THE TRUE COST OF EFFICIENCY:

WHAT WE'VE SACRIFICED WITH THE RISE OF LED LUMINAIRES IN ARCHITECTURE

by

Roberto Esquenazi Alkabes

A Master's Thesis
Submitted to the Department of Lighting Design
at Parsons School of Design at The New School
In Fulfillment of the Requirements
for a Master's in Fine Arts in Lighting Design Degree
May 2025

Thesis Studio Faculty: Dr. Craig Bernecker & Francesca Bastianini
Thesis Seminar Faculty: Kyle McGahan

Table of Contents

1. Introduction

Lighting technology has been intrinsically linked to human civilization. Early humans, beginning with *Homo Erectus*, harnessed fire, initiating a legacy of controlled illumination passed down through *Homo Sapiens* and beyond (Guarneri, 2015). As civilization developed so did lighting technology, rapidly advancing especially throughout the industrial revolution. This revolution led to the discovery and widespread use of coal gas, a by-product from coal refinement, which facilitated widespread industrial lighting adoption across Europe (Bowers, 1998). As the general industry became more advanced, entertainment endeavors swiftly followed. The entertainment industry, however, is not tied down by limitations that the wider industry faced. As such, theatre specifically, often served as the proving ground for emerging lighting technologies before their broader architectural implementation. Frederick Winsor's 1804 installation of the first gas lighting system at London's Lyceum Theatre exemplifies this phenomenon (Penzel, 1978). Parallel developments included kerosene lamps from crude oil extraction, and chemically produced sources like limelight, popular in theatre, and acetylene lamps, widely used domestically and industrially (Guarneri, 2015).

Electrical innovations emerged subsequently, such as arc lighting, which was first fully installed at the Opéra Theatre in Paris, where carbon electrodes separated by an electric current generated consistent illumination. This evolved into incandescent technology, notably implemented at London's Savoy Theatre in 1881. This marks another instance of theatrical pioneering (Penzel, 1978). The 20th and 21st centuries saw further developments, including discharge, fluorescent, and compact fluorescent lamps, each driven primarily by energy efficiency, and other limiting factors.

Today's prevalent lighting technology is the light emitting diode (LED). Following each historical advancement that prioritized energy efficiency, LEDs aimed to reduce heat waste and maximize luminous output. However, this efficiency-centric evolution has overlooked critical perceptual qualities. Each technological shift introduced spectral compromises: fluorescent lamps had discontinuous spectra inferior to incandescents, yet energy and cost efficiency justified their adoption despite perceptual drawbacks.

Similarly, High Pressure Sodium lamps, despite poor color rendering, gained preference due to economic and energy considerations. (Tabaka, 2021; Cho et al., 2017)

Theatre, however, was never much concerned with energy efficiency, but rather with the human experience and the illumination of the stage. As a consequence, theatre has been reluctant to move on from technologies such as tungsten halogen, that arguably render skin tones and the built show environment in a more favorable light. Thus, technological innovation in theatrical settings has been concerned with improving the theatre-goer experience and not energy considerations. More recently, Electronic Theater Controls (ETC), an industry leader in theatrical lighting, developed a fixture with 8 colored chips to provide theatre practitioners with an LED luminaire that provides a spectral power distribution (SPD) closer to tungsten and daylight sources. No such luminaire exists inside an architectural setting, but we are seeing more and more technologies developed that mix more than one LED to create a broader spectral distribution.These advancements in theatre seem to have illuminated a major drawback of architectural lighting innovation;

We've sacrificed lighting quality, especially when rendering skin tones and the built environment in favor of more energy and cost efficient light sources. Metrics, although useful, have failed to fully acknowledge the extent of this issue in the human experience. This study tests whether observers can perceive meaningful differences among LED luminaires that share the same correlated color temperature, illuminance and Δuv but differ only in their spectral power distributions (SPDs).

In doing so, this research also seeks to provide actionable feedback to the lighting industry—particularly in how luminaires are designed, evaluated, and selected. Are we actively attempting to recreate the perceptual qualities of legacy sources like tungsten, or even emulate the universally flattering tones of "Golden Hour" daylight? If so, what technologies or configurations actually come closest to achieving this goal, and are those differences perceptible to end users? This study aims to assess those perceptual gaps, and to begin imagining how lighting standards and practices might evolve to better reflect human-centric priorities in design.

2. Background

2.1 Evolution of Color Metrics

Lighting quality has historically been assessed through objective metrics designed to standardize performance across a wide variety of sources. The Color Rendering Index (CRI), introduced by the CIE in the 1960s, evaluates how accurately a light source renders a set of fifteen standardized color samples compared to a reference illuminant (CIE, 1995). Scores (Ra) range from 0 to 100, with higher values indicating closer agreement with the reference. Despite its ubiquity, CRI's limitations have been well documented: it neglects spectral power distribution (SPD) nuances, relies on a small sample of idealized pastel colors, and fails to predict rendering of saturated or skin-tone-relevant hues (Ohno, 2005) As a result, sources with spiky or discontinuous SPDs such as fluorescent and early LED phosphor-converted lamps can achieve deceptively high CRI values despite perceptual deficiencies.

In response, the Illuminating Engineering Society introduced TM-30 in 2015, later updated to TM-30-20, which benchmarks 99 color evaluation samples to report fidelity (Rf) and gamut (Rg) indices, supplemented by color vector graphics that reveal hue and saturation shifts across the visible spectrum (David et al., 2015). TM-30 improves upon CRI by capturing a wider array of colors, notably including deep reds and greens critical for accurate skin-tone and material rendering and by visualizing deviations in a chromaticity diagram format. The CIE themselves seemed to agree with some of the metrics that TM-30 used in their position statement published earlier this year (CIE, 2025).

However, like CRI, TM-30 remains a scientific metric and does not directly incorporate psychophysical responses or user-centered evaluations. There are currently no universally accepted TM-30 thresholds for “acceptable” rendering, and practical interpretation requires careful analysis of combined Rf/Rg values and vector plots rather than reliance on a single aggregated score like CRI.

2.2 Spectral Power Distribution and Human Perception

Spectral Power Distribution (SPD) graphs illustrate the relative energy output of a light source at each wavelength, offering clear insight into which portions of the spectrum are emphasized or lacking. Broader-spectrum sources, which more closely approximate the continuous distribution of natural daylight, tend to enhance visual comfort, material differentiation, and accurate skin-tone reproduction (Rea & Freyssinier, 2010). Conversely, sources with gaps, such as typical phosphor-converted white LEDs with limited red output, often render warm hues inaccurately, leading to diminished visual appeal in built environments.

Peer-reviewed investigations have directly linked SPD characteristics to perceptual outcomes. For instance, studies comparing narrow-band phosphor LEDs with multi-channel LED arrays found significant improvements in skin-tone naturalness, color discrimination tasks, and overall observer preference when long-wavelength content was enhanced (Royer et al., 2017; Narendran et al., 2004). These findings are supported by psychophysical research indicating that human chromatic adaptation is optimized by continuous spectral content, which reduces visual fatigue and enhances color constancy under changing ambient conditions (Brainard et al., 2002).

