U.S. DEPARTMENT OF ENERGY | Energy Efficiency & Renewable Energy

CALiPER

Report 20.1:

Subjective Evaluation of Beam Quality, Shadow Quality, and Color Quality for LED PAR38 Lamps

October 2013

Prepared for:

Solid-State Lighting Program
Building Technologies Office
Office of Energy Efficiency and Renewable Energy
U.S. Department of Energy

Prepared by:

Pacific Northwest National Laboratory

PACIFIC NORTHWEST NATIONAL LABORATORY
operated by
BATTELLE
for the
UNITED STATES DEPARTMENT OF ENERGY
under Contract DE-AC05-76RL01830

Printed in the United States of America

Available to DOE and DOE contractors from the
Office of Scientific and Technical Information,
P.O. Box 62, Oak Ridge, TN 37831-0062;
ph: (865) 576-8401
fax: (865) 576-5728
email: reports@adonis.osti.gov

Available to the public from the National Technical Information Service
5301 Shawnee Rd., Alexandria, VA 22312
ph: (800) 553-NTIS (6847)
email: orders@ntis.gov <http://www.ntis.gov/about/form.aspx>
Online ordering: http://www.ntis.gov

This document was printed on recycled paper.

(8/2010)

Outline of CALiPER Reports on PAR38 Lamps

Application Summary Report 20: LED PAR38 Lamps (November 2012)[1]
An initial sample of 38 LED PAR38 lamps, as well as 5 halogen and CFL benchmarks, underwent photometric testing according to IES LM-79-08. CALiPER Application Summary Report 20, which also includes an addendum covering 6 additional LED products and 3 additional benchmarks that were similarly tested, focuses on the basic performance characteristics of the LED lamps compared to the benchmarks, as well as performance relative to manufacturers' claims. This report follows numerous similar reports on different product types that have been published by the CALiPER program.

Following the initial CALiPER report on LED PAR38 lamps, several additional special investigations were initiated. The results of these investigations are divided into four reports, each of which includes new data for all or a subset of the products described in Application Summary Report 20. Information on basic performance characteristics—such as efficacy, lumen output, and beam angle—is generally not repeated in any of the additional reports.

Report 20.1: Subjective Evaluation of Beam Quality, Shadow Quality, and Color Quality for LED PAR38 Lamps
This report focuses on human-evaluated characteristics, including beam quality, shadow quality, and color quality. Using a questionnaire that included rank-ordering, opinions on 26 of the Report 20 PAR38 lamps were gathered during a demonstration event for members of the local Illuminating Engineering Society (IES) chapter. This was not a rigorous scientific experiment, and the data should not be extrapolated beyond the scope of the demonstration. The results suggest that many of the LED products compared favorably to halogen PAR38 benchmarks in all attributes considered. LED lamps using a single-emitter design were generally preferred for their beam quality and shadow quality, and the IES members' ranking of color quality did not always match the rank-order according to the color rendering index (CRI).

Report 20.2: Flicker, Dimming, and Power Quality Characteristics of LED PAR38 Lamps (pending)
Dimming curves and flicker waveforms, measured at the Pacific Northwest National Laboratory (PNNL), will be presented for all of the PAR38 lamps included in Application Summary Report 20.

Report 20.3: Stress Testing of LED PAR38 Lamps (pending)
A small sample of each of the Application Summary Report 20 PAR38 lamp types is currently undergoing stress testing that includes substantial temperature and humidity changes, electrical variation, and vibration. The results will not directly address expected lifetime, but can be compared with one another, as well as with benchmark conventional products, to assess the relative robustness of the product designs.

Report 20.4: Lumen and Chromaticity Maintenance of LED PAR38 Lamps (pending)
The lumen depreciation and color shift of 40 lamps is currently being monitored for an extended period of time using the Lumen Maintenance Test Apparatus (LMTA) at PNNL.

[1] Available at: http://apps1.eere.energy.gov/buildings/publications/pdfs/ssl/caliper_20_summary.pdf

1 Introduction

The occupants of an architectural space cannot sense the efficacy or power factor of the light sources, but they are likely to be aware of the quality of light in the space. Lighting quality is often about choosing the right product for the application, and in many cases that includes minimizing the effects of product shortcomings. Some of a lamp's limitations may be trivial and easily averted, but identifying areas for potential improvement is important to the continued development of energy-efficient light sources, such as LEDs. Even conventional lighting products, which are often considered the standard for comparison, have limitations that should not be replicated in integral LED lamps or dedicated LED luminaires. In fact, reducing or eliminating any quality limitations could lead to increased adoption of energy-efficient technologies. In other words, energy-efficiency upgrades may be more likely to occur if they are coupled with quality improvements, rather than trading off lighting quality for long-term cost savings.

Testing results from the initial CALiPER report on PAR38 lamps showed that the LED lamps had much higher efficacies than conventional sources, while offering comparable lumen output and luminous intensity distributions—at least in the target range of lamps comparable in output to 75 W halogen PAR38s. Most of the LED PAR38 lamps had CCTs that were similar to the conventional lamps they were intended to replace and CRIs that were good (80s) or excellent (90s). By the basic numbers, many LED PAR38 lamps are a good alternative to halogen lamps, but focusing only on standard data from photometric testing does not reveal a lamp's performance in other areas, such as the smoothness or color consistency of the beam. Unfortunately, many of these facets of lighting quality are not easily quantified using readily available metrics. Thus, it can be difficult for manufacturers to make improvements and for consumers to differentiate products. Subjective, comparative evaluations can help to identify product attributes that lead to better lighting quality, as well as what attributes specifiers consider desirable.

PAR38 Demonstration

In order to learn more about lighting preferences, as well as to gain a greater understanding of the difference in performance for a set of LED and halogen PAR38 lamps, a demonstration mock-up was constructed at PNNL's demonstration facility in Fairview, Oregon. Members of the local IES chapter were invited to view the 26 different lamps being demonstrated, and were asked to complete a questionnaire. The questionnaire focused on beam quality, shadow quality, and color quality.

This report focuses on the outcomes of the observers' evaluations. Although it was not a rigorous scientific experiment, the resulting feedback can be used to identify areas of concern related to lighting quality aspects of LED PAR38 lamps. Specifically, the subjective evaluations have been used to identify the relative performance of the LED lamps compared to benchmark halogen lamps, as well as to determine which features of LED products may lead to more desirable lighting quality.

2 Methods

Demonstration Space

In June 2013, a space for demonstrating PAR38 lamps was constructed at the PNNL demonstration facility. The space included a series of lamp holders mounted on an adjustable truss system (Figure 1). A total of 26 lamps were installed and aimed at a wall covered with off-white paint typical of a basic commercial space. The lamps were aimed at pushpins placed in the wall, which were spaced approximately 10 inches on center to match the spacing of the lamps. Visual evaluation ensured that each lamp was aimed at the associated pushpin, with the lamp only tilted on one axis.

Evaluation Types

For the demonstration, three performance attributes were considered: beam quality, shadow quality, and color quality. Each performance attribute had a different demonstration configuration: beam quality was evaluated with lamps aimed at a blank wall; shadow quality was evaluated using a series of synthetic flowers positioned in a vase about 12 inches from the wall; and color quality was evaluated using an X-Rite Color Checker Classic Card, fabric samples, and a board with Rosco Scenic Paint (Figure 2). Shadow quality and color quality were evaluated

Figure 1. The PAR38 lamps were installed on a truss system at the PNNL demonstration facility in Fairview, Oregon.

Figure 2. The color quality demonstration setup. The color samples included fabric swatches, Rosco Scenic Paint, and X-Rite Color Checker Classic cards. Participants could not position themselves within the beam; that is, they could not use their skin as a reference to aid evaluation. The setup for the other evaluations was similar, but did not include the color samples.

each using just one set of six lamps, whereas beam quality was evaluated using three distinct sets of lamps that were separated based on their nominal beam sizes. In total, there were four groups of lamps, with one group being used for both beam quality and shadow quality. To minimize bias, the evaluators were shielded from having a direct view of the lamps, and were instructed not to discuss their opinions or observations with others.

Lamps

The four groups of lamps, which were evaluated separately, included:

- Six narrow flood lamps—five LED and one halogen—that were specifically chosen for their CRI, CCT, and D_{uv} and adjusted (mechanically or electrically) so that the illuminance at the center of the beam was approximately equal. The manufacturers' listed beam angles were between 25° and 30°, with measured beam angles ranging between 20° and 30°. This was the only set of lamps used to evaluate color quality.

- Six narrow flood lamps—five LED and one halogen—that were listed by the manufacturer as having a 25° beam angle. The measured beam angles ranged from 20° to 29° (all within American National Standards Institute [ANSI] tolerances for a 25° nominal beam angle), whereas output ranged between 848 and 1,157 lm, with center beam candle power (CBCP) values of 2,753 to 6,945 cd. The group included LED lamps with one emitter, multiple emitters, and diffusing elements. This group of lamps was used to evaluate beam quality and shadow quality during two separate assessments.

- Eight "spot" lamps; that is, lamps that were given a spot designation by the manufacturer. Five of the lamps were LED-based, whereas three were halogen—including one 130 V lamp.[2] The measured beam angles ranged from 8° to 18° (the manufacturers listed beam angles between 8° and 17°), with CBCPs between 7,609 and 19,858 cd. Output ranged from 737 to 1,088 lm, the approximate range for 75 W halogen PAR38 spot lamps. These lamps were only evaluated in terms of beam quality.

- Six "flood" lamps; that is, lamps that were given a flood designation by the manufacturer. Five of the lamps were LED-based, and one was a halogen product. The measured beam angles ranged between 39° and 57°, whereas the manufacturers' listed beam angles were between 40° and 55°. Output ranged between 781 and 977 lm, with CBCPs of 655 to 1,779 cd. These lamps were only evaluated in terms of beam quality.

