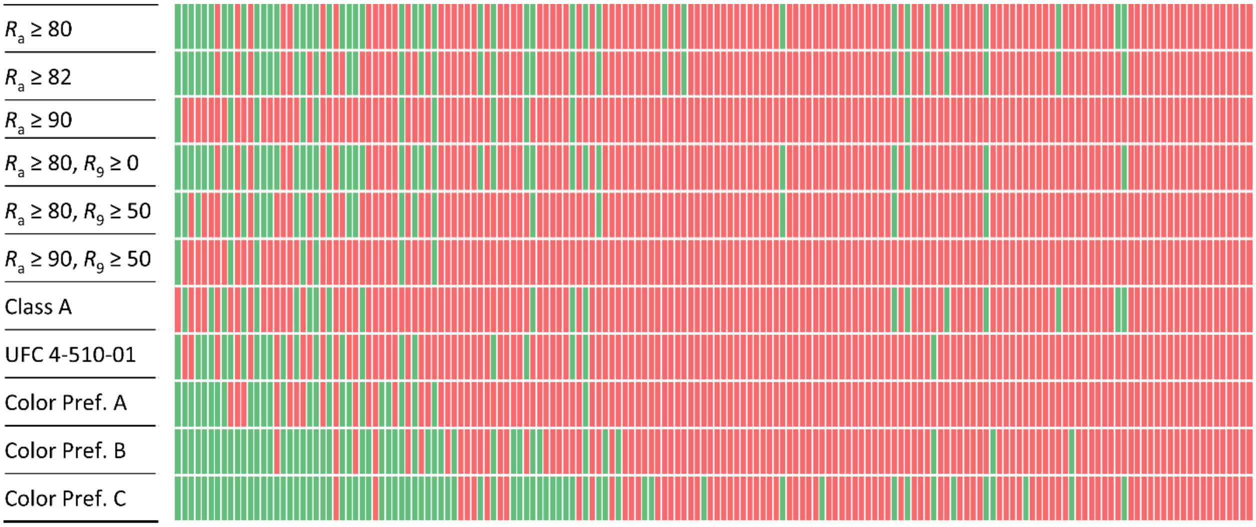

Figure 3. Performance of color rendition specification criteria for the Zhang et al. dataset [71]. Each bar represents one SPD. The SPDs are arranged in rank order from most preferred (left) to least preferred (right).

Figure 4. Performance of color rendition specification criteria for the combined Royer et al. dataset [10, 23, 24]. Each bar represents one SPD. The SPDs are arranged in rank order from most preferred (left) to least preferred (right).

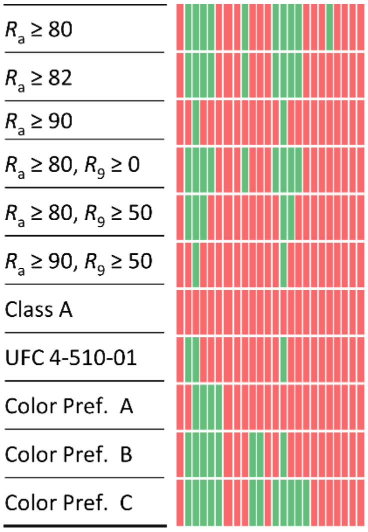

Figure 5. Performance of color rendition specification criteria for the Esposito and Houser dataset [5]. Each bar represents one SPD. The SPDs are arranged in rank order from most preferred (left) to least preferred (right).

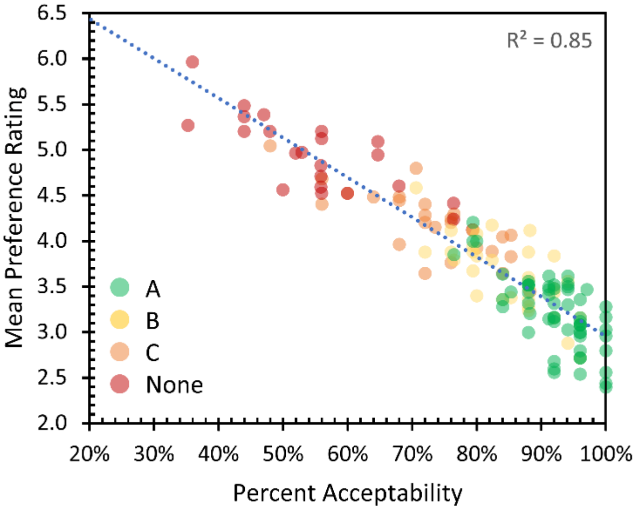

Figure 6. Correlation between mean preference rating and percent acceptability for the 140 SPDs shown in the second and third studies at PNNL.

2.3 Specification criteria applied to large SPD sets

It is important to consider how conversion to new color rendition specification criteria will affect existing products, a majority of which was developed under different criteria. Figure 7 displays $R_{cs,h1}$ versus R_f for three types of SPDs: approximately 165,000 theoretical SPDs comprised of randomly-generated Gaussian primaries mixed to create nominally white light [83], the 354 experimental SPDs from the three datasets used in this analysis, and 484 commercially-available LED SPDs that came from the LED Lighting Facts database (those that reported SPDs as of December 2018), U.S. DOE CALiPER Testing, U.S. DOE Next Generation Luminaire Systems Competition submissions, the IES TM-30-18 Example SPD Library, or personal communication with LED manufacturers.