Of particular interest is the comparison to "Golden Hour" daylight—the brief period shortly after sunrise or before sunset where sunlight takes on a warm, flattering quality due to the low solar angle and increased atmospheric scattering. Anecdotally, this time of day is widely regarded as the most photogenic, and its effect on skin tone and environmental color saturation has influenced both film lighting and photography. Some lighting manufacturers have implicitly attempted to replicate this effect using high-R9 red channels or dynamic CCT tuning, though few studies directly test whether people actually perceive these attempts as successful or naturalistic. A study on facial attractiveness and lighting conditions found that warm-toned faces produced more favorable ratings of facial appearance (Jones et at., 2004), aligning with broader industry assumptions. However, comprehensive evaluation of how closely artificial sources approximate this spectral ideal remains limited.

By drawing analogies between natural daylight conditions and LED spectral design, this section emphasizes the perceptual consequences of SPD trade-offs. Lighting designers often aim for fidelity without fully considering how perceptual “success” might be context-dependent, emotional, or even symbolic—elements that Golden Hour naturally captures, but current SPD targets often overlook.

2.3 Lighting Quality and Psychological and Cognitive Performance

Beyond color rendering, lighting quality profoundly influences human well-being, productivity, and cognitive function. In healthcare settings, exposure to broad-spectrum, high-fidelity lighting has been associated with reduced patient recovery times, improved sleep–wake cycles, and lower rates of medical errors among staff (Boyce et al., 2000 - Perceptions of safety at night in different lighting conditions). Educational research demonstrates that classrooms illuminated with lighting approximating daylight can enhance student alertness, concentration, and reading comprehension, while reducing behavioral issues and eye strain (Mott et al., 201). Similarly, workplace studies report that employees under broad-spectrum lighting exhibit higher task performance, lower stress levels, and increased satisfaction compared to those under standard fluorescent or narrow-band LED sources (Viola et al., 2008). These outcomes underscore the multifaceted benefits of prioritizing spectral quality alongside quantitative efficiency metrics.

2.4 Theatrical Innovations and Multi-Chip Luminaires

The theatrical lighting industry has long driven innovations in spectral quality. Early adopters of arc and incandescent systems prioritized visual impact and fidelity over efficiency, leveraging technologies such as limelight and tungsten filaments. The advent of LED introduced new challenges, but also opportunities: multi-chip LED luminaires, like ETC’s LUSTR 3, incorporate eight individually controllable LEDs, spanning deep red to indigo, to achieve an SPD that more closely approximates natural light while maintaining adjustable control (ETC, 2023). The inclusion of a "deep red" chip specifically addresses red chromaticity deficiencies, allegedly resulting in perceptibly improved skin-tone rendering and material saturation without sacrificing overall illumination levels.

2.5 Barriers to Architectural Adoption

Despite demonstrated perceptual advantages, multi-chip luminaires face significant obstacles in architectural contexts. Regulatory frameworks, such as building energy codes and green certifications, primarily reward lumen-per-watt metrics, disincentivizing higher-order spectral designs that may reduce overall efficacy (US Department of E., 2020). Economically, the increased upfront costs of multi-chip fixtures which are often double or triple that of standard LED alternatives, pose financial barriers for large-scale installations. Logistically, the larger physical size, complex heat management requirements, and proprietary control systems complicate integration into existing architectural infrastructures. Cultural factors further entrench these barriers: architectural lighting specifications frequently rely on manufacturers' photometric data sheets and CRI/TM-30 scores without room for perceptual or experiential considerations, while control systems remain vendor-specific and lack the plug-and-play interoperability seen in most theatrical DMX systems, for example.

By synthesizing advances in colorimetric theory, psychophysical research, and industry practice, this literature review illuminates the critical gap between technical lighting standards and human perceptual needs. The following methods outline an empirical approach to assess whether wider-spectrum multi-chip luminaires can meaningfully enhance perceptual quality in architectural settings when compared against conventional phosphor-converted and RGB LED solutions.

3. Methods

3.1 Participants

A convenience sampling strategy yielded fifteen volunteers (12 female, 3 male) recruited from the Parsons School of Design community in New York City. Eligible participants were between 23 and 42 years old (mean = 26.6 years, SD = 5.1) and reported normal or corrected-to-normal vision, with no history of color vision deficiencies. Recruitment materials described the study as an investigation into lighting perception, and participants provided written informed consent in accordance with Institutional Review Board protocols. Demographic data was collected, including age, gender, and self-identified ethnicity (8 White/Caucasian [one Latino], 5 Asian, 1 Black/African American, 1 undisclosed). Although many participants had backgrounds in design or lighting, no formal exclusion criteria beyond visual capability were imposed. No financial compensation was offered, but participants were provided with a summary of overall study findings upon completion.

The Human Research Protection Program at The New School determined that this activity does not constitute "human subjects research" as defined by federal regulations for the protection of human subjects. As such, Institutional Review Board (IRB) review was not required.

3.2 Experimental Setup and Apparatus

The study environment was constructed within the Light Lab at 25 E 13th Street, Manhattan, using a modular booth design to standardize viewing conditions. The primary structure comprised three 4’ × 6’ panels of medium-density fiberboard (MDF), painted matte white, joined at 90° angles to form a U-shaped enclosure. A secondary black-painted U-shape, with a central 24” × 36” mirror and two 2’ × 6’ side panels, was positioned 1.5’ away from the white enclosure to create a reflective viewing portal. The booth was draped with diffusing white fabric to eliminate hotspots and minimize glare, while lateral entry apertures and peripheral surfaces were shielded with black opaque material to prevent ambient light intrusion. A 40” × 30” abstract painting (Fig. 1b) mounted on the rear white panel provided a consistent environmental reference point and visual interest.