Complete performance data for each lamp is available in Appendix A. Every effort was made to reduce differences in lamps unrelated to the characteristic being evaluated for each group—although some differences could not be avoided, given the need to use only commercially available lamps. All of the lamps were purchased seven to eight months before the demonstration event[3] and in some cases may no longer be sold—or may have been upgraded. The delay was necessary to perform photometric testing, build the apparatus, develop the protocol, and arrange for the evaluation. Although the lamps were several months old, they generally represented what was (and still is) currently available.

Control System

The four groups of lamps were individually controlled by EnOcean®-based equipment provided Leviton. A wireless, plug-in LevNet dimmer/relay receiver—which could provide both on-off and dimming control—was installed in-line with each lamp holder. During commissioning, each receiver was programmed to respond to one or more control signal sources.

The evaluation process called for lamps to be viewed in their respective groups, as well as individually. Groups were created by programming multiple LevNet receivers to respond to a single, wireless, self-powered, remote switch. Viewing individual lamps was facilitated by programming each receiver to respond to unique commands sent from a laptop computer enabled with a LevNet RF USB Computer Link, and coding the desired sequence and timing of the *on* and *off* commands in DolphinView.

The color quality evaluation required some of the lamps to be dimmed, so that all of the lamps provided equal illuminance at the center of the display. After the required dimming level for each lamp was determined, a command sequence was created in DolphinView to individually turn on each lamp and then dim it to an appropriate level—all occurring prior to the evaluation participants' arrival. Achieving precise, repeatable dimming levels (and target illuminances) was somewhat challenging with this control equipment, as the dimming control signal is not created by specifying and sending a numeric value, but rather by sending two signals separated by a specified amount of time. The time separating the two signals effectively defines the dimmed level, with longer times providing lower dimmed levels. The accuracy of this approach is limited by the

[2] When operated at standard voltages (nominally 120 V), 130 V halogen lamps offer extended life, but also consume less energy, have a slightly lower CCT, and emit fewer lumens than their rated values; basically, they function in a dimmed state.
[3] Procurement followed the standard CALiPER procedure of anonymous purchases from typical lighting vendors.

time resolution of the control system, while the precision of this approach is limited by the stability of the wireless communication network latency.

Evaluation Methods

In total, the demonstration included five separate evaluations: three for beam quality, one for color quality, and one for shadow quality. For each characteristic, the participants were asked to ignore differences not directly related to that attribute.

For the beam quality evaluations, feedback was requested in three stages. First, the participants were asked to rank-order the lamps from least desirable to most desirable, with the results later converted to a numerical scale; for example, a rank-value of six indicates the least favorable product in the narrow flood group, where as a rank-value of one indicates the most desirable product in the group. Next, the participants were asked to indicate which lamps they considered "unacceptable," "acceptable," or "outstanding," using check boxes below the ranks. Finally, blank spaces were left for the participants to indicate what aspects of the beam quality most influenced their ranking, both positively and negatively, using a numbered list of attributes. The attributes included hardness or softness of the beam edge, falloff/gradient from the center to the edge of the beam, brightness pattern/consistency, color pattern/consistency, center beam intensity, and stray light or lack thereof outside the main beam. It was also possible to provide feedback on other areas of beam quality that were not listed. A sample response form is provided in Appendix B.

For shadow quality, the participants were asked to "rank-order the lamps based on the quality of the shadow projected on the wall behind the objects." Similarly, for color quality the visitors were asked to "rank-order the lamps based on color appearance of the lighted color sample boards for a retail application." Neither of these evaluations asked the participants to provide more feedback about specific attributes affecting their judgment, although space was provided to allow participants to indicate if the performance was unacceptable, acceptable, or outstanding.

Participants and Evaluation Procedure

Two groups of visitors were brought to the demonstration facility in July and August 2013, totaling 21 participants (8 in one group, 13 in the other). The participants were not screened for color vision deficiencies or any other conditions, and no data on gender or other identifying attributes were collected. When visitors arrived, they were provided with a written set of instructions and asked to provide some basic information, then given an oral introduction to the activity. Next, they were taken as a group to the demonstration room, where they proceeded to evaluate the five groups of lamps (color quality, beam quality spot, beam quality narrow flood, beam quality flood, and shadow quality). Following the demonstration, there was time for discussion and review of all the lamps.

For the color quality and shadow quality demonstration groups, the lamps were all on at one time for the entirety of the evaluation period. For each of the beam quality groups, all the lamps were turned on at the same time for approximately one minute, then turned on one at a time for 30 seconds each, before again all being illuminated simultaneously. Participants were free to take notes during the entire process, and were given approximately 5–10 minutes to complete the evaluation during the final interval with all the lamps illuminated. The entire evaluation took approximately 45 minutes, not including the subsequent discussion.

3 Results and Discussion

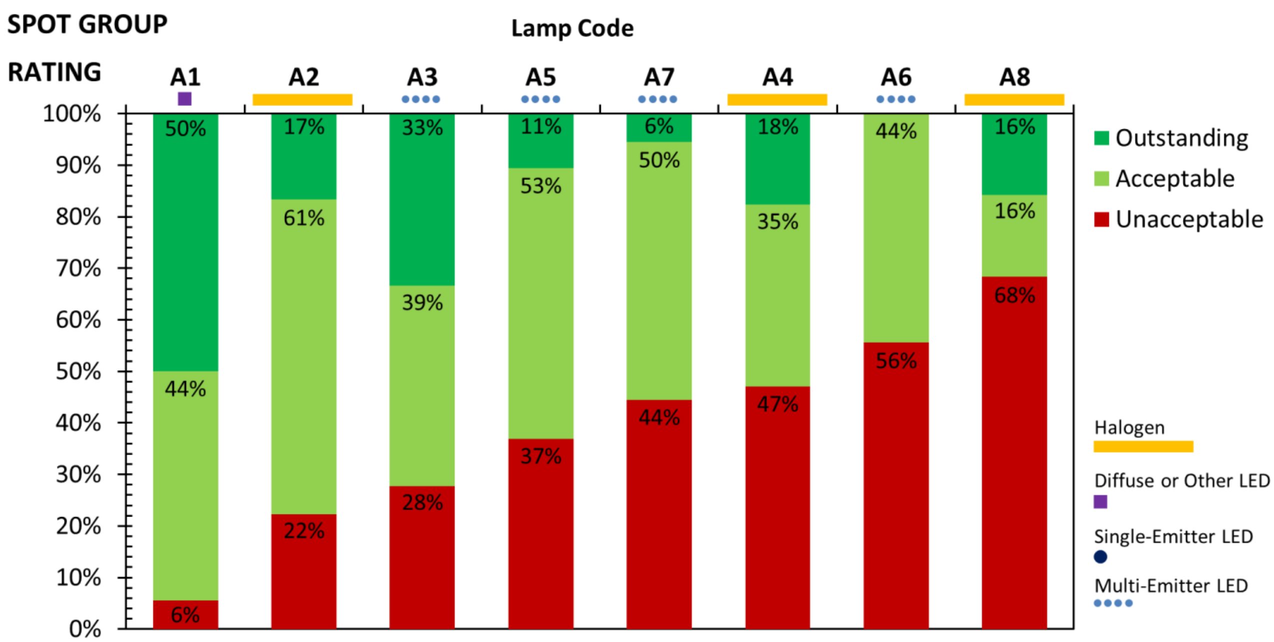
It is important to reiterate that this was not a rigorously controlled scientific experiment. Thus, no statistical inferences were made to determine the significance of different rankings and ratings. Further, the results should be treated as anecdotal, rather than extrapolated to include products outside the scope of this exercise. Nonetheless, the collective observations obtained during this demonstration offer valuable insight into the state of LED PAR38 lamps, and they can help guide continued product development.

Beam Quality

Beam quality was evaluated for three different groups of lamps, which were divided based on their distribution type. The narrow flood lamps all had the same nominal beam angle, whereas the spot and flood groups exhibited a range of beam angles. Although the evaluators were explicitly directed to ignore beam angle—as well as all other attributes not directly related to beam quality—it is likely that beam angle had some influence on participants' opinions. This is especially true for the spot distribution group, which had measured beam angles between 8° and 18°, resulting in noticeably different-sized beam patterns projected on the wall.

Spot Lamps

The spot lamps had the least acceptable beam quality of the three distribution types evaluated; on average, 38% of the given ratings were in the unacceptable category (34% for the LED lamps only), as shown in Figure 3.



RANKING

	A1	A2	A3	A5	A7	A4	A6	A8
Mean	2.5	4.2	3.1	4.4	4.9	4.5	5.6	6.7
Mode	1	3	1	5	3	4	8	8
Min	1	2	1	1	2	1	2	3
Max	8	7	7	7	8	7	8	8

Figure 3. Summary of results for the spot lamp group beam quality evaluation. Whether arranged by rating or ranking, the results are very similar. Lamp type A3 had a better mean and mode rank than lamp type A2, but also a higher number of unacceptable ratings. (Lower ranks were better; a rank of one was the most favorable, whereas a rank of eight was the least favorable.)