The theoretical SPDs in Figure 7 are colored based on their performance using criteria of $R_a \geq 80$ with $R_9 \geq 0$ or $R_a \geq 90$ with $R_9 \geq 50$; this shows regions within the TM-30 R_f-$R_{cs,h1}$ space where such products could fall. SPDs meeting the $R_a \geq 80$ or $R_a \geq 90$ criteria could have much lower R_f values (41 or 65, respectively), due to the technical flaws of R_a [5]. The chart also demonstrates the bias of R_a against SPDs that increase red chroma: the clouds of data based on R_a and R_9 criteria are not centered vertically on the $R_{cs,h1} = 0\%$ line. When used alone, R_f (nor R_a) is not capable of ensuring the most preferred color rendition of any given hue, even with a criterion of 95.

Most commercially-available SPDs (black points) fall toward the lower boundary of the $R_a \geq 80$, $R_9 \geq 0$ region (orange points). This is because the gamut shape produced by these SPDs is the most spectrally efficient way to meet the $R_a \geq 80$ specification [83], which has been most widely used. Compared to listings from LED Lighting Facts as of January 14, 2019 [84], the 484 commercially-available SPDs overrepresent products with $R_a \geq 85$ (Figure 8). Of the 32 commercially-available SPDs with $R_{cs,h1} \geq 0$ % in Figure 7, 26 were variants of four white-tuning products that have low efficacy (≤ 53 lm/W) [85]. The commercially-available product with the highest $R_{cs,h1}$ value (7%) was specifically designed based on the research used to create the Color Preference specifications.

Nearly half of the products listed by LED Lighting Facts, and approximately 80% of the products with $R_a \geq 80$ have an R_a value between 80 and 85—represented in Figure 7 by the large cluster of commercially-available products (black points). The ubiquitous nature of this design (featuring a blue InGaN LED and a yellow-green YAG phosphor) is an important consideration when establishing color rendition criteria for immediate use. The Color Preference C criteria were devised to qualify a vast majority of these products, even though experimental SPDs with equivalent color rendition characteristics were considered acceptable to fewer than 70% of participants [23]. Implementing color rendition criteria that improves user satisfaction, preference, and acceptability, will require moving away from this type of product, perhaps by adding a red phosphor, red LED, yellow filters, or converting to a color-mixed approach.

Figure 7 only shows two of the three measures included in the Color Preference criteria, excluding R_g. However, products failing to meet the R_g criterion while meeting the other criteria is a small minority of the possibilities. Figure 7 shows the concentric nature of the Color Preference criteria, with the center shifted to positive values of $R_{cs,h1}$.