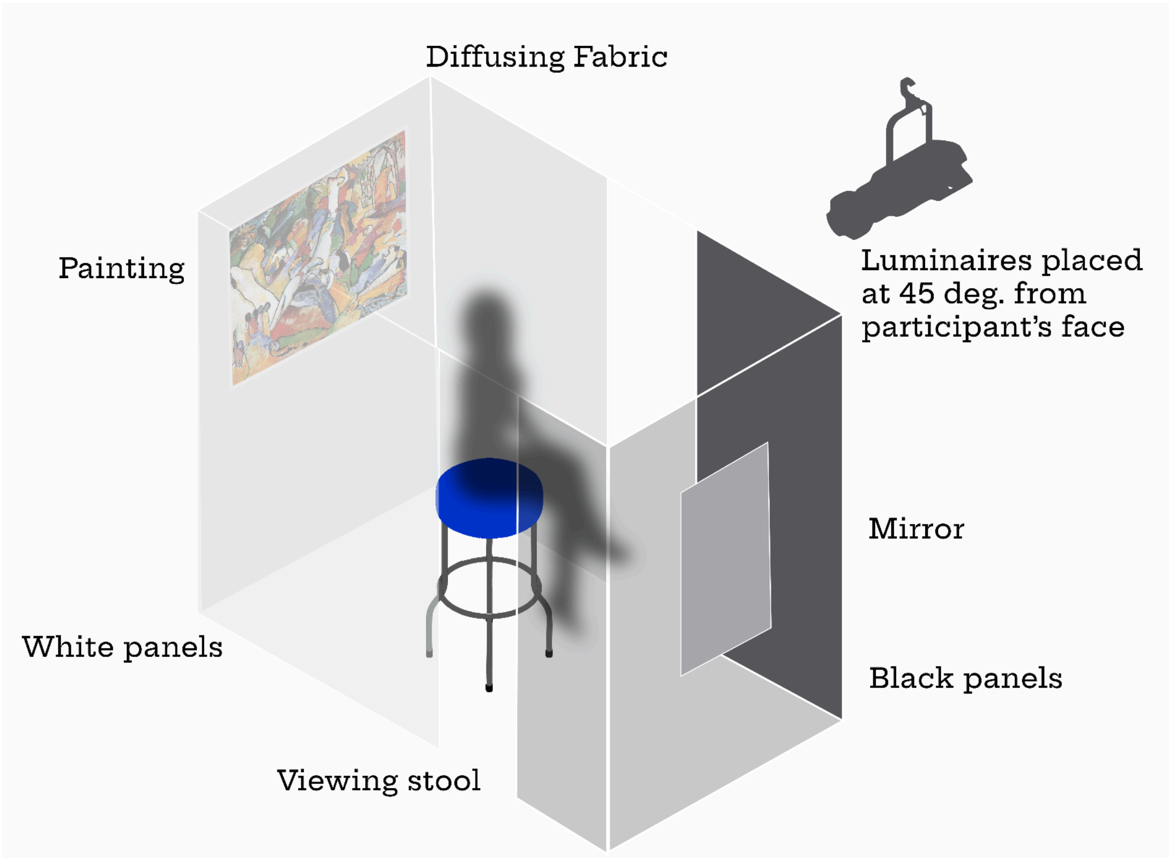

Figure 1a. *Axonometric drawing of experimental setup.*

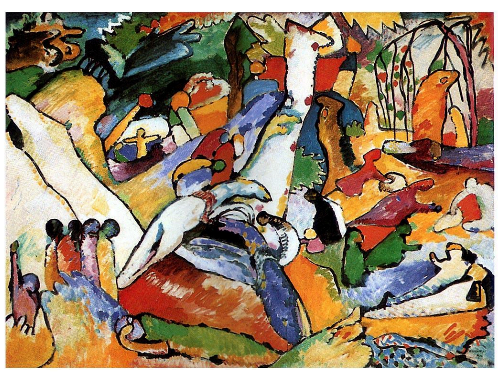

Figure 1b. *Sketch for "Composition II" by Wassily Kandinsky, 1910 (Wikimedia Commons 2019).*

Three luminaire types were mounted to the room's unistrut grid (≈9' ceiling height) at a 45° elevation relative to the participant's seated eye level on a stool positioned 1.5’ from the mirror. Fixtures included: (L1) an ETC LUSTR 3 eight-chip LED luminaire; (L2) an ETC ColorSource Spot V in three-channel RGB mode; and (L3) 3 SATCO PAR38 Flood static white LED at each of the CCTs. All fixtures were calibrated to nominal correlated color temperatures of 2700K, 3000K, and 3500K, and matched for illuminance (~75 lux at eye level) and chromaticity (Δuv < 0.005) using a Sekonic C-4000 spectral and illuminance meter. SPD and TM-30 metrics for each condition were recorded and archived. A DMX control console ensured consistent intensity output, while fixtures were aligned laterally within 6” of one another to minimize directional differences. (See Fig. 1a for booth schematic and fixture layout; Fig. 2a for sample SPD profiles, and Fig. 2b for full illuminant data.)

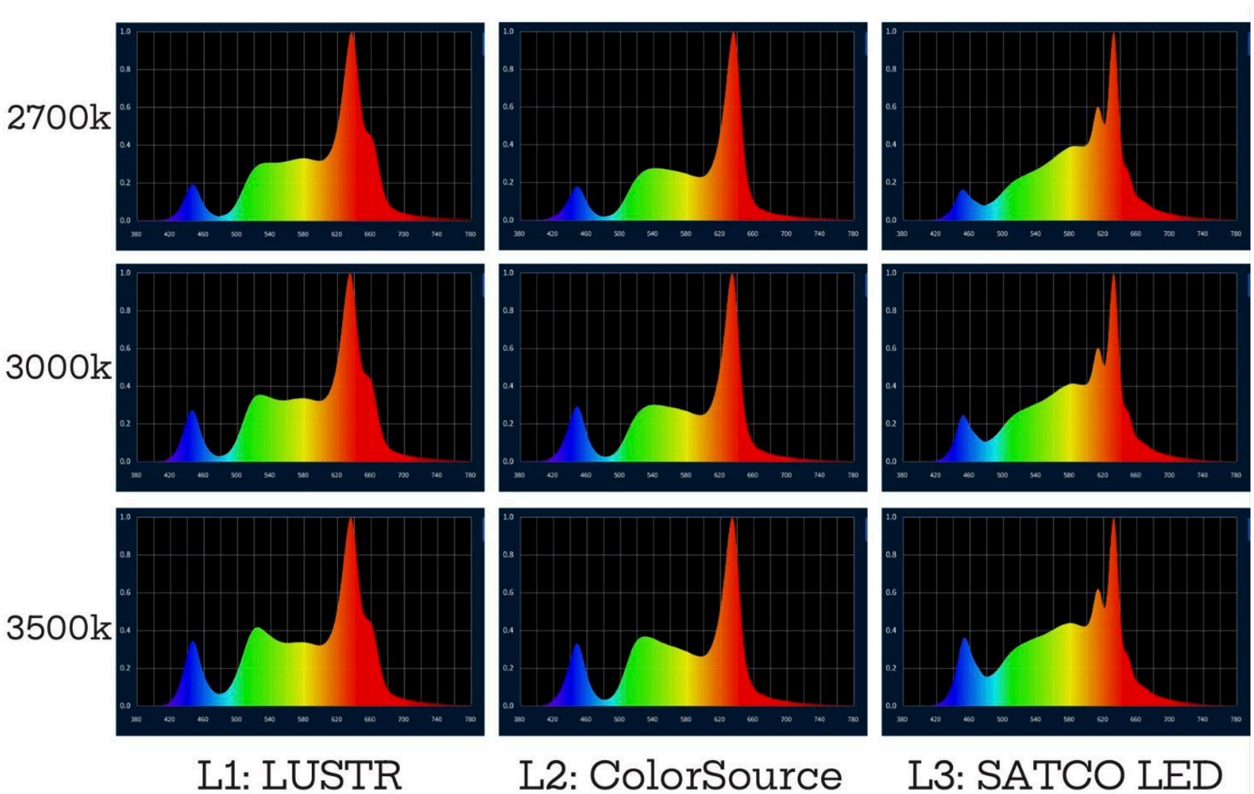

Figure 2a. *SPD graphs measured with Sekonic C-4000 light meter.*

Illuminant Code	lux	TCP	Δuv	Ra	Rf	Rg
LUSTR_2700	76.1	2642	0.0001	86	86	112
CS_2700	76.5	2650	-0.0023	82.9	84	115
LED_2700	76.2	2618	0.0005	94.4	92	100
LUSTR_3000	72.7	2883	-0.0014	84.6	85	113
CS_3000	72	2877	-0.0087	83.8	82	116
LED_3000	72.5	2886	-0.0005	95.3	92	101
LUSTR_3500	72	3245	-0.0018	82.7	86	113
CS_3500	72.6	3279	-0.0021	82.7	85	114
LED_3500	72.7	3258	-0.0014	96.5	92	101

Figure 2b. *Full illuminant data for each condition, taken at participant eye level.*

3.3 Procedure

Participants completed an informed consent form and demographic survey before entering the booth alone. Standardized instructions, read verbatim by the experimenter, explained the sequence of lighting presentations and response tasks. Each trial block consisted of one CCT condition (2700K, 3000K, or 3500K), with block order randomized per participant using a computer-generated sequence. Within each block, participants experienced the three luminaires in randomized order to mitigate anchor bias.