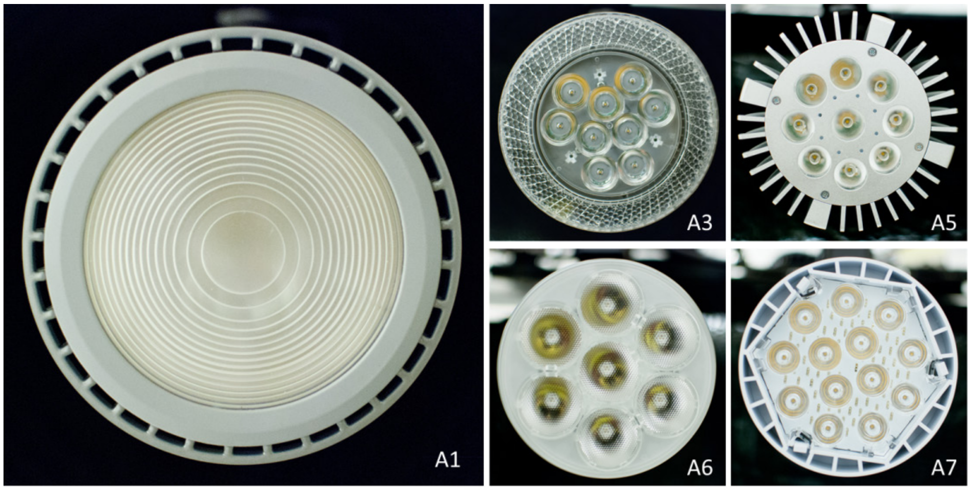


Figure 4. Photographs of the lamps in the spot group. The optical system for lamp type A1 was very different from the other LED lamps in the group, and the beam quality was ranked as the best. Notably, it had the widest beam angle of any product in the group, and narrower distributions may not be possible using this relatively diffuse optical system.

Furthermore, only one of the eight lamps (A1) had less than 20% of the given ratings in the unacceptable category. Undoubtedly, the least favorite attribute of these lamps was stray light outside the beam, which was noted at least once for each of the lamps and made up 35% of the total listings for undesirable attributes (Appendix C), regardless of the source type. The second least favorable attribute was color pattern/consistency, which made up 21% of the noted undesirable characteristics—again, this percentage was nearly the same whether considering all of the lamps or just the LED lamps. The most noted favorable attributes were generally evenly divided between the beam edge, falloff, brightness pattern, color pattern, and CBCP.

The one lamp that stood out was A1, which received acceptable or outstanding ratings from 94% of the observers and had the best mean rank in the group (complete ranking data is available in Appendix D). This product utilized a diffusing lens, unlike the other LED lamps that used multiple, undiffused emitters (Figure 4; photographs of the optical systems for all lamps are shown in Appendix E). This resulted in a noticeably "cleaner" pattern of light on the wall, with a soft but well-defined edge. It is noteworthy that this product had the largest measured beam angle at 18°.

Three of the eight products in the spot group were halogen lamps, and they placed third (A2), fifth (A5), and eighth (A8) in the group based on mean ranking, or second (A2), sixth (A5) and eighth (A8) based on unacceptability rating.[4] Their combined acceptability ratings[5] ranged from 32% to 78%, indicating substantial differences in opinion for a seemingly homogeneous and fully developed technology. As shown in Figure 5,

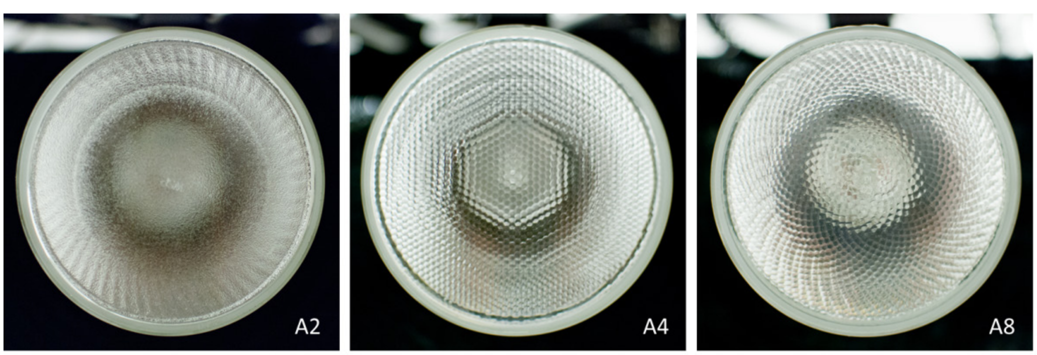


Figure 5. The lenses of the halogen lamps in the spot group were different, leading to widely different rankings. The highest-ranked product, A2, had only a slightly diffusing lens, whereas the other two lamps had more noticeable dimple patterns.

[4] For all but the spot lamp beam quality evaluation, the order of preference was the same based on mean rank or unacceptability rating. For the spot lamp group, there were small juxtapositions of two pairs of lamps.

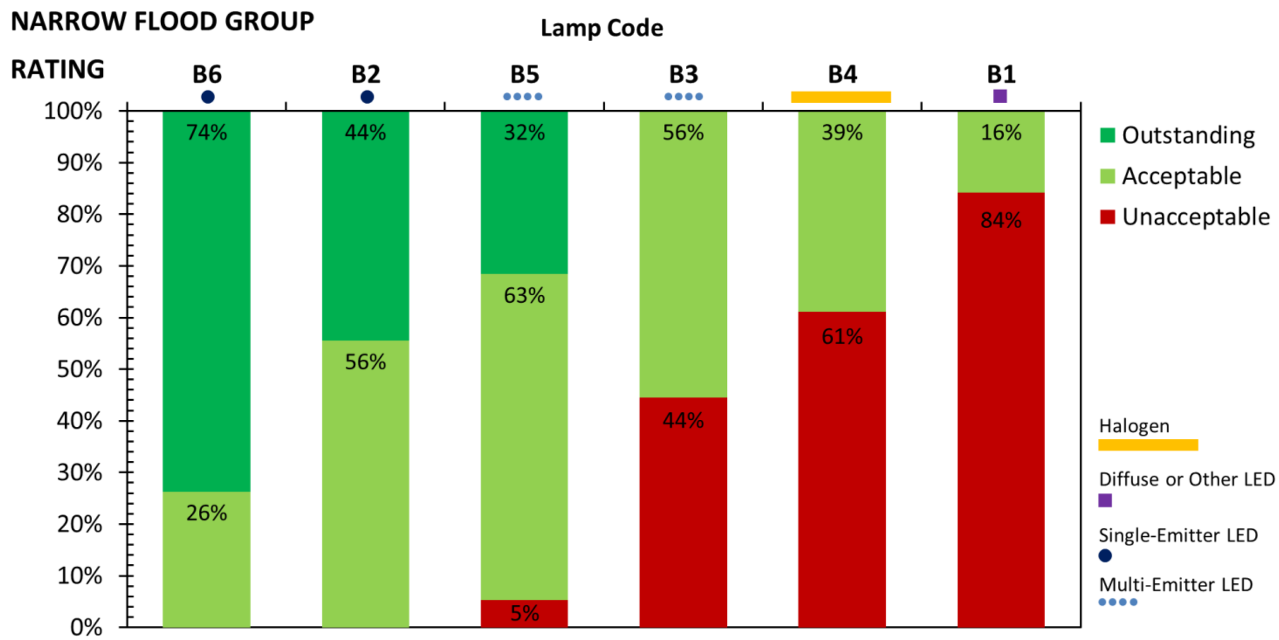
[5] Combined acceptability ratings combine the *acceptable* and *outstanding* categories, providing an overall sense of the suitability of the product.

however, the lamps use noticeably different optical systems. Lamp A2, which received substantially more favorable ratings than the other two products, had a slightly diffuse clear lens rather than a dimpled lens; this resulted in a smoother beam with less stray light, and correspondingly fewer notations of those attributes being unfavorable than the other two halogen lamps in the group.

Narrow Flood Lamps

A much clearer picture emerged from the evaluation of beam quality for the narrow flood lamps than for the spot lamps. While the average combined acceptability rating of 67% (Figure 6; 73% for LED lamps only) was not much higher than for the spot lamps (62%; 66% for LED lamps only), there was a well-defined split in opinions regarding the narrow flood lamps. Three of the narrow flood lamps (B2, B5, and B6) received no more than 5% unacceptable ratings from the observers, whereas the other lamps received 44%, 61%, and 84% unacceptable ratings.

The top two ranked narrow flood lamps (B2 and B6) utilized single-emitter designs (Figure 7). These two lamps were ranked number one or two by a vast majority of the evaluators (65% and 85%, respectively), and were given outstanding or acceptable ratings by 100% of the observers. The other narrow flood lamp that was viewed favorably, B5, was most frequently ranked number three and carried the third-best mean ranking. It had multiple emitters, and was the only lamp with such a design in the entire study to get such high combined



RANKING

	B6	B2	B5	B3	B4	B1
Mean	1.4	1.7	3.1	4.5	4.8	5.7
Mode	1	2	3	4	5	6
Min	1	1	2	3	4	4
Max	2	3	5	6	6	6

Figure 6. Summary of results for the narrow flood lamp group beam quality evaluation. Whether arranged by rating or ranking, the order of preference was the same. The evaluators preferred single-emitter LED products and did not like the benchmark halogen lamp. (Lower ranks were better; a rank of one was the most favorable, whereas a rank of six was the least favorable.)