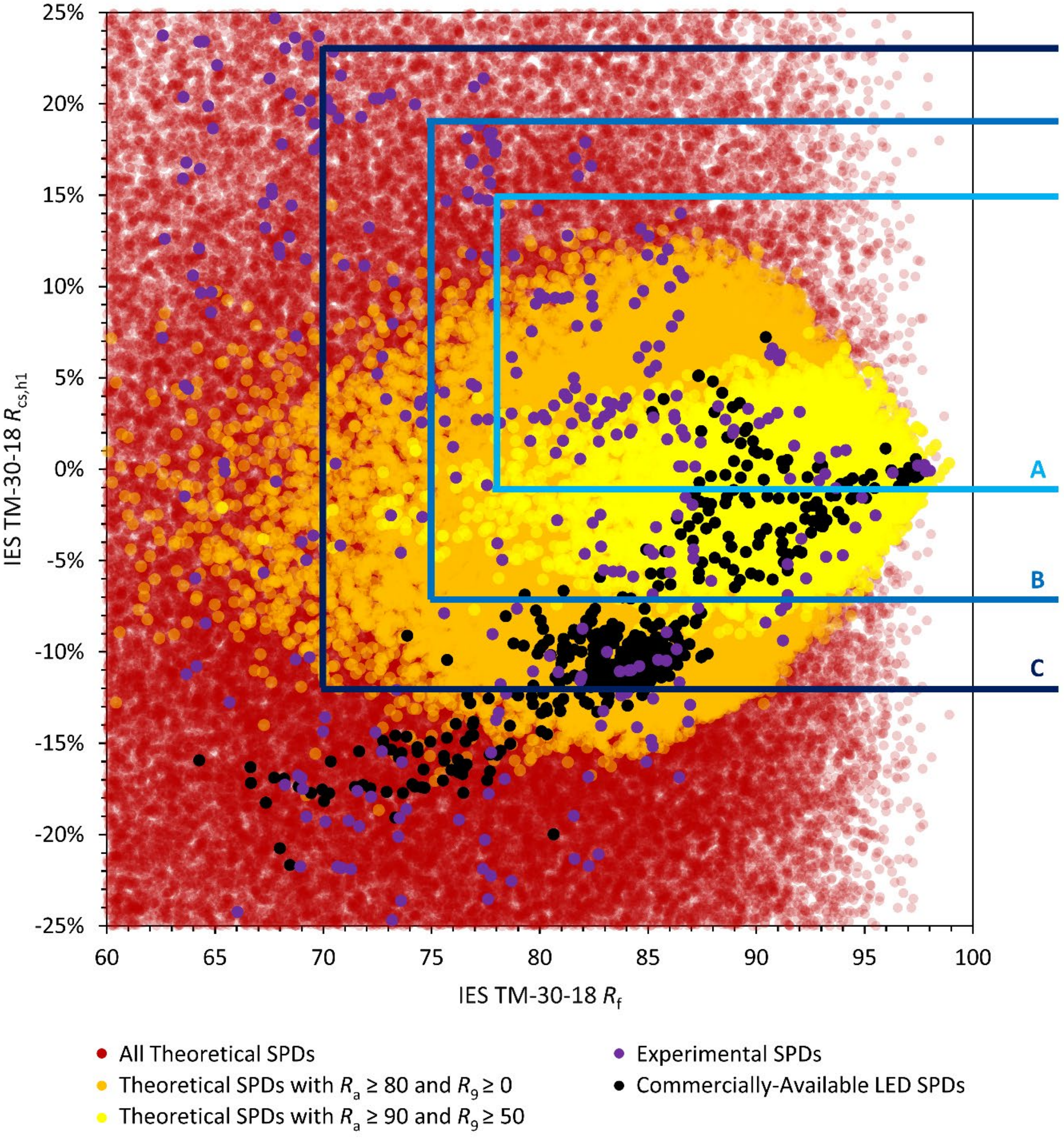

Figure 7. Color rendition characteristics of theoretical, experimental, and commercially-available LEDs relative to color rendition specification criteria. Note that only two of the three measures in the Color Preference criteria are represented; R_g. is excluded.

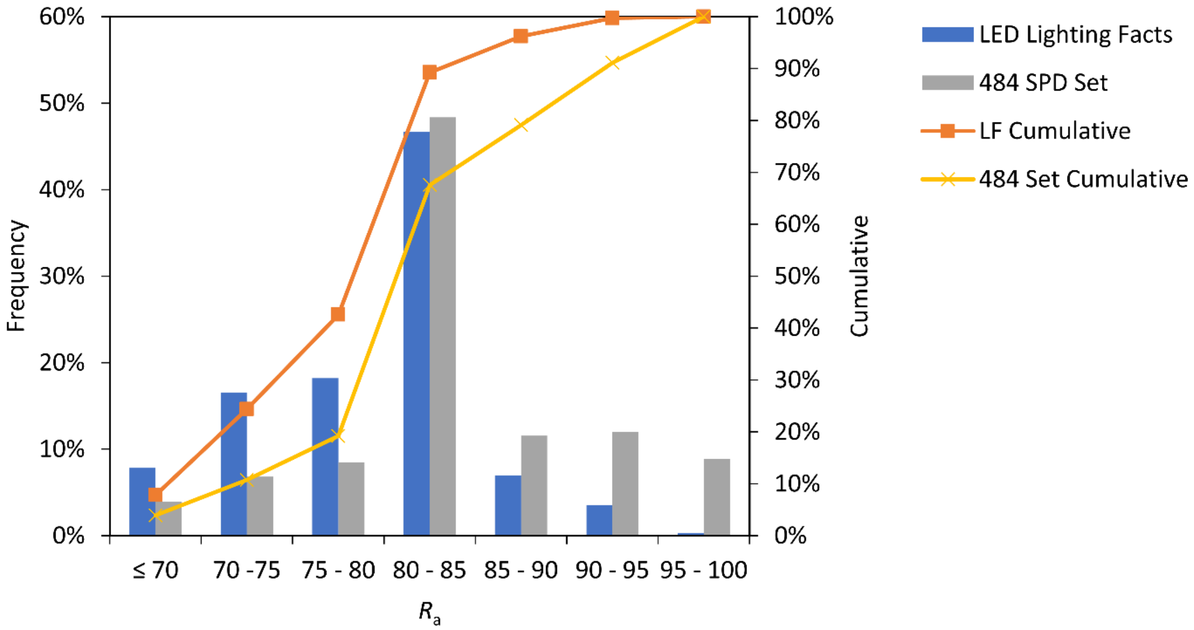

Figure 8. R_a values of the set of 484 commercially-available SPDs versus the listings from LED Lighting Facts [4].

3. DISCUSSION

Existing criteria utilizing R_a, regardless of how stringent or if combined with R_9, are unable to consistently qualify products with preferred color rendition. Some disliked products qualify, while some preferred products do not qualify. This is due to inaccuracies and limitations of R_a and R_9 resulting from outdated color science and the types of characteristics they quantify. In contrast, there is now a tiered system of criteria based on ANSI-approved measures of color rendition that offers superior performance for identifying light sources that were found to be preferred in experiments. Tiers allow for adjusting the stringency of the criteria; more restrictive criteria can be used if color rendition takes precedence over other lighting considerations, or more lenient criteria can be used to allow for greater balance with other factors. Given that the goal of many specifications that include limits for color rendition is to promote high-quality products that will encourage use (and subsequently save energy), it would be logical to update existing criteria as soon as it is feasible.