Each presentation began with a 30-second dark adaptation (all fixtures off), followed by 20 seconds of illumination from each luminaire. A 10-second blackout interval preceded subsequent fixtures. After viewing all three fixtures, ambient illumination was provided so that participants could complete a six-item 7-point Likert questionnaire for each luminaire, assessing: (1) skin-tone appearance; (2) environmental color vividness and accuracy; (3) perceived warmth; (4) visual comfort; (5) overall liking; and (6) ambiance. Participants then indicated their forced-choice preference among the three fixtures and provided brief verbal feedback, which was transcribed for thematic analysis.

∞	AMBIENT LIGHT (SETUP)	∞
30s	BLACKOUT	30s
20s	ILLUMINANT "A" (L1, L2 OR L3)	20s
20s	ILLUMINANT "B" (L1, L2 OR L3)	20s
20s	ILLUMINANT "C" (L1, L2 OR L3)	20s
∞	AMBIENT LIGHT (QUESTIONNAIRE)	∞

Figure 3. *Task schedule*

This process was repeated for the remaining two CCT blocks, with identical timing and response procedures. Total session duration per participant averaged 20 minutes, including setup and debriefing. After the final block, participants were debriefed, thanked for their participation, and invited to ask questions about the study's aims.

3.4 Data Analysis

This study employed a mixed-methods approach to assess participant responses to various LED luminaires under controlled perceptual conditions. Quantitative data from Likert-scale questionnaires and forced preference tasks were complemented by qualitative data from open-ended participant feedback. Both sets of data were analyzed using rigorous methods to explore potential perceptual differences between illuminants.

<u>3.4.1 Quantitative Analysis</u>

a. ***Descriptive and Non-Parametric Statistics***

Likert-scale responses for each of the six perceptual dimensions (e.g., skin tone accuracy, warmth, comfort) were analyzed across three luminaire types: ETC LUSTR 3 (8-chip), ETC ColorSource V (3-chip), and SATCO PAR38 (static white). Responses were first averaged per participant and per fixture. Given the ordinal nature of Likert data and the small sample size (n=15), non-parametric tests were selected.

The Friedman test was used to detect statistically significant differences in rankings across the three lighting conditions. Where applicable, Wilcoxon signed-rank tests were conducted as post-hoc pairwise comparisons between luminaires to further examine specific preference trends. A significance threshold of $p < .05$ was applied, though effect sizes and practical relevance were emphasized given the exploratory nature of the study.

b. Bayesian Inference

Due to the limitations of small sample sizes and the exploratory framing of this research, a Bayesian statistical framework was also used to assess forced preference task outcomes. Instead of relying solely on frequentist p-values, Bayesian methods allowed for a probabilistic interpretation of participant preferences and the relative plausibility of one luminaire being favored over another.

A Beta-binomial model was used to generate posterior distributions for each luminaire's likelihood of being chosen. From these distributions, Bayes Factors (BF_{10}) were calculated to quantify the strength of evidence for differences between luminaires. For example, the posterior probability that LUSTR was preferred over the SATCO LED was approximately 91%, yielding a Bayes Factor of ~10.3, indicating moderate to strong evidence in favor of LUSTR. In contrast, LUSTR vs. ColorSource yielded BF ≈ 1, suggesting negligible difference.

Bayesian inference was particularly useful in contextualizing trends that were not statistically significant under frequentist analysis but nonetheless showed consistent patterns across participants.

3.4.2 Qualitative Analysis

To deepen the analysis, verbal participant comments provided during the forced preference task were thematically analyzed to identify common perceptual impressions associated with each luminaire.

a. Coding Framework

An initial set of codes was developed inductively from a preliminary reading of the full response set. Themes included color accuracy, skin tone rendering, warmth, visual comfort, and neutrality. Each comment was assigned one or more codes depending on content.

b. Thematic Synthesis

Themes were aggregated and cross-tabulated by luminaire type. Notably, comments related to skin tone rendering were most frequently associated with the LUSTR (n=4) vs (n=2 for ColorSource and SATCO), while color vividness appeared unexpectedly in relation to the ColorSource (n=8). Although subjective, these impressions provided valuable insight into how participants interpreted spectral differences, complementing the quantitative findings.

3.4.3 Integration of Findings

By triangulating forced-choice data, Likert ratings, and thematic responses, the analysis aimed to answer whether spectral improvements offered by 8-chip luminaires yield perceptual benefits detectable by end users. While statistical significance was limited, recurring patterns in both preference and language suggest that participants could perceive subtle differences—particularly in skin tone rendering—supporting the need for further research into human-centric lighting metrics that evaluate more than just quantitative approaches to lighting.

4. Results

This section presents the quantitative and qualitative outcomes of the study, including descriptive statistics, inferential testing, Bayesian inference, and thematic analysis. Visual aids such as chromaticity plots, TM-30 graphics, and bar charts are referenced to support interpretation.

4.1 Quantitative Findings

4.1.1 Likert Scale Ratings

Participants evaluated six perceptual attributes under each illuminant condition using a 7-point Likert scale. Given the ordinal nature of the data and within-subjects design, the Friedman test was used to assess statistically significant differences across illuminants at three CCTs (2700K, 3000K, 3500K), totaling nine conditions per question (df=8).

Perceptual Attribute	Friedman χ^2 (Q)	p-value	Kendall's W
Q1. The lighting made my skin tone look good.	15.3	0.054	0.128
Q2. The colors in the environment appeared vivid and accurate.	20.11	0.0099	0.168
Q3. The lighting felt warm.	42.9	< 0.001	0.357
Q4. The lighting was comfortable to look at.	15.17	0.056	0.126
Q5. I liked this lighting overall.	12.49	0.131	0.104
Q6. The lighting created a pleasant or inviting atmosphere.	13.54	0.095	0.113

A statistically significant difference was found in *Color Accuracy* (Q2) and *Warmth* (Q3), with the latter showing the strongest agreement among participants naming SATCO as the "warmest" even after all CCTs, Δuv and Illuminance was matched (Kendall's W = 0.357).