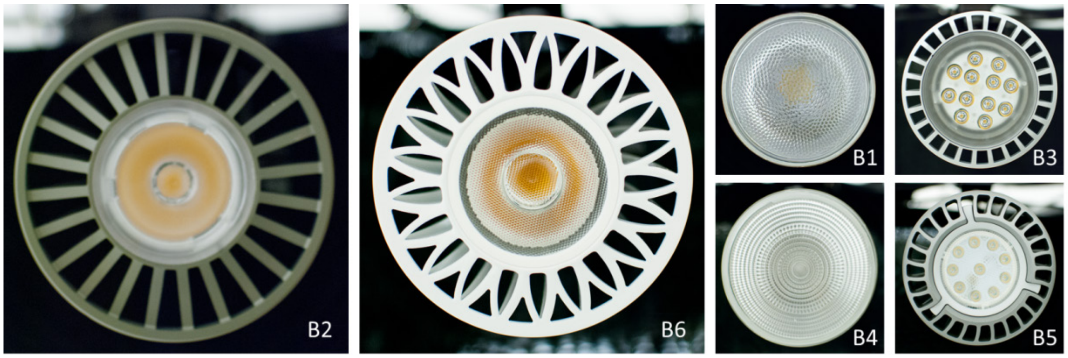


Figure 7. Photographs of the narrow flood lamps. The single-emitter designs (B2 and B6) were strongly favored in terms of beam quality.

acceptability ratings. This is perhaps because it had a slightly diffusing lens in front of the emitters, which may have reduced some of the undesirable attributes found with other multi-emitter lamps. Although all of the lamps in the group had the same *listed* beam angle (25°) and were very close in lumen output, lamp B5 had the widest *measured* beam angle (29°) and correspondingly had noticeably lower CBCP—which was generally noted as a negative quality by the observers and potentially contributed to its third-place ranking. This is further illustrated in Figure 8.

By almost any measure, lamp B6 was one of the evaluators' favorite lamps—for both beam quality and shadow quality. The only lamp in the entire evaluation without a single unacceptable rating,[6] it had the best (adjusted[7]) mean ranking, the highest percentage of outstanding ratings (74%), and no ranking lower than third within the group of six. Lamp B6 had a measured beam angle of 20°—compared to the listed beam angle of 25° for all of the narrow flood lamps (including B6)—which contributed to it having the highest CBCP of the group, despite not emitting the most lumens.[8] This extra "punch" was noted by the observers, with over half of them listing it as one of the lamp's desirable attributes—which was at least twice as much as for any other lamp in the group.

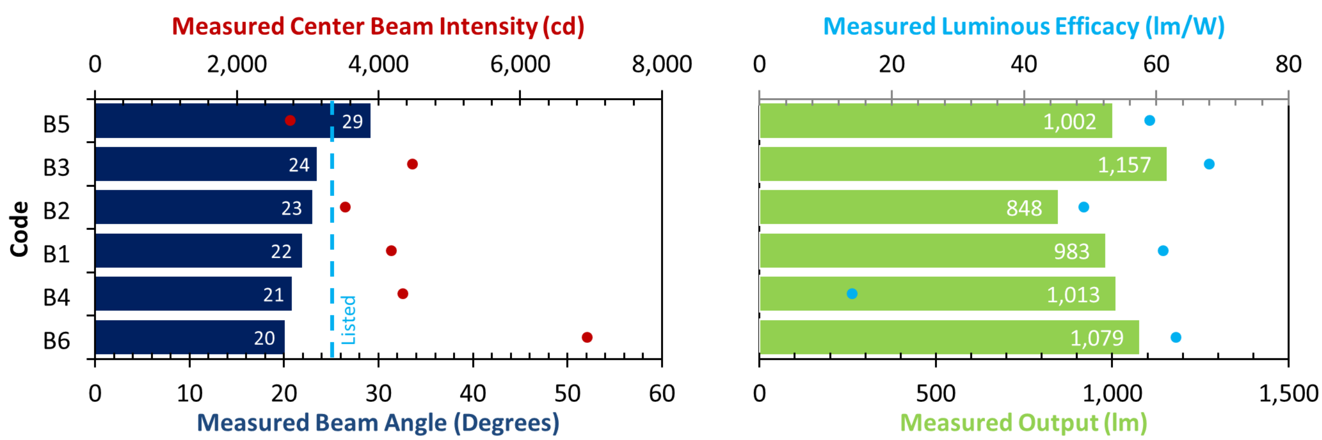


Figure 8. Beam angle (blue bars), CBCP (red dots), lumen output (green bars), and luminous efficacy (cyan dots) of the narrow flood lamps that were evaluated for beam quality. Despite all having a listed beam angle of 25°, the lamps had very different CBCPs. This was due to differences in measured beam angle, as well as differences in lumen output.

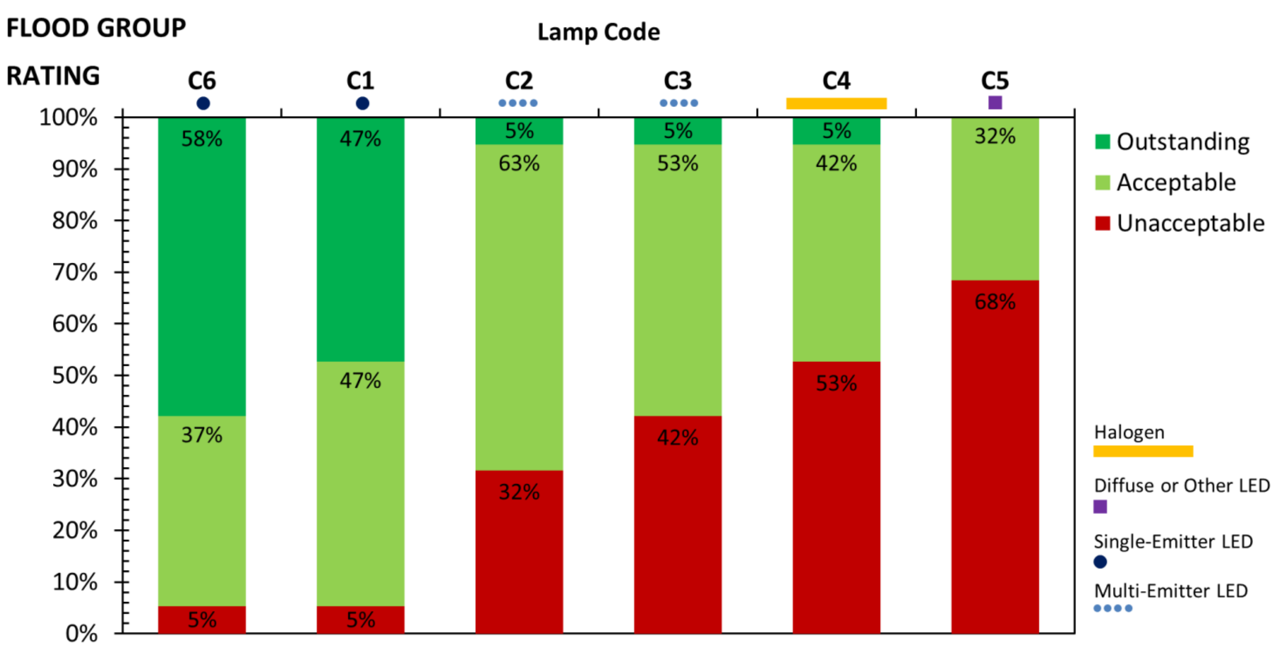
[6] Lamp B2 also had no unacceptable ratings for beam quality, but had one unacceptable rating for shadow quality. The narrow flood lamps (Group B) were the only ones that were evaluated for both beam quality and shadow quality.
[7] The mean rankings were adjusted to a common minimum-maximum scale to account for the different size of the groups. That is, the mean rankings for the eight lamps in the spot lamp group were rescaled to a one-through-six scale, rather than a one-through-eight scale. All of the other groups had six lamps.
[8] Despite having a smaller measured beam angle than the value listed by the manufacturer, lamp type B6—and all of the lamps in the narrow flood group—had measured beam angles that were within ANSI tolerances (C78.379-2006) for a nominal 25° beam angle (± 6°).

As with the spot lamps, there was no obvious attribute among the narrow flood lamps that was noted as being more desirable than the others. That is, there was little difference in the number of notations for beam edge, falloff, brightness pattern, color pattern, and CBCP—although the only other attribute, stray light or lack thereof, was noted as positive much less often. In contrast, stray light outside the beam was again a clear predictor of unacceptability; in fact, over half of the respondents listed stray light as an undesirable quality that influenced their ranking for the three lowest-ranked products.

Flood Lamps

Two lamps in the flood group, C1 and C6, were clearly favored; for both, 95% of the observers rated the beam quality as acceptable or outstanding (Figure 9). In keeping with the other results, both C1 and C6 utilize a single-emitter design. The three other LED lamps were noted by over half of the observers as having undesirable color consistency within the beam pattern, and the halogen lamp was noted by over half of the responders as having unfavorable brightness consistency (i.e., striations) within the beam. There was much less negative reaction to stray light outside the beam, which might be expected given the large beam angles.

It is not clear if the observed color inconsistency of the three multi-emitter LED lamps was a property of the lamps themselves, or was simply more visible because of the wider distribution patterns. It is possible that the wider dispersion of light and generally lower intensities resulted in the color change across the projected beams being more noticeable than for the narrow flood or spot lamps.



RANKING

	C6	C1	C2	C3	C4	C5
Mean	1.8	2.1	3.8	3.8	4.5	5.1
Mode	1	2	5	4	6	6
Min	1	1	1	2	1	3
Max	6	5	5	6	6	6

Figure 9. Summary of results for the flood lamp group beam quality evaluation. Whether arranged by rating or ranking, the order of preference was the same. The evaluators preferred single-emitter LED products. (Lower ranks were better; a rank of one was the most favorable, whereas a rank of six was the least favorable.)

Overall Discussion

Combining the beam quality evaluation results from all three groups reveals several clear trends. First, lamps using a single-emitter design were almost unanimously favored, regardless of distribution type (Figure 10). While they were applicable to different degrees for the different lamp types, the characteristics that most strongly affected unfavorable rankings were stray light outside the main beam and color consistency within the beam pattern. The features that the evaluators found favorable were more varied.