There has been some debate about whether the variety of factors influencing subjective evaluations of color rendition would preclude use of a criteria that is independent of demographics and lighting application. Besides the difference in SPDs, some of the small differences in performance observable in Figures 3-5 can be attributed to different experimental conditions, including the objects being viewed, their context (or lack thereof), participant demographics, and procedures. The question of application can be addressed to some degree by creating specification criteria for different design intents (*e.g.*, preference, vividness, high fidelity) without resorting to identifying specific applications. This article focuses only on color preference, but similar approaches can be taken for other design intents. The strong performance of the new criteria across three independent investigations with varied observer populations is indicative of their potential for broad use, as has been found for other color rendition measures [64]. Note that the difference in the total percentage of SPDs that met the various tiers of criteria is not an indication of the performance of the criteria, but rather reflects the range and types of SPDs shown in each experiment.

While the benchmark for criteria performance in this analysis was a defined division of blocks identified as pass and fail, it is unrealistic to expect such performance for experimental datapoints arranged in rank order. While the ratings themselves are continuous when ordered, the experimental data is a sample of a population and has an associated

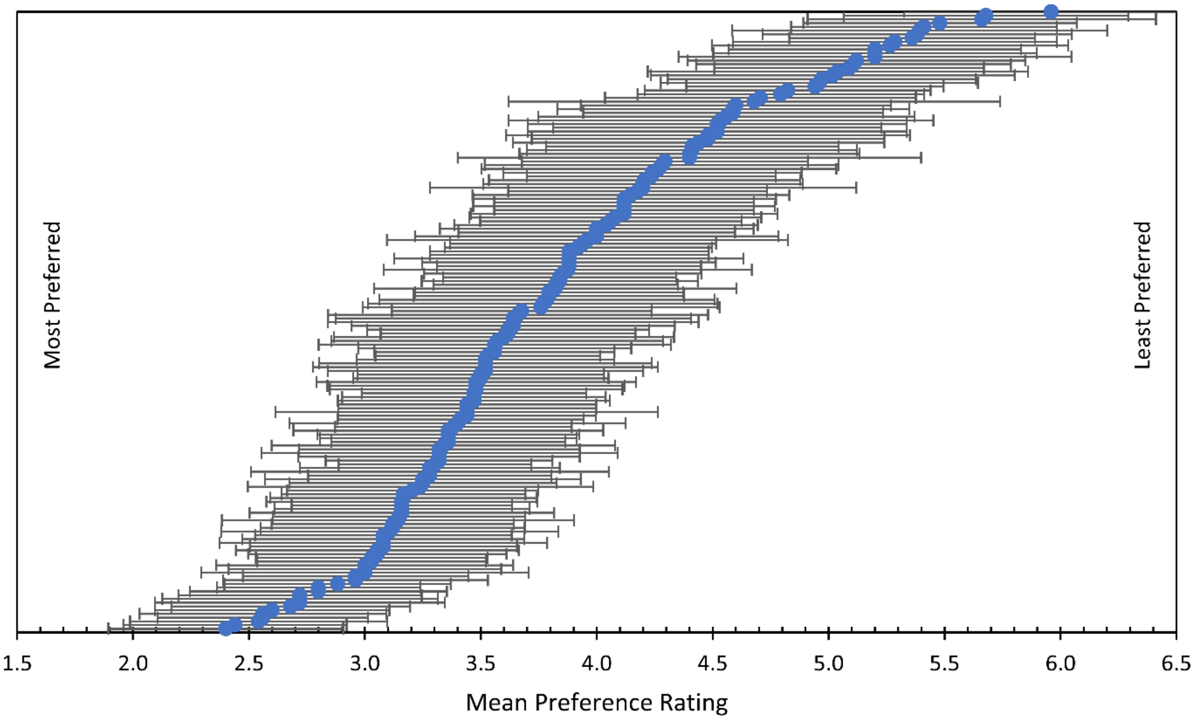

Figure 9. Mean preference rating and 95% confidence interval for each of the 166 SPDs comprising the PNNL dataset.

variance. Figure 9 shows means and 95% confidence intervals for the 166 SPDs in the Royer *et al.* dataset. Some of the visually-evaluated false positive and false negative findings can be attributed to displacement in the rank order due to sampling variance. In other cases, some of the false negatives are produced by SPDs that narrowly miss meeting the criteria. It is possible that any newly adopted criteria could be revised over time as more evidence emerges, just as thresholds for other values, such as efficacy, are adjusted on a regular basis. It is also possible that additional measures, such as the Metameric Uncertainty Index (R_t) [86] could supplement the existing criteria to improve criteria performance.

Users of any color rendition specification criteria should be aware of their limitations, purpose, audience, and uncertainties. The Color Preference criteria are intended for use with polychromatic environments having illuminance levels between 200 and 650 lux, without significant variation in chromaticity for the lighting within the space. By focusing on criteria, rather than an optimization model that produces a single index, it is expected that users can rely on additional discretion when applying the criteria to specific situations.

3.1 Characteristics of SPDs optimized for spectral efficiency under color rendition constraints

An important consideration when changing color rendition specification criteria is the effect on engineering targets. $R_a \geq 80$ has been a goalpost for lighting products intended for architectural interiors throughout the development of LED technology, which is reflected in typical product characteristics (see Figure 7). Substantial resources are consumed trying to improve upon existing designs and changing the engineering targets potentially could require a change in direction.