Other dimensions such as *Skin Tone Rendering* (Q1) and *Comfort* (Q4) approached significance, suggesting emerging trends that merit further investigation. Below is each Likert question plotted in a box and whisker chart:

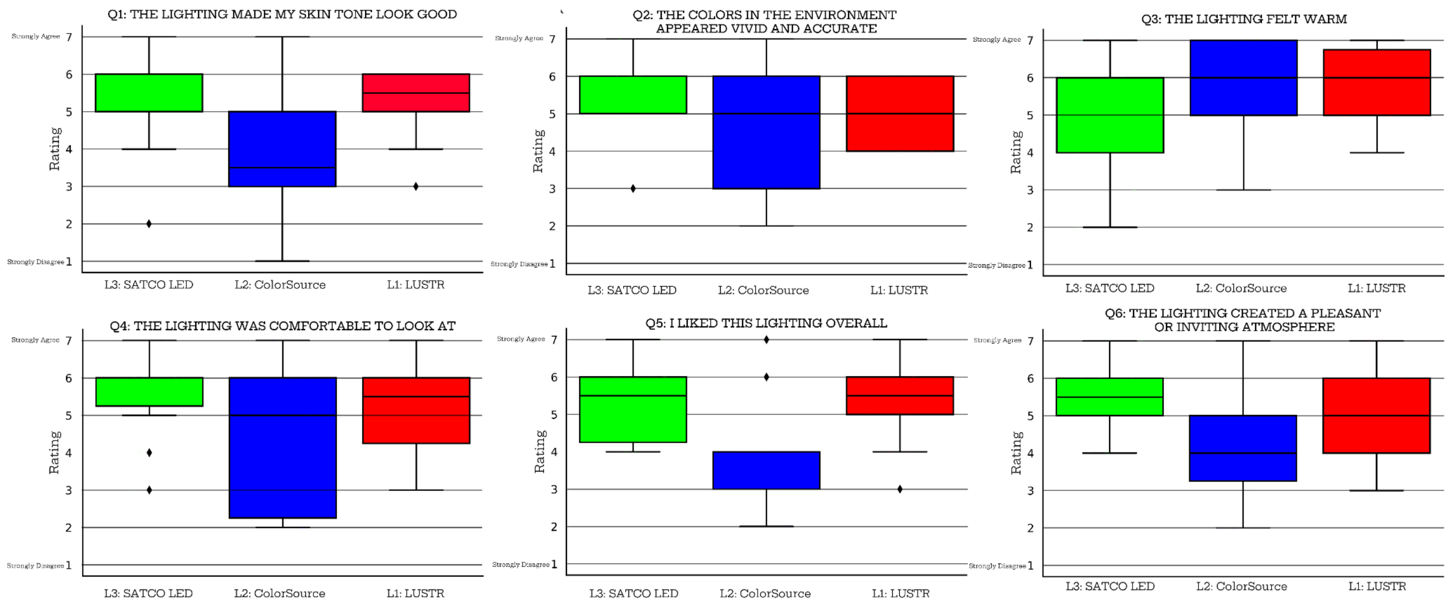

Figure 4. *Box and Whisker Plots for each question.*

4.1.2 Forced Preference Task

In each CCT block, participants selected which of the three illuminants they preferred. Preferences were aggregated across three trials.

Luminaire	Chosen 1x (%)	Chosen 2x (%)	Chosen 3x (%)
LUSTR	40%	33%	7%
ColorSource	27%	20%	0%
SATCO LED	33%	13%	0%

Only 1 participant (7%) consistently chose the same luminaire across all three trials, while 13 participants (87%) chose the same fixture twice, suggesting moderate consistency. The LUSTR was chosen moderately more consistently than the ColorSource and much more consistently than the SATCO LED.

4.2 Bayesian Analysis

Given the small sample size and exploratory nature of the study, Bayesian inference was used to complement traditional statistics. Each participant completed three forced-choice trials, selecting the luminaire they preferred in each CCT block. Out of 45 total selections the results were:

Luminaire	**Number of Times Chosen**
LUSTR	17
ColorSource (CS)	17
SATCO LED (LED)	11

A Bayesian beta-binomial model with uniform priors was used to estimate posterior probabilities for each luminaire's likelihood of being preferred.

Luminaire	**Posterior Mean**	**95% Credible Interval**	**α (successes + 1)**	**β (failures + 1)**
LUSTR	0.383	[0.251, 0.525]	18	29
CS	0.383	[0.251, 0.525]	18	29
LED	0.255	[0.143, 0.388]	12	35

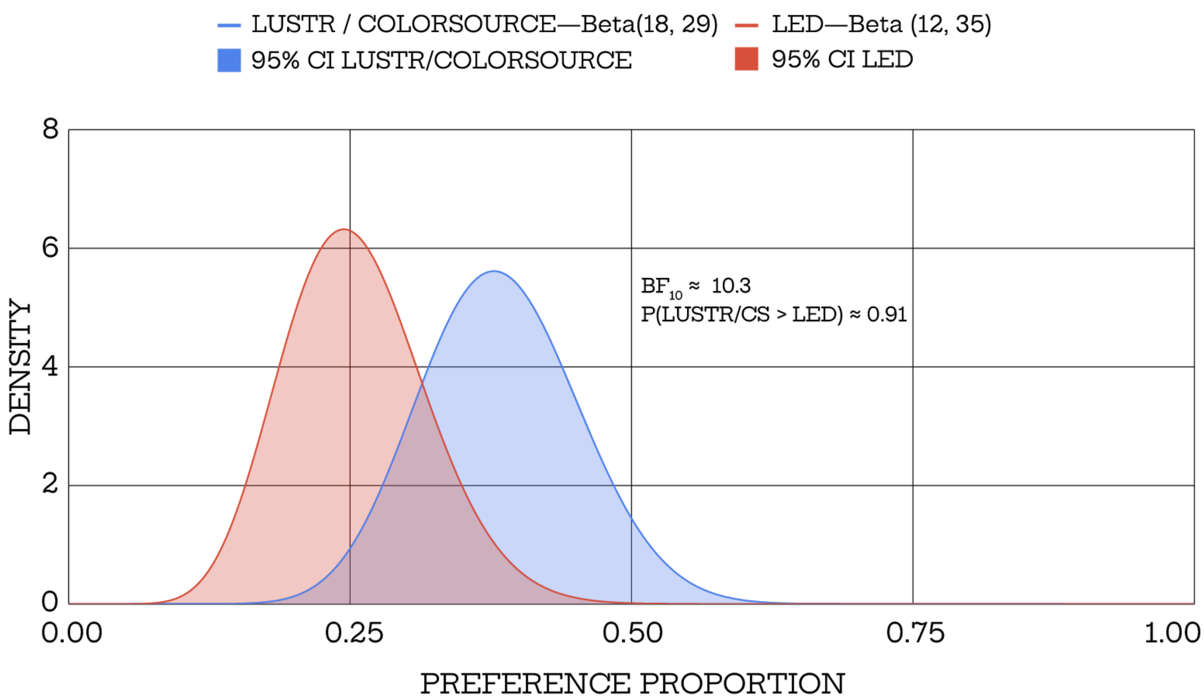

Figure 5. *Posterior distributions for preference (Bayesian)*

Comparison	P(A > B)	Bayes Factor (BF_{10})	Interpretation
LUSTR > LED	0.911	10.2	Strong evidence in favor of LUSTR
CS > LED	0.911	10.22	Strong evidence in favor of CS
LUSTR > CS	0.498	≈ 1.00	No evidence for either being better

These results suggest a significant perceptual advantage for both LUSTR and ColorSource over the static white LED. However, there is no discernible difference between LUSTR and ColorSource in this dataset.