In general, the spot distribution lamps were viewed the least favorably. For several years, narrow spot distributions have been a challenge for LED products. In the initial CALiPER report on this set of PAR38 lamps, the availability of multiple lamps with a beam angle of less than 10° was considered a noteworthy achievement. However, the present evaluation shows that more work may be needed before narrow spot LED PAR38 lamps achieve their full potential. One challenge is that there are limitations to using a true single-emitter design to create a spot distribution, especially within a specific form factor. To get enough lumen output, the area of emission must be sufficiently large, which to date has required a combination of several emitters—and as the evaluation results indicate, this tends to reduce beam quality. On a related note, manufacturers must be careful when upgrading the LED packages used in directional lamps, as small differences in the size of the emitting area can change the luminous intensity distribution, unless the optical system is also adjusted.

At least two LED PAR38 lamps in each category were ranked higher than the halogen benchmarks. Although halogen lamps may be considered the gold standard by some, the participants in this demonstration were not completely satisfied with their performance. The evaluators typically did not like the striations and luminance patterns that extended beyond the main beam pattern for many of the halogen benchmarks. Given the markedly higher efficacy of LED PAR38 lamps, those with better beam quality than the halogen lamps have a multifaceted advantage for retrofit applications or new construction. Nonetheless, many factors must be addressed when choosing a lamp, including electronic compatibility with control equipment, dimming performance, form factor, cost, and others.

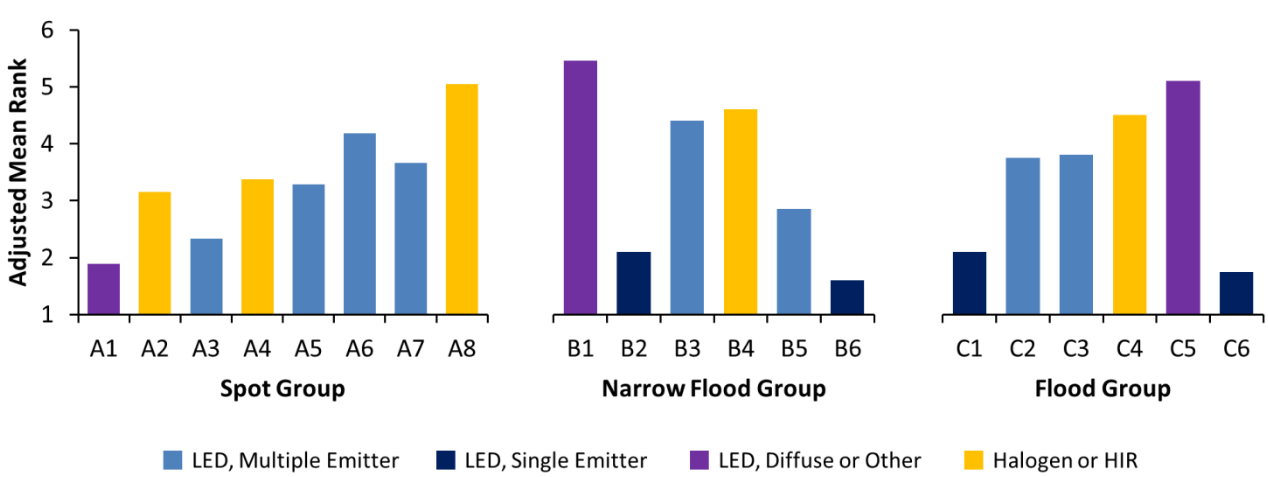


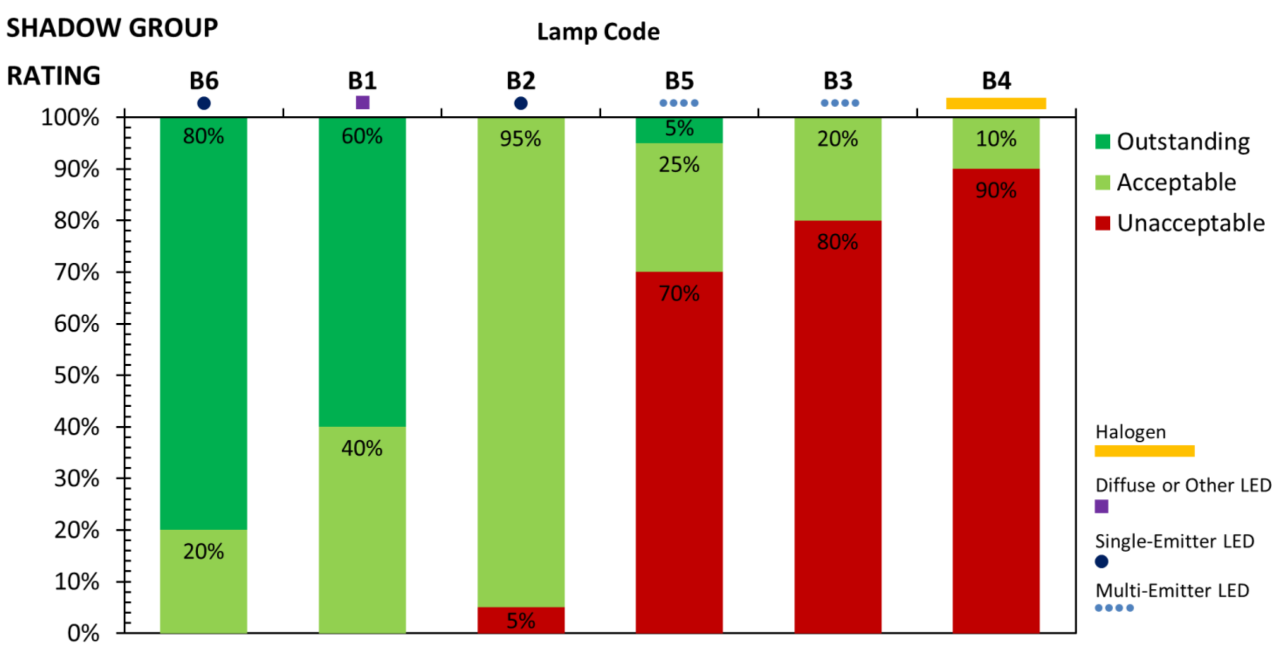
Figure 10. Mean rank (adjusted to a maximum of six for all three groups) for each product that was evaluated for beam quality. Single-emitter LED products were strongly favored (lower ranks are better). Ranks should not be compared between groups.

Shadow Quality

Separate from the evaluation of their beam quality, the Group B narrow flood lamps were ranked and rated based on their shadow quality. The lamps in this group were chosen for this evaluation because they had a variety of optical systems, including LED products with single emitters, multiple emitters, and a diffusing lens, as well as a halogen lamp.

The results were very clear, with perhaps the most agreement among evaluators of all the tasks. Lamp types B6 and B1 were ranked first or second for shadow quality by all but one person, and each had a 100% combined acceptability rating, with at least 60% of the evaluators rating the lamps' shadow quality outstanding (Figure 11). The shadow quality of lamp B2 was rated as acceptable by 95% of the participants, but nobody rated it as outstanding. Lamp types B3, B4, and B5 all had combined acceptability ratings of 30% or less.

Lamp type B6 had the best mean ranking for shadow quality, complementing its top ranking for beam quality. The other single-emitter lamp in the group (B2) ranked third, behind second-place lamp B1, which had a reflector cone and diffusing lens similar to a halogen PAR38 (Figure 8). In contrast with lamp B1, however, the halogen benchmark was ranked as producing the least-desirable shadows by a vast majority of the demonstration participants. Despite a similar design, the LED lamp (B1) likely emitted a higher proportion of its total lumens directly from the single LED package, rather than having the lumens reflected within the optical



RANKING

	B6	B1	B2	B5	B3	B4
Mean	1.4	1.7	3.1	4.5	4.8	5.7
Mode	1	2	3	4	5	6
Min	1	1	2	3	4	4
Max	2	3	5	6	6	6

Figure 11. Summary of results for the shadow evaluation. Whether arranged by rating or ranking, the order of preference was the same. The evaluators preferred single-emitter LED products. (Lower ranks were better; a rank of one was the most favorable, whereas a rank of six was the least favorable.)

system. Accordingly, the shadow for LED lamp type B1 was much more defined than the shadow created by the halogen lamp, B4 (Figure 12).

At least for this evaluation in a simulated scenario, the observers indicated a preference for crisp shadows, rather than for soft or multiple shadows. In some cases, the desire for soft or hard shadows is dependent on the specific application. Nonetheless, for most applications, single-emitter or diffused LED lamps seem to have shadow qualities that are not distracting, and are therefore more desirable. Fortunately, it appears that single-emitter designs have become more prevalent in the past year—probably because they are more feasible, thanks to the increasing efficacy and output of LED packages. Going forward, LED lamp manufacturers will have to confront the challenge of using single-emitter designs in higher-output lamps and lamps with narrow spot