A recent analysis examined the maximum luminous efficacy of radiation (LER)—or spectral efficiency—that could be achieved under various color rendition constraints, including many of those listed in Table 1 [83]. Figure 10 illustrates these limits for SPDs consisting of up to four-primary mixtures having either theoretical (≥ 1 nm full-width-half-maximum [FWHM], 2700 K ≤ CCT ≤ 6500 K, $D_{uv} \leq 0.006$) or more realistic (wavelength-specific minimum FWHM between 12 and 25 nm, 2700 K ≤ CCT ≤ 6500 K, $D_{uv} \leq 0.000$) spectral characteristics. Increasing the tier of the Color Preference criteria (*e.g.*, C to B) has less of an effect on maximum LER than increasing average color fidelity alone. This occurs because the Color Preference criteria focus on ensuring high levels of the psychologically-important [87] red chroma, while remaining relatively flexible for average color fidelity and gamut area. Fortunately, the spectral characteristics of optimized SPDs do not substantially change when the color rendition specification criteria change from R_a based systems to TM-30 based system, because key features of the optimized SPDs are driven by color matching functions, rather than color rendition measures. For new and old alike, narrow red (610-620 nm peak wavelength) and blue (450-460 nm peak wavelength) primaries are the most spectrally efficient. SPDs optimized with higher average color fidelity thresholds require wider FWHM green/amber primaries, whereas SPDs optimized for the preference-based criteria feature a narrower green primary.

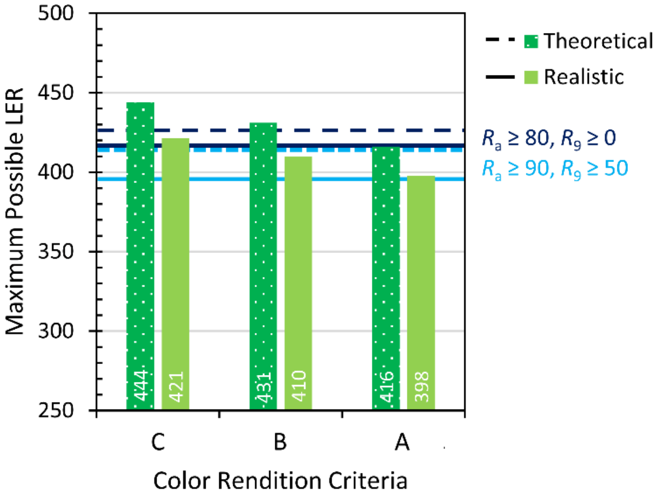

Figure 10. Maximum achievable luminous efficacy of radiation (LER) for different types of SPDs under various color rendition constraints.

4. CONCLUSION

This analysis explored the performance of color rendition specification criteria. It applied 11 different criteria to three sets of data from evaluations of color preference. The three datasets were generated independently and followed best practices for psychophysical experiments on color rendition perception; they featured different populations of observers and different objects. The criteria that most effectively qualified the most preferred products and excluded the least preferred products were the recently proposed Color Preference criteria based on ANSI/IES TM-30-18. This is due to the combination of improved color science and the new types of measures offered by the TM-30 method. There are three tiers with varying levels of stringency:

A “Best”: $R_f \ge 78$, $R_g \ge 95$, $-1\% \le R_{cs,h1} \le 15\%$
B “Good”: $R_f \ge 74$, $R_g \ge 92$, $-7\% \le R_{cs,h1} \le 19\%$
C “Acceptable”: $R_f \ge 70$, $R_g \ge 89$, $-12\% \le R_{cs,h1} \le 23\%$

These criteria are suggested for use where color preference (or naturalness or acceptability) are the most important consideration for lighting color rendition quality. The lowest tier qualifies a vast majority of products that currently meet the frequently used $R_a \ge 80$ with $R_9 \ge 0$ specification. The higher tiers increase the likelihood of building occupants finding the lighting to their liking. Using these color rendition criteria can deliver desired color qualities without a detrimental effect on theoretical limits for energy efficiency, compared to current practices. Products are available today that meet all three tiers of performance. Adoption of these color rendition specification criteria by influential organizations should spur the development of new products that offer more preferred color rendition, which can improve building occupants’ satisfaction and work outcomes while saving energy. Future experimental work and experience with their application may help to further refine these color preference criteria.

ACKNOWLEDGEMENTS

The work of Michael Royer was supported by the U.S. Department of Energy's Solid-State Lighting Program, part of the Office of Energy Efficiency and Renewable Energy.