4.3 Spectral & Photometric Comparison

All luminaires were calibrated to match CCT (±50 K), chromaticity (Δuv ≤ 0.002), and illuminance levels (±5 lux), where possible. Despite this calibration, spectral power distribution (SPD) and TM-30 metrics varied across fixtures. (See Fig. 2a & 2b)

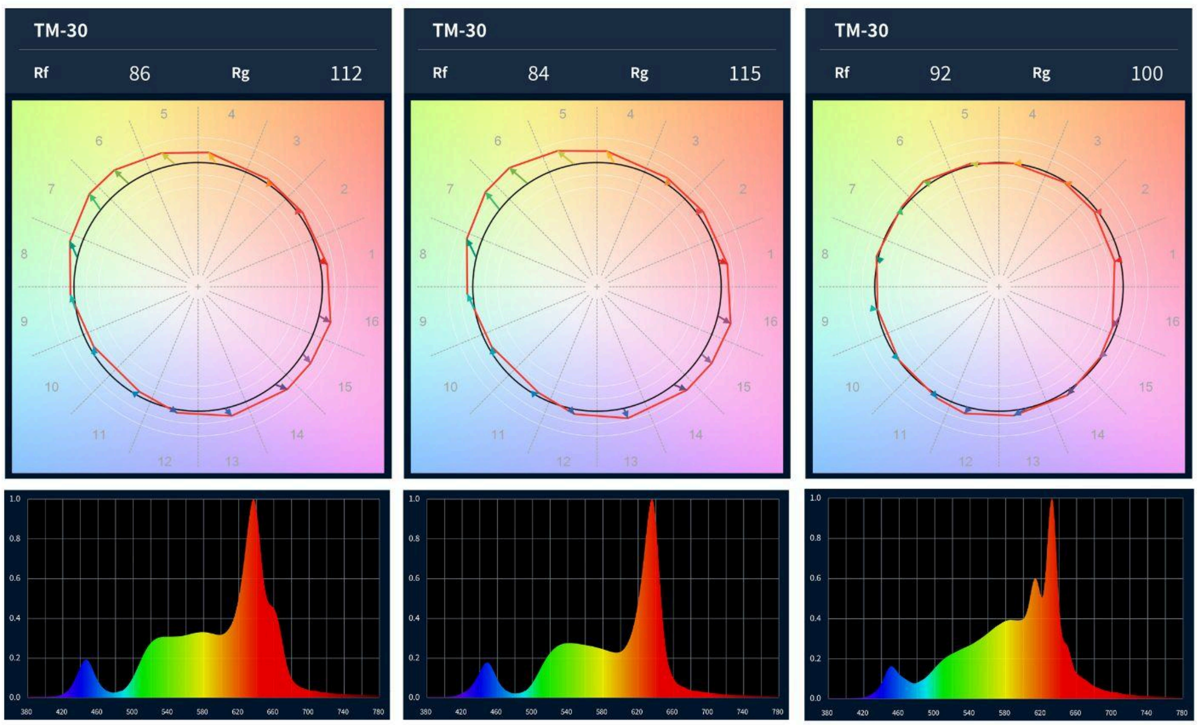

Figure 6. *TM-30 Scores and SDP graphs for all luminaires at 2700k. From left to right, L1, L2, & L3.*

As seen in Fig. 6, in the SPD the LUSTR demonstrated greater spectral fullness in deep reds (≈650-720nm) and dark purples (≈420nm). When comparing the TM-30 Rf values, however, we can see that the ColorSource = 82, LUSTR = 86 and SATCO = 92. The Rg values, on the other hand, varied from SATCO = 100, LUSTR = 112 and ColorSource = 116. These measurements seem to contradict impressions in qualitative responses, especially regarding skin tone rendering and color vividness.

4.4 Thematic Analysis of Open-Ended Responses

Participant comments from the forced preference task were coded inductively to identify recurring perceptual themes.

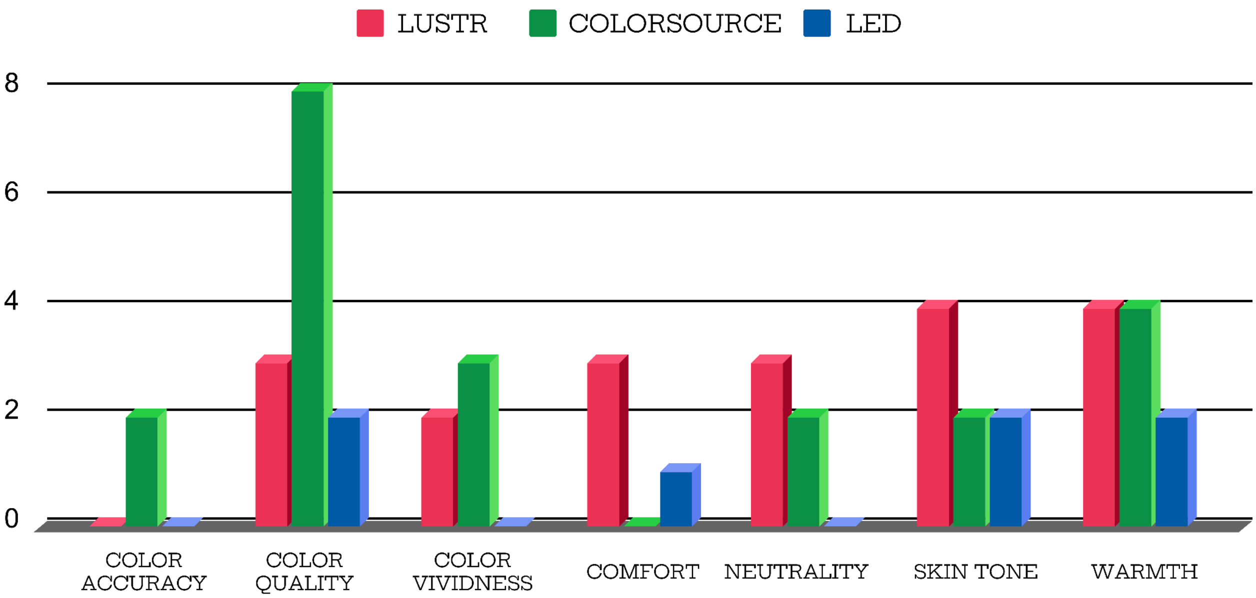

Figure 7. *Theme incidence by illuminant.*

Theme	Lustr	ColorSource	LED
Color Accuracy	1	2	0
Color Quality	3	8	2
Color Vividness	2	3	0
Comfort	3	0	1
Neutrality	3	2	0
Skin Tone	4	2	2
Warmth	4	4	2

Except of Participant Qualitative Responses:

1. On LUSTR: *"It's more neutral in the color tone — I can see a little bit more red in my skin tone. Another thing, the painting looks more vivid."*
2. On ColorSource: *"It rendered the colors better, the others were too warm and made me look yellow"*
3. On SATCO LED: *"The color rendering was better compared to the other two."*

4.5 Summary of Key Results

While the metrics provided a relevant baseline for the luminaires, they failed to provide the whole story. Even though all luminaires were matched for CCT, Δuv and illuminance, the SATCO LED appeared statistically warmer ($p=0.047$, $r=.75$) than the ColorSource and the LUSTR. Even being the “best” luminaire on paper based on traditional metrics like CRI and TM-30, there was moderate evidence for participant preference toward LUSTR and ColorSource over SATCO. Qualitative feedback favored LUSTR for skin tone rendering (n=4) and ColorSource for Color Accuracy (n=8), though perception varied by individual. Spectral data confirmed broader SPD coverage by the LUSTR, aligning with some subjective impressions but not all.