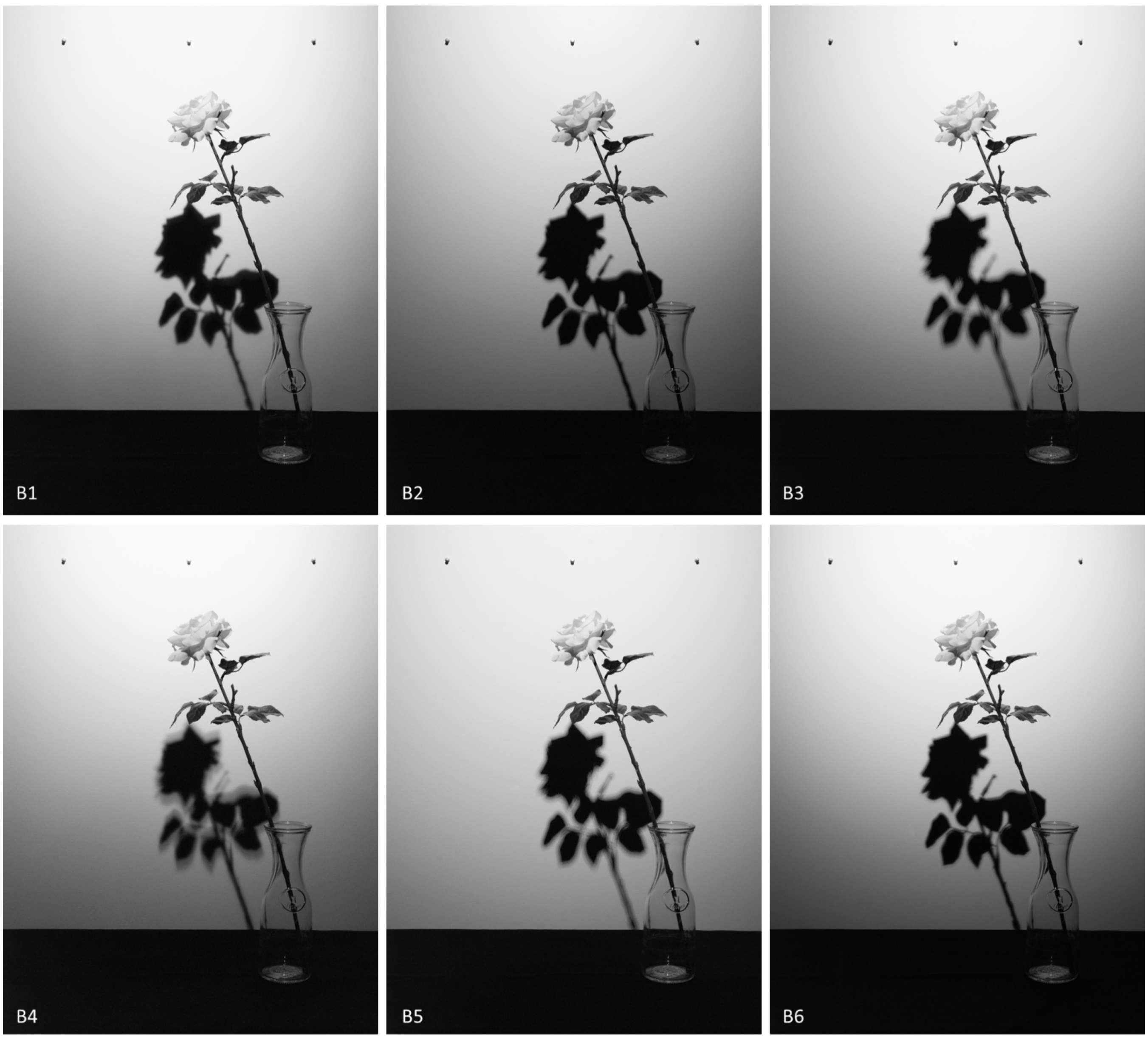


Figure 12. Photographs of the shadows cast by the six lamps for which shadow quality was evaluated. The photographs may not perfectly convey the results viewed by the study participants.

distributions. However, these issues will naturally abate if output and efficacy continue to rise, so the need for novel lamp designs may be diminished.

Color Quality

Of all the evaluations that were part of this demonstration, the participants appear to have had the most difficult time evaluating the six products in the color quality group. Unlike many scientific experiments, the spectral power distributions were not carefully tuned to test a specific hypothesis; rather, lamps with a range of different characteristics—representative of a situation that might be faced by a specifier—were simultaneously used to light identical display boards filled with a variety of color samples.

It should be reiterated that, where necessary, the lamps were adjusted to produce approximately equal illuminance at the center of the beam pattern—the halogen infrared (HIR) lamp was not dimmed, because it already had the lowest CBCP. Although the lamps had similar narrow flood distributions, there were likely illuminance differences moving away from the center of the beam, however. Because the display boards were approximately 24 inches by 36 inches, this may have had an effect on color appearance, due to the Hunt effect—whereby objects appear more colorful at a higher luminance, despite having the same chromaticity. However, varying illuminance—due to different distances between the lamp and target, different lumen output, or other factors—is another factor that is dealt with during product specification and installation. It should be explicitly noted that this was not intended to be a scientific experiment, and extrapolation of the results is inappropriate.

One tenet of CRI is that comparisons are valid only between lamps of the same CCT. In this case, however, multiple CCTs were included to simulate the real-world choices faced by specifiers. Every installation is different, and a knowledgeable specifier would likely choose a different lamp to wash a red wall than a blue wall, but in many retail applications, the target can frequently change. Thus, when color is a driving factor, lamps may need to be specified based on their overall color quality, which includes both color rendering and color temperature. This demonstration investigated how both aspects of color quality might interact to affect specifiers' choices. Importantly, the analysis considers the color quality of all the lamps together, as well as each nominal CCT group separately.

The six products in the group included two nominally 2700 K LED lamps (2743 K and 2771 K), three nominally 3000 K LED lamps (3022 K, 3033 K, 3037 K), and one HIR benchmark lamp having a typical CCT of 2842 K (see Table 1 and Figure 13). Each product had a D_{uv} between -0.001 and 0.001—a very tight tolerance—with lamp D3 the only one with a D_{uv} slightly less than zero. Within each of the nominal CCTs, there were multiple CRI levels. For the 2700 K group, the products had measured CRIs of 82 and 96, while for the 3000 K group, the products had measured CRIs of 83, 86, and 93. The values listed in this report are based on the average of two samples tested by CALiPER;[9] the lamps used in the evaluation were not one of those two samples, but there was very little (if any) difference between the tested samples, indicating consistency.

The variation in participants' ranking for each lamp was substantial: five of the six lamps received at least one most-desirable ranking, and all of the lamps had at least one rank of least or second-least desirable. Likewise, every lamp had at least one unacceptable rating, and all but one of the lamps had at least one outstanding rating. These results are shown in Figure 14. At least some of the variation may have resulted from different observers' preference for either 2700 K or 3000 K, but the spread of rankings likely demonstrates the observers'

[9] All of the measured CRIs for a pair of lamps were within one point of each other, and were generally in line with each manufacturer's listed value.

Table 1. **Color characteristics of the lamps in the color evaluation group.**

Code	D1	D2	D3	D4	D5	D6
Type	LED	LED	LED	HIR	LED	LED
x	0.4542	0.4358	0.4336	0.4502	0.4574	0.4358
y	0.4096	0.4063	0.4010	0.4110	0.4119	0.4071
u'	0.2593	0.2489	0.2497	0.2561	0.2604	0.2522
v'	0.5261	0.5221	0.5197	0.5261	0.5275	0.5232
CCT	2771	3037	3033	2842	2743	3022
D_{uv}	0.0001	0.0010	-0.0010	0.0008	0.0007	0.0003
CRI	96	93	83	99	82	86
R_9	77	59	27	96	2	28
CQS Q_g	101	99	97	97	94	98
CQS Q_a	95	93	83	97	83	86
CQS Q_f	94	93	83	97	83	86
GAI	52	56	57	52	47	56

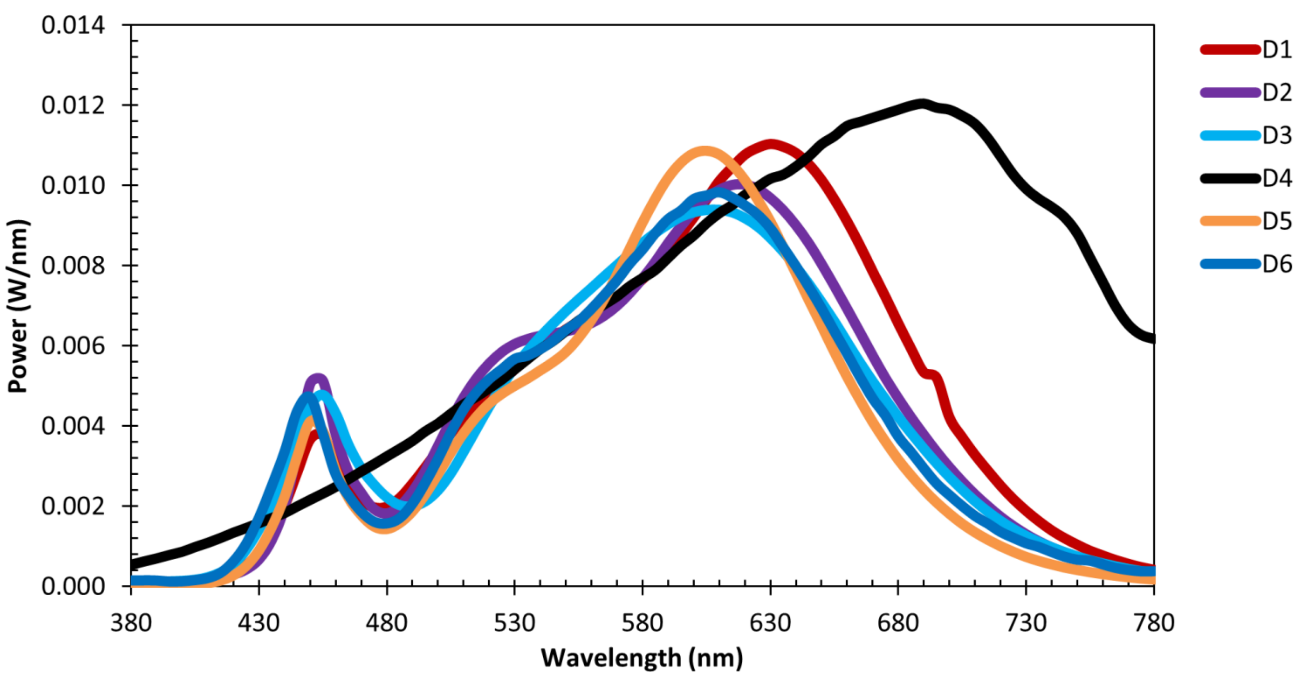
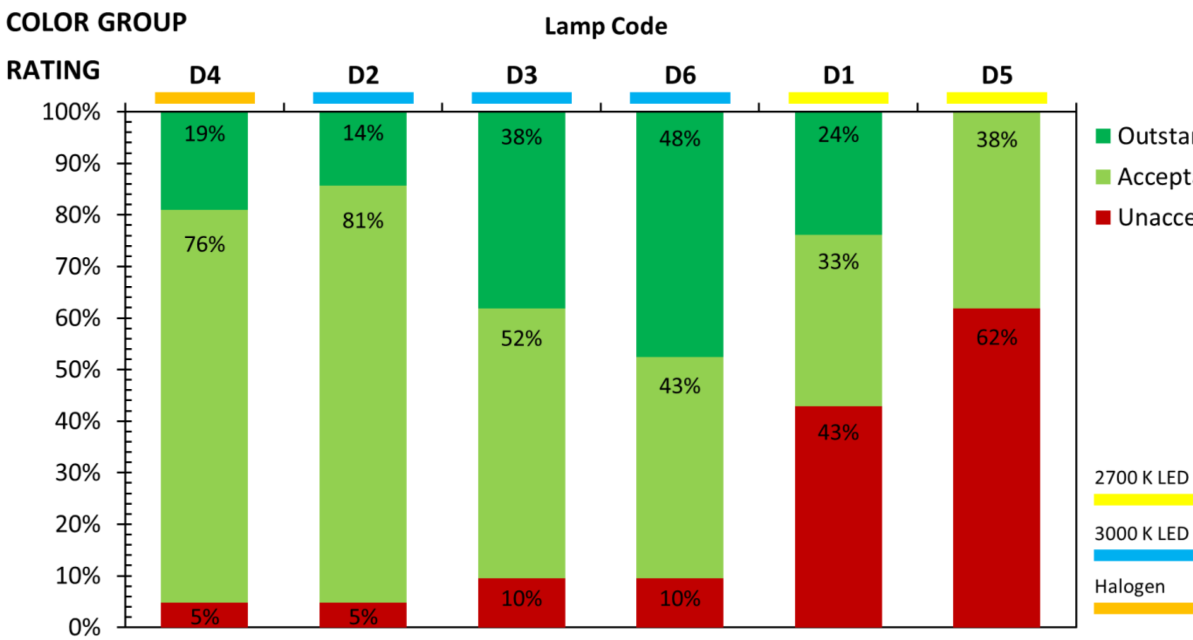