REFERENCES

[1] [CIE], "CIE 13.3-1995: Method of measuring and specifying colour rendering properties of light sources," Commission Internationale de l'Eclairage, Vienna, Austria (1995).
[2] C. van Trigt, "Color rendering, a reassessment," Color Research & Application, 24(3), 197-206 (1999).
[3] K. Houser, M. Mossman, K. Smet *et al.*, "Tutorial: Color rendering and its applications in lighting," LEUKOS, 12(1-2), 7-26 (2016).
[4] W. Davis, and Y. Ohno, "Color quality scale," Optical Engineering, 49(3), 033602 (2010).
[5] M. Royer, "Comparing measures of average color fidelity," LEUKOS, 14(2), 69-85 (2017).
[6] X. Guo, and K. W. Houser, "A review of colour rendering indices and their application to commercial light sources," Lighting Research & Technology, 36(3), 183-199 (2004).
[7] [IES], "IES-TM-30-18 Method for evaluating light source color rendition," The Illuminating Engineering Society of North America, New York, NY (2018).
[8] [CIE], "CIE 224:2017 CIE 2017 Colour Fidelity Index for accurate scientific use," Commision Internationale de L'Eclairage, Vienna, Austra (2017).
[9] A. David, P. T. Fini, K. W. Houser *et al.*, "Development of the IES method for evaluating the color rendition of light sources," Opt Express, 23(12), 15888-906 (2015).
[10] M. Royer, A. Wilkerson, M. Wei *et al.*, "Human perceptions of colour rendition vary with average fidelity, average gamut, and gamut shape," Lighting Research & Technology, 49(8), 966-991 (2016).
[11] K. Smet, A. David, and L. Whitehead, "Why color space uniformity and sample set spectral uniformity are essential for color rendering measures," LEUKOS, 12(1-2), 39-50 (2015).
[12] [CIE], "CIE 177:2007 Colour rendering of white LED light sources," Commision Internationale de L'Eclairage, Vienna, Austria (2007).
[13] [IES], "IES-TM-30-15 Method for evaluating light source color rendition," The Illuminating Engineering Society of North America, New York, NY (2015).
[14] M. Royer, "Comparing measures of gamut area," LEUKOS, Online Before Print. DOI: 10.1080/15502724.2018.1500485 (2018).
[15] M. Royer, K. Houser, and A. David, "Chroma shift and gamut shape: going beyond average color fidelity and gamut area," LEUKOS, 14(3), 149-165 (2018).
[16] [NEMA], "ANSI/NEMA C78.377-2017 American National Standard for Electric Lamps—Specifications for the Chromaticity of Solid State Lighting Products," National Electral Manufacturers Association, Rosslyn, Virginia (2017).
[17] M. R. Ledbetter, "Origins of CRI > 80," Personal Communication (2017).
[18] [DOE], "ENERGY STAR Program Requirements for CFLs Partner Commitments, Version 2.0," Department of Energy, Washington, DC.(2001).
[19] [CEC],. "Building Energy Standard Title 24, Part 6, Appendix JA8, Qualification Requirements for High Efficacy Light Sources," California Energy Commission (2016).
[20] M. S. Rea, and J. P. Freyssinier, "Color rendering: Beyond pride and prejudice," Color Research & Application, 35(6), 401-409 (2010).
[21] M. S. Rea, and J. P. Freyssinier-Nova, "Color rendering: A tale of two metrics," Color Research & Application, 33(3), 192-202 (2008).
[22] J. P. Freyssinier, and M. Rea, "The Class A color designation for light sources."
[23] M. Royer, A. Wilkerson, M. Wei *et al.*, "Experimental validation of color rendition specification crtieria based on IES TM-30-18," Lighting Research & Technology, Submitted for publication, (2019).
[24] M. Royer, A. Wilkerson, and M. Wei, "Human perceptions of color rendition at different chromaticities," Lighting Research & Technology, Online before print, DOI: 10.1177/1477153517725974 (2017).
[25] J. A. Veitch, M. G. M. Stokkermans, and G. R. Newsham, "Linking Lighting Appraisals to Work Behaviors," Environment and Behavior, 45(2), 198-214 (2013).

[26] J. A. Veitch, G. R. Newsham, P. R. Boyce *et al.*, “Lighting appraisal, well-being and performance in open-plan offices: A linked mechanisms approach,” Lighting Research & Technology*,* 40(2), 133-151 (2008).

[27] C. Sanders, “Color preferences for natural objects,” Illuminating Engineering*,* 47(1), 452-456 (1959).

[28] C. Sanders, “Assessment of color rendition under an iIlluminant using color tolerances for natural objects,” Journal of the Illuminating Engineering Society(54), 640-646 (1959).

[29] S. M. Aston, and H. E. Belichambers, “Illumination, colour rendering and visual clarity,” Lighting Research & Technology*,* 1(4), 259-261 (1969).

[30] C. W. Jerome, “Flattery vs color rendition,” Journal of the Illuminating Engineering Society*,* 1(3), 208-211 (1972).

[31] W. A. Thornton, “A validation of the color-preference index,” Journal of the Illuminating Engineering Society*,* 4(1), 48-52 (1974).

[32] M. S. Rea, A. R. Robertson, and W. M. Petrusic, “Colour rendering of skin under fluorescent lamp illumination,” Color Research & Application*,* 15(2), 80-92 (1990).

[33] J. A. Veitch, L. A. Whitehead, M. Mossman *et al.*, “Chromaticity-matched but spectrally different light source effects on simple and complex color judgments,” Color Research & Application*,* 39(3), 263-274 (2014).

[34] J. A. Veitch, D. K. Tiller, I. Pasini *et al.*, “The effects of fluorescent lighting filters on skin appearance and visual performance,” Journal of the Illuminating Engineering Society*,* 31(1), 40-60 (2002).