5. Discussion

This study investigated whether participants could perceive meaningful differences among three luminaires with distinct spectral characteristics. While the quantitative results were modest, both statistical and Bayesian analyses suggest slight perceptual advantages for broader-spectrum luminaires—particularly the ETC LUSTR 3 and ColorSource V over a conventional static white SATCO LED. These findings align with prior literature that identifies richer SPDs as more favorable for tasks involving color rendering and skin tone perception (Houser et al., 2015; Davis & Ohno, 2010).

The Lustr, with its 8-chip configuration, demonstrated broader SPD coverage, particularly in the red and violet regions, which are often truncated in typical phosphor-converted white LEDs. This spectral completeness may explain why participants frequently cited skin tone accuracy (n=4), comfort (n=3), and neutrality (n=3) when selecting the LUSTR as their preference. Interestingly, the ColorSource received the most mentions related to color quality and vividness (n=8, n=3), despite being a 3-chip RGB system and the “worst” luminaire based on traditional metrics alone. This suggests that perceptual impact is not solely determined by the spectral breadth, but also by how saturation and balance interact in the user’s context.

This perceptual nuance highlights a critical limitation in relying on metrics like TM-30 or CRI as sole predictors of visual quality. While TM-30 provides a richer dataset than CRI, it still abstracts perceptual experience into generalized fidelity and gamut indices. Studies like Rea & Freyssinier (2010) argue that subjective preference often diverges from objective metrics, especially when spatial or facial cues are involved. The high TM-30 Rf values of the LUSTR and ColorSource did not fully account for their distinct qualitative impacts, indicating a persistent gap between measured and experienced light.

For architectural lighting practice, these results encourage reconsideration of fixture selection criteria. Design decisions based purely on efficiency or compliance and traditional metrics, risk neglecting human-centric values like skin tone fidelity and environmental rendition. Regulatory frameworks could evolve to incorporate perceptual testing or user feedback into standards, particularly in spaces where human appearance matters (e.g., retail, hospitality, healthcare).

Several limitations constrain the generalizability of this study. The sample size (n= 15) was small and opportunistically drawn from a design school context, likely introducing field-specific bias and limiting demographic diversity. The experimental environment, while standardized, lacked the full complexity of real-world architectural conditions. To improve future studies, larger and more demographically diverse samples should be recruited. Additionally, testing could occur in contextualized environments such as lobbies or galleries, and perhaps incorporate eye-tracking or color appearance modeling to deepen the perceptual dataset.

Another limitation is that the task was difficult to complete due to the complex and quick nature of the experiment. A small retrial of about 50% (n=7) of the original participants was conducted using a revised methodology. Instead of going through the three illuminants and answering questions after the fact, participants were allowed to sit in each lighting state indefinitely until they had answered the 6 Likert ratings. After each one, they were allowed to move on to the next and so on. After all three illuminants they were shown the three lighting states in quick succession before answering the forced preference task. This procedure was completed three times for each of the CCT blocks.

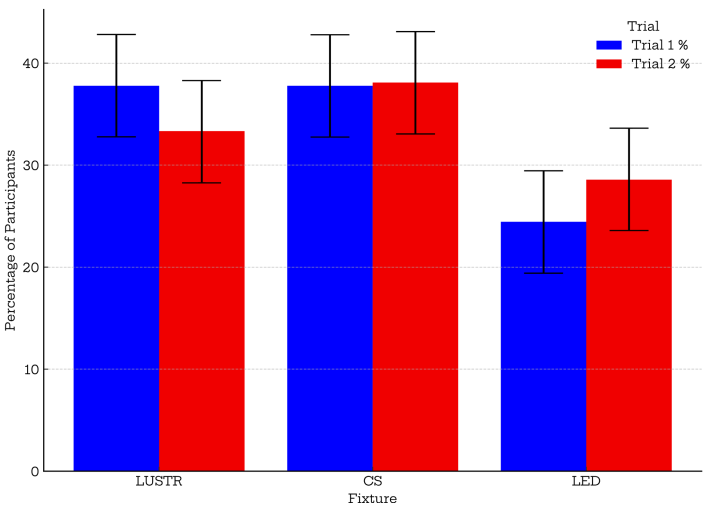

Figure 8. *Fixture preference by trial with 5% error bars.*

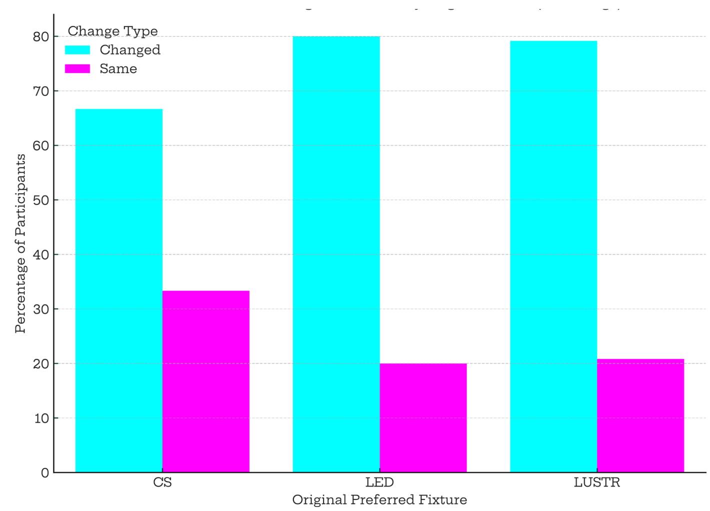

Figure 9. *Preference change by original fixture (%).*

The results of this limited replication point towards similar data points, still maintaining the overall trends on the original data set. What was observed, however, was that participants did not stay consistent with their original choices. No other aspect of the experimental setup changed, which again shows the unreliability of human perception.

Despite these constraints, the findings suggest that we should not dismiss the importance of spectral nuance. Perception is not merely about energy or other limitations in the architectural lighting design workflow, but about experience.

The relatively low consistency in forced-choice selections, with only one participant choosing the same luminaire across all three trials, suggests either a high sensitivity to environmental shifts (e.g., changes in CCT context) or a perceptual ambiguity that current SPD differences cannot fully resolve. This variability could be interpreted as a lack of meaningful perceptual difference, or alternatively, as evidence that even minor spectral shifts evoke different emotional or visual responses depending on context. Importantly, participants still expressed strong qualitative preferences in their open-ended qualitative responses, highlighting the complexity of mapping lighting perception through quantitative metrics alone.