Figure 13. Spectral power distributions (SPDs) for the lamps in the color quality evaluation group. The SPDs have been normalized to equal lumen output.

difficulty, as a whole, in distinguishing between the lamps. Histograms of the observers' responses, provided in Appendix F, further illustrate a lack of consensus rankings for many of the lamps in the group.

One of the goals of this evaluation was to determine if the users would correctly rank-order lamps based on their CRI, both within and between given CCT groups. Based on mean rankings, the observers ranked the 3000 K group in *opposite* order of what CRI would suggest, considering the 3000 K, 83 CRI lamp to deliver the best color



RANKING

	D4	D2	D3	D6	D1	D5
Mean	2.9	3.5	2.7	2.7	4.1	5.3
Mode	2	4	1	1	5	6
Min	1	1	1	1	1	3
Max	5	6	5	6	6	6

Figure 14. Summary of results for the color quality evaluation. The order of preference was dependent on the evaluation method; the mean ranking, mode ranking, and acceptability rating methods all resulted in different outcomes. (Lower ranks were better; a rank of one was the most favorable, whereas a rank of six was the least favorable.)

quality (Figure 15).[10] The mean ranking for the 3000 K, 86 CRI lamp was very similar to the 83 CRI lamp, and both were slightly ahead of the mean ranking for the halogen lamp. The 3000 K, 92 CRI LED lamp had the fourth-worst mean ranking in the group, followed by the 2700 K, 96 CRI lamp and the 2700 K, 82 CRI lamp. Unlike the 3000 K lamps, the two 2700 K lamps were rank-ordered as would be predicted based on their CRI.

In contrast with the order of preference according to the mean ranking, the order of preference based on combined acceptability score was more in line with what CRI would predict—at least within each nominal CCT group. Based on that methodology, the halogen lamp was the most preferred, followed by the 3000 K LED lamps with CRIs of 96, 86, and 83 (the latter two tied), and finally the 2700 K, 96 CRI lamp and the 2700 K, 82 CRI lamp (Figure 14). However, what may be most striking about the results, regardless of the method used for ordering, is the clear separation of the nominal CCT groups. As a group, the observers displayed a strong preference for 3000 K over 2700 K, regardless of CRI—although some individuals exhibited the opposite preference. The strong effect of CCT on color appearance may have overshadowed the effect of different CRIs. In a real installation, chromatic adaptation would likely reduce this effect, but for this demonstration, chromatic adaptation was mixed.

[10] The differences in mean ranking may not be statistically significant. No statistical inference was made, given the nature of the data collection.

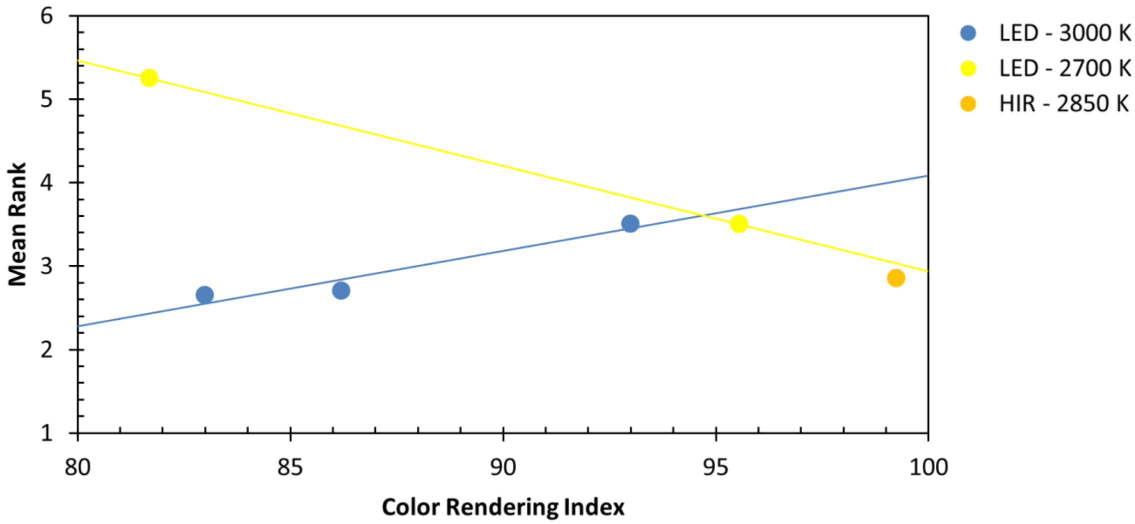


Figure 15. Mean rank versus CRI. Although some of the differences may be insignificant, the rank-order determined by the observers for the 3000 K lamps did not match what the CRI metric indicates. (Lower ranks were better; a rank of one was the most favorable, whereas a rank of six was the least favorable.)

Table 1 shows some additional color rendering metrics for the lamps in the color quality group, including the Color Quality Scale (CQS) and the Gamut Area Index (GAI). Based on rank-order alone—rather than on correlation—the GAI metric was somewhat more effective at predicting how the observers ranked the products than was CRI, indicating that color gamut may have been an important factor for the participants. However, no such relationship was found between the CQS index that measures color gamut (Q_g) and the observed data. It is possible that the relationship between GAI and the rank-order data occurs because GAI tends to penalize lower-CCT sources, and the observers also ranked the 2700 K sources worse, on average. This small dataset is not sufficient to draw widespread conclusions about the validity of any given color metric, but it does provide one anecdotal example of the imperfect nature of CRI and the general difficulty in quantifying human color preference.

There are several other notable observations based on the color quality evaluation:

- The CCT of the halogen lamp (2842 K, which is typical of halogen PAR38 lamps) is essentially at the boundary between the ANSI bins for 2700 K and 3000 K LED products. This could result in LEDs in neither category being effective in applications where halogen lamps are also used.
- The halogen lamp, having a CRI of 99, was the third-best-ranked lamp, on average, but also had the fewest unacceptable ratings. Both of the higher-ranked LED lamps had a CCT of approximately 3000 K and a CRI in the mid-80s, which are common CCT and CRI values for LED lamps. These lamps had marginally higher unacceptability ratings, but many more observers felt the color quality was outstanding.
- The LED lamps with a CCT of 2700 K were the lowest-ranked products, on average. Of those two lamps, the one with a CRI of 82 was consistently rated as having unacceptable color quality. Although CRI and R_9 are highly correlated for phosphor-coated white LEDs, this was the only lamp in the group with an R_9 value less than 20. The 2700 K LED lamp with a CRI of 96 and R_9 of 77—both the highest of any of the

LED products—had at least five rankings in the top two and was rated as having outstanding color quality by five observers.

- The only lamp with a negative value for D_{uv} had the best average rank, although all of the lamps were very close to the black body locus.

The findings on color quality are very anecdotal, but they illustrate one of the major challenges with LEDs: in many cases, visual evaluation is the best tool for judging color quality. Existing color-rendering metrics, such as CRI, may be insufficient, and using such metrics in standards could actually be detrimental to lighting quality. This evaluation is not the first to result in the same conclusion, and the CIE (International Commission on Illumination) and IES are both working on a solution to the color-rendering-metric problem.