[35] E. M. Quellman, and P. R. Boyce, “The light source color preferences of people of different skin tones,” Journal of the Illuminating Engineering Society*,* 31(1), 109-118 (2002).

[36] N. Narendran, and L. Deng, "Color rendering properties of LED light sources."

[37] J. Schanda, and N. Sandor, “Colour rendering, Past – Present – Future,” Proc. Int. Lighting and Colour Conf, Cape Town, South Africa (2003).

[38] J. Schanda, and G. Madar, "Light source quality assessment." Proc CIE 26th Session, Beijing, China (2007).

[39] J. Vanrie, “Appendix 4: Technical report to the user committee of the IWT-TETRA project (80163): The effect of the spectral composition of a light source on the visual appreciation of a composite objectset,” Diepenbeek, Belgium (2009).

[40] S. Jost-Boissard, P. Avouac, and M. Fontoynont, “Assessing the colour quality of LED sources: Naturalness, attractiveness, colourfulness and colour difference,” Lighting Research & Technology*,* 47(7), 769-794 (2014).

[41] S. Jost-Boissard, M. Fontoynont, and J. Blanc-Gonnet, “Perceived lighting quality of LED sources for the presentation of fruit and vegetables,” Journal of Modern Optics*,* 56(13), 1420-1432 (2009).

[42] F. Szabó, P. Csuti, and J. Schanda, "Color preference under different illuminants—new approach of light source coloour quality." PWDAS-43.

[43] K. Smet, W. R. Ryckaert, M. R. Pointer *et al.*, “Memory colours and colour quality evaluation of conventional and solid-state lamps,” Opt Express*,* 18(25), 26229-44 (2010).

[44] A. Liu, A. Tuzikas, A. Zukauskas *et al.*, “Cultural preferences to color quality of illumination of different artwork objects revealed by a color rendition engine,” IEEE Photonics Journal*,* 5(4), 6801010 (2013).

[45] A. Zukauskas, R. Vaicekauskas, P. Vitta *et al.*, “Color rendition engine,” Opt Express*,* 20(5), 5356-67 (2012).

[46] Y. Imai, T. Kotani, and T. Fuchida, “A study of color rendering properties based on color preference in adaptation to LED lighting” Proc of CIE2012 Lighting Quality & Energy Efficiency, Hangzhou, China (2012).

[47] J. M. Spaulding, "Evaluation of desirability assessment techniques for tunable solid state lighting applications." 56, 643-647.

[48] K. A. G. Smet, W. R. Ryckaert, M. R. Pointer *et al.*, “Optimization of colour quality of LED lighting with reference to memory colours,” Lighting Research & Technology*,* 44(1), 7-15 (2012).

[49] R. R. Baniya, R. Dangol, P. Bhusal *et al.*, “User-acceptance studies for simplified light-emitting diode spectra,” Lighting Research & Technology*,* 47(2), 177-191 (2015).

[50] R. Dangol, M. S. Islam, M. Hyvarinen *et al.*, “User acceptance studies for LED office lighting: Preference, naturalness and colourfulness,” Lighting Research & Technology*,* 47(1), 36-53 (2015).

[51] M. S. Islam, R. Dangol, M. Hyvarinen *et al.*, “User preferences for LED lighting in terms of light spectrum,” Lighting Research & Technology*,* 45(6), 641-665 (2013).

[52] R. Dangol, M. Islam, M. Hyvarinen *et al.*, “Subjective preferences and colour quality metrics of LED light sources,” Lighting Research & Technology*,* 45(6), 666-688 (2013).

[53] Y. Kawashima, Y. Ohno, and S. Oh, "Vision experiment on verification of hunt effect for lighting." CIE x039 (2014)

[54] Y. Ohno, G. Fein, and C. Miller, "Vision experiment on chroma saturation for color quality preference." Proc CIE 28th Session, (2015).

[55] Y. Imai, T. Kotani, and T. Fuchida, "A study of color rendering properties based on color preference of objects in adaptation to LED lighting." CIE Centenary Conference "Towards a New Century of Light" (2013).

[56] A. Tsukitani, "Optimization of colour quality for landscape lighting based on feeling of contrast index." CIE Centenary Conference "Towards a New Century of Light" (2013).

[57] F. Szabo, R. Keri, J. Schanda *et al.*, "A study of preferred colour rendering of light sources: Home lighting," Lighting Research & Technology, 48(2), 103-125 (2014).

[58] M. Wei, W. Bao, and H. Huang, "Consideration of light level in specifying light source color rendition," LEUKOS, Online before print. DOI: 10.1080/15502724.2018.1448992 (2018).

[59] M. Wei, and K. W. Houser, "Systematic changes in gamut size affect color preference," LEUKOS, 13(1), 23-32 (2017).

[60] M. Wei, K. Houser, A. David *et al.*, "Colour gamut size and shape influence colour preference," Lighting Research & Technology, Online before print. DOI: 10.1177/1477153516651472 (2016).

[61] Y. Lin, M. Wei, K. Smet *et al.*, "Colour preference varies with lighting application," Lighting Research & Technology, 49(3), 316-328 (2015).

[62] M. Wei, K. W. Houser, G. R. Allen *et al.*, "Color preference under leds with diminished yellow emission," LEUKOS, 10(3), 119-131 (2014).