To better align lighting practices with human perceptual needs, the industry must shift toward a dual-metric approach: one that preserves the energy and efficiency standards that dominate architectural lighting, while also embedding perceptual benchmarks derived from studies like this one. Manufacturers could develop SPD profiles that intentionally include red and violet content to mimic legacy sources like tungsten or the qualities of Golden Hour sunlight. Meanwhile, regulatory frameworks such as ASHRAE or Title 24 might benefit from the inclusion of perceptual indices, such as minimum TM-30 Rg thresholds or even user trial validations, in settings where human appearance is critical. Designers should feel empowered to specify fixtures based on visual experience, not just compliance, particularly in hospitality, healthcare, and cultural institutions where subjective impressions drive the user experience.

5.1 Conflict of Interest Disclosure

The theatrical luminaires (L1 & L2) used in this study were provided temporarily by ETC (Electronic Theatre Controls) for the purposes of experimental evaluation. ETC had no involvement in the study design, data collection, data analysis, interpretation of results, or preparation of this manuscript.

6. Conclusion

This study did not find significant perceptual differences between broader-spectrum LED luminaires like the ETC LUSTR 3 and other “worse” luminaires. Yet, some subjective experiences seem to point at the ColorSource (an RGB source) as the best at Color Quality and consistently rated the SATCO luminaire as the “worst” in most categories. These results highlight a disconnect between perceptual experience and commonly used lighting metrics such as TM-30 or CRI. As a result, designers should prioritize human-centric qualities, especially in visually sensitive environments, while regulatory bodies should consider incorporating perceptual data into performance standards. Future research should test larger, more diverse samples in real-world architectural settings to better align lighting practices with how spaces are actually seen and experienced. Further research should also consider the context when conducting studies and acknowledge the deeply subjective experience of individuals when attempting to make broad assumptions about the general population.

7. Acknowledgements

Special thanks to Dr. Craig Bernecker and Francesca Bastianini for their relentless support throughout this process. To Kyle McGahan for his challenging yet rewarding conversations. To Craig Fox and Howard Leuthold at ETC NYC for graciously providing the necessary equipment to conduct this study. To Wendy Luedtke for her invaluable insights very early on in the process. To Dr. Louise Alkabes-Esquenazi for her help in completing this massive undertaking. To Shoshanna Segal for her continued support from the very beginning.

8. References

1. Davis W, Ohno Y. 2009. Approaches to color rendering measurement. J Mod Opt. 56(13):1412–1419. https://doi.org/10.1080/09500340903023733.
2. Narendran N, Gu Y, Freyssinier JP, Yu H, Deng L. 2004. Solid-state lighting: failure analysis of white LEDs. J Cryst Growth. 268(3):449–456. https://doi.org/10.1016/j.jcrysgro.2004.04.071.
3. Mott MS, Robinson DH, Walden A, Burnette J, Rutherford AS. 2012. Illuminating the effects of dynamic lighting on student learning. SAGE Open. 2(2):2158244012445585. https://doi.org/10.1177/2158244012445585.
4. [DOE] US Department of Energy. 2020. Energy conservation program: energy conservation standards for general service fluorescent lamps and incandescent reflector lamps. Fed Regist. [accessed 2024 May 14];85(85):25326–25410. https://www.govinfo.gov/content/pkg/FR-2020-05-01/pdf/2020-08851.pdf.

5. Cho J, Park JH, Kim JK, Schubert EF. 2017. White light-emitting diodes: history, progress, and future. Laser Photonics Rev. 11(2):1600147. https://doi.org/10.1002/lpor.201600147.
6. Rutherford MD, Brainard DH. 2002. Lightness constancy: a direct test of the illumination-estimation hypothesis. Psychol Sci. 13(2):142–149. [accessed 2024 May 14]. https://www.jstor.org/stable/40063856.
7. [CIE] Commission Internationale de l'Éclairage. 1995. Method of measuring and specifying colour rendering properties of light sources. Vienna (AT): CIE Central Bureau. 16 p.
8. Penzel F. 1978. Theatre lighting before electricity. Middletown (CT): Wesleyan University Press. 179 p.
9. Bowers B. 1998. Lengthening the day: a history of lighting technology. Oxford (UK): Oxford University Press. 221 p. [accessed 2024 May 14]. http://www.loc.gov/catdir/enhancements/fy0637/97048433-d.html.
10. Viola AU, James LM, Schlangen LJ, Dijk DJ. 2008. Blue-enriched white light in the workplace improves self-reported alertness, performance and sleep quality. Scand J Work Environ Health. 34(4):297–306. https://doi.org/10.5271/sjweh.1268.
11. Rea MS, Freyssinier JP. 2010. Color rendering: beyond pride and prejudice. Color Res Appl. 35(6):401–409. https://doi.org/10.1002/col.20562.
12. Ohno Y. 2005. Spectral design considerations for white LED color rendering. Opt Eng. 44(11):111302. https://doi.org/10.1117/1.2130694.

13. David A, Fini PT, Houser KW, Royer MP, Whitehead L, Wei M, Smet KA. 2015. Development of the IES method for evaluating the color rendition of light sources. Opt Express. 23(12):15888–15906. https://doi.org/10.1364/OE.23.015888.
14. Guarnieri M. 2018. An historical survey on light technologies. IEEE Access. 6:25881–25897. https://doi.org/10.1109/ACCESS.2018.2834432.
15. [CIE] Commission Internationale de l'Éclairage. 2017. CIE position statement on colour quality metrics. 2nd ed. Vienna (AT): CIE. https://doi.org/10.25039/ps.6qdfenyc.
16. Royer MP, Wilkerson A, Wei M, Houser K, Davis R. 2017. Human perceptions of colour rendition vary with average fidelity, average gamut, and gamut shape. Light Res Technol. 49(8):966–991. https://doi.org/10.1177/1477153516663615.
17. Tabaka P. 2021. Influence of replacement of sodium lamps in park luminaires with LED sources of different closest color temperature on the effect of light pollution and energy efficiency. Energies. 14(19):6383. https://doi.org/10.3390/en14196383.
18. [ETC] ETC Electronic Theatre Controls. Source Four LED Series 3 features. Middleton (WI): ETC; [accessed 2024 May 14]. https://www.etcconnect.com/Products/Entertainment-Fixtures/Source-Four-LED-Series-3/Features.aspx.
19. Jones BC, Little AC, Burt DM, Perrett DI. 2004. When facial attractiveness is only skin deep. Perception. 33(5):569–576. https://doi.org/10.1068/p3463.

20. Boyce PR, Eklund NH, Hamilton BJ, Bruno LD. 2000. Perceptions of safety at night in different lighting conditions. Light Res Technol. 32(2):79–91. https://doi.org/10.1177/096032710003200205.
21. Wikimedia Commons. 2029. File: Vassily Kandinsky, 1910 - Sketch for Composition II [Internet]. Wikimedia Foundation; [accessed 2024 May 14]. https://commons.wikimedia.org/w/index.php?title=File:Kandinsky_-_Sketch_for_%E2%80%9CComposition_II%E2%80%9D,_1909%E2%80%9310.jpg&oldid=991946700