[63] M. Wei, K. Houser, A. David *et al.*, "Perceptual responses to LED illumination with colour rendering indices of 85 and 97," Lighting Research & Technology, 47(7), 810-827 (2014).

[64] K. Smet, and P. Hanselaer, "Impact of cross-regional differences on color rendition evaluation of white light sources," Opt Express, 23(23), 30216-26 (2015).

[65] X. Tang, and C. Teunissen, "The appreciation of LED-based white light sources by Dutch and Chinese people in three application areas," Lighting Research & Technology, Online before print. DOI: 10.1177/1477153517754130 (2018).

[66] C. Teunissen, F. van der Heijden, S. H. M. Poort *et al.*, "Characterising user preference for white LED light sources with CIE colour rendering index combined with a relative gamut area index," Lighting Research & Technology, 49(4), 461-480 (2016).

[67] T. Khanh, P. Bodrogi, Q. Vinh *et al.*, "Colour preference, naturalness, vividness and colour quality metrics, Part 1: Experiments in a room," Lighting Research & Technology, Online before print: DOI: 10.1177/1477153516643359 (2016).

[68] T. Khanh, and P. Bodrogi, "Colour preference, naturalness, vividness and colour quality metrics, Part 3: Experiments with makeup products and analysis of the complete warm white dataset," Lighting Research & Technology, Online before print. DOI: 10.1177/1477153516669558 (2016).

[69] T. Khanh, P. Bodrogi, Q. Vinh *et al.*, "Colour preference, naturalness, vividness and colour quality metrics, Part 2: Experiments in a viewing booth and analysis of the combined dataset," Lighting Research & Technology, Online before print: DOI: 10.1177/1477153516643570 (2016).

[70] P. Bodrogi, S. Brückner, T. Q. Khanh *et al.*, "Visual assessment of light source color quality," Color Research & Application, 38(1), 4-13 (2013).

[71] F. Zhang, H. Xu, and H. Feng, "Toward a unified model for predicting color quality of light sources," Applied Optics, 56(29), 8186-8195 (2017).

[72] T. Esposito, and K. Houser, "Models of colour quality over a wide range of spectral power distributions," Lighting Research & Technology, Online Before Print. DOI: 10.1177/1477153518765953., (2018).

[73] T. Khanh, and P. Bodrogi, "Colour preference, naturalness, vividness and colour quality metrics, Part 3: Experiments with makeup products and analysis of the complete warm white dataset," Lighting Research & Technology, 50(2), 218-236 (2018).

[74] T. Khanh, P. Bodrogi, Q. Vinh *et al.*, "Colour preference, naturalness, vividness and colour quality metrics, part 4: experiments with still life arrangements at different correlated colour temperatures," Lighting Research & Technology, 50(6), 862-879 (2018).

[75] T. Khanh, P. Bodrogi, X. Guo *et al.*, "Towards a user preference model for interior lighting. Part 2: Experimental results and modelling," Lighting Research & Technology, Online before print, DOI: 10.1177/1477153518816474 (2018).

[76] T. Khanh, P. Bodrogi, X. Guo *et al.*, “Towards a user preference model for interior lighting Part 1: Concept of the user preference model and experimental method,” Lighting Research & Technology, Online before print, DOI: 10.1177/1477153518816469 (2018).

[77] T. Khanh, P. Bodrogi, Q. Vinh *et al.*, “Colour preference, naturalness, vividness and colour quality metrics, Part 2: Experiments in a viewing booth and analysis of the combined dataset,” Lighting Research & Technology, 49(6), 714-726 (2017).

[78] T. Khanh, P. Bodrogi, Q. Vinh *et al.*, “Colour preference, naturalness, vividness and colour quality metrics, Part 1: Experiments in a room,” Lighting Research & Technology, 49(6), 697-713 (2017).

[79] Q. Liu, Z. Huang, M. R. Pointer *et al.*, “Evaluating colour preference of lighting with an empty light booth,” Lighting Research & Technology, 1477153517727330 (2017).

[80] Z. Huang, Q. Liu, S. Westland *et al.*, “Light dominates colour preference when correlated colour temperature differs,” Lighting Research & Technology, Online before print, DOI: 10.1177/1477153517713542 (2017).

[81] M. Royer, and M. Wei, “The role of presented objects in deriving color preference criteria from psychophysical studies,” LEUKOS, 13(3), 143-157 (2017).

[82] [WELL], “WELL Building Standard v2 pilot” International WELL Building Institute (2018).

[83] M. Royer, “Evaluating tradeoffs between energy efficiency and color rendition,” Optics Express, Submitted for publication (2019).

[84] D. R. International, “LED Lighting Facts Database,” Available from http://www.lightingfacts.com (2019).

[85] [DOE], “CALiPER Report 23: Photometric testing of white-tunable LED luminaires,” Pacific Northweset National Laboratory, Richland, WA (2018).

[86] A. David, T. Esposito, K. Houser *et al.*, “A vector field color rendition model for characterizing color shifts and metameric mismatch,” LEUKOS, In press (2019).

[87] A. J. Elliot, and M. A. Maier, “Color psychology: effects of perceiving color on psychological functioning in humans,” Annu Rev Psychol, 65, 95-120 (2014).