Everything else being equal, a higher CRI may still indicate better color quality, but such clear-cut scenarios are rarely the case when specifying commercially available products. As the results of this evaluation indicate, the effect of CCT and D_{uv} on color appearance should not be overlooked in favor of choosing a product based solely on CRI—especially when the CRI is above a minimum threshold (e.g., 80). Going forward, the adoption of new color-quality metrics will be important for LED products. It should help manufacturers design better products and specifiers choose better products.

4 Conclusions

After basic testing according to LM-79, CALiPER concluded that many available LED PAR38 lamps outperformed halogen and other types of PAR38 lamps—just by the photometric testing results. To extend this conclusion, CALiPER orchestrated a demonstration of PAR38 lamps where lighting professionals evaluated LED and halogen PAR38 lamps based on their beam quality, color quality, and shadow quality. The results of the evaluation can be summarized as follows:

- In each of the quality categories, at least one LED lamp was rated more favorably than the benchmark halogen lamp(s). Halogen lamps should not always be considered the ideal source for lighting quality.
- Single-emitter LED lamps were favored in terms of both beam quality and shadow quality.
- Poor color consistency within the beam, and stray light outside the main beam pattern, were the attributes most likely to be noted by the observers as negative features.
- LED lamps with narrow spot distributions were generally viewed as having less-acceptable beam quality than their narrow-flood or flood counterparts, although there was substantial variation in perceived quality within any of the groups.
- For color quality, the observers generally preferred 3000 K LED lamps over 2700 K LED lamps, but their ranking of color quality did not always correlate with the CRI of the lamps, or with any other color rendering metric. Based on the wide variation in rankings, the observers likely had a difficult time distinguishing between the color rendering capability of the various lamps, despite differences in performance according to established color quality metrics.

Although many of the LED PAR38 lamps were ranked higher than the halogen benchmarks, there is room for improvement for even the best-performing lamps. Furthermore, there remains substantial variability among LED PAR38 lamps—as was also seen among halogen lamps—which will require buyers to make careful purchases.

Appendix A: Lamp Properties

Table A1. Identifying information for the lamps included in the CALiPER PAR38 Demonstration Assessment.

Code	Group	CALiPER ID	Type	Brand	Model
A1	Spot	12-75	LED	GE Lighting	LED17P38S830/17
A2	Spot	12-143	Hal.	OSRAM SYLVANIA	75PAR/CAP/SPL/SP9 130V
A3	Spot	12-95	LED	MSI Solid State Lighting	IPAR3830101D
A4	Spot	12-71	Hal.	Philips	75PAR38/HAL/SP10
A5	Spot	12-102	LED	LED Waves	LW10-NYC-008-WW-DM
A6	Spot	12-67	LED	Cree LED Lighting	LRP38-10L-30K-12D
A7	Spot	12-96	LED	Array Lighting	AE26PAR38183010
A8	Spot	12-69	Hal.	Sylvania	75PAR/CAP/SPL/WSP12 120V
B1	Narrow Fl.	12-144	LED	Solais Lighting	LRP38/25/30
B2	Narrow Fl.	12-79	LED	Westinghouse	18PAR38/LED/DIM/30
B3	Narrow Fl.	12-92	LED	Samsung	SI-P8V181DB0US
B4	Narrow Fl.	12-70	Hal.	Philips	75PAR38/HAL/FL25
B5	Narrow Fl.	12-91	LED	Verbatim	P38ES-L1000-C30-B25
B6	Narrow Fl.	12-64	LED	TCP	LED17E26P3830KNFL
C1	Flood	12-85	LED	LEDnovation	LED-PAR38-90-1WD-1WF
C2	Flood	12-146	LED	Zenaro	SL-PAR38C/H/P16/50/E30/TD/26/LAC
C3	Flood	12-87	LED	Honeywell	HWL1FP3811301BDIM
C4	Flood	12-142	Hal.	Eiko	75PAR38/H/WFL-120V
C5	Flood	12-74	LED	Satco Products	S8853
C6	Flood	12-62	LED	Ecosmart	ECS 38 WW FL 120
D1	Color	12-90	LED	Acculamp	ALSP38 900L R9
D2	Color	12-72	LED	Sylvania	LED21PAR38/DIM/P/930/FL30
D3	Color	12-82	LED	Litetronics	LP15566FL4D
D4	Color	12-68	HIR	GE	60PAR/HIR/FL30
D5	Color	12-65	LED	Lighting Science Group	DFN 38 W27 V2 NFL 120
D6	Color	12-140	LED	Philips	BC19.5PAR38/AMB/3000K/ FL25 DIM 120V

Table A2. Pertinent performance data for the lamps included in the demonstration assessment. Complete data is available in *CALiPER Application Summary Report 20*.

	CALiPER DATA						MANUFACTURER DATA		
Code	**Input Power** (W)	**Output** (lm)	**Luminous Efficacy** (lm/W)	**CBCP** (cd)	**Beam Angle** (deg)	**Field Angle** (deg)	**Input Power** (W)	**Output** (lm)	**Beam Angle** (deg)
A1	16.3	840	52	7,609	18	31	17.0	820	17
A2[1]	65.8	737	11	10,330	11	22	75.0	1,060	9
A3	16.3	799	49	19,858	8	13	16.0	800	10
A4	76.1	1,036	14	15,060	10	21	75.0	1,050	10
A5	13.7	787	57	17,543	8	18	20.0	959	8
A6	12.4	976	79	12,947	12	25	13.5	1,000	12
A7	14.4	738	51	10,178	11	24	18.0	932	10
A8	75.3	1,073	14	15,360	11	22	75.0	1,060	12
B1	16.2	983	61	4,183	22	45	16.0	1,025	25
B2	17.2	848	49	3,529	23	47	18.0	870	25
B3	16.9	1,157	68	4,480	24	49	18.0	1,100	25
B4	74.9	1,013	14	4,347	21	43	75.0	1,050	25
B5	17.0	1,002	59	2,753	29	54	17.0	1,000	25
B6	17.1	1,079	63	6,945	20	38	17.0	1,100	25
C1	19.4	874	45	894	52	86	19.0	870	55
C2	15.8	890	56	1,690	39	65	16.0	960	50
C3	10.7	603	57	1,173	38	68	11.0	610	40
C4	76.4	1,017	13	949	53	109	75.0	1,100	50
C5	17.0	906	53	942	39	120	17.0	875	40
C6	18.4	977	53	1,779	42	65	18.0	850	
D1	16.6	932[2]	56	5,805[1]	20	34	20.0	900	25
D2	20.1	1,150[2]	57	3,807[1]	24	54	21.0	1,150	30
D3	14.7	883[2]	60	4,228[1]	25	41	15.0	850	28
D4	62.0	933[2]	15	2,590[1]	30	47	60.0	1,050	30
D5	24.1	1,195[2]	50	6,064[1]	22	41	24.0	1,250	25
D6	18.9	1,353[2]	71	7,161[1]	23	39	19.5	1,200	25

1. The manufacturer data for lamp A2 is for 130 V operation. CALiPER tested the lamp at 120 V.
2. Some of the lamps used for the color quality evaluation (Group D) were dimmed to produce approximately equal illuminance at the center of the beam. The HIR lamp (D4) was not dimmed, since it had the lowest CBCP. All of the values shown in this table were measured at full output.

Appendix B: Example Questionnaire

BEAM QUALITY: PAR38 LAMPS (Group X)

Participant ID: _______

1. Each lamp shown has a unique label from X1 through X6. Please rank order the lamps based on the beam quality that you see in the pattern of light it projects on the wall, by writing the lamp labels in the appropriate spaces below. Write the label for the pattern of light that you find *least* desirable on the end space on the left, and write the label for the pattern of light that you find *most* desirable on the end space on the right, with the others listed between them in order of your ranking. Each space must contain one identifier—no ties. Assume that you are ranking these lamps for use in a retail application that requires a __________ distribution. In other words, if all other factors (such as energy efficiency or cost) were the same, determine the order in which you would choose each lamp for a retail application based on its beam pattern.
2. Beneath each lamp label in your ranking, check the appropriate box to indicate whether you find this lamp unacceptable, acceptable, or outstanding for use in a retail application, based only on the beam quality. Do not check more than one box for each lamp.
3. Beneath each lamp label in your ranking, please identify one or more attributes that are the most desirable and least desirable (in order), using the number key at the bottom of the page. In other words, identify the attributes that most influenced your ranking.

Least Desirable	X4	X1	X6	X2	X3	X5	**Most Desirable**
Unacceptable:	☒	☒	☒	☐	☐	☐	
Acceptable:	☐	☐	☐	☒	☒	☐	
Outstanding:	☐	☐	☐	☐	☐	☒	
+ Most Desirable Attribute:	1	5	6	2	3	5	
2nd Most Desirable Attribute:		6		4		1	
3rd Most Desirable Attribute:						6	
3rd Least Desirable Attribute:	2		5		4		
2nd Least Desirable Attribute:	4	1	4	5			
− Least Desirable Attribute:	3	2	3	7			

something we missed

List of Attributes

1. Hardness or softness of the beam edge
2. Falloff/gradient from the center to the edge of the beam
3. Brightness pattern/consistency
4. Color pattern/consistency
5. Center beam intensity (CBCP)
6. Stray light or lack thereof outside the main beam
7. Other (please specify)

Additional Comments: Have more to say? Let us know what you think about any of these lamps here!

Appendix C: Ranking Factors

Table C1. Count of attributes listed as positively or negatively affecting the beam quality ranking for each lamp. The percentages provided for the LED and halogen lamps in each group are based on the total number of positive or negative notations.

		POSITIVE ATTRIBUTES							NEGATIVE ATTRIBUTES						
Code	Source Type	Beam Edge	Falloff/Gradient	Brightness Pattern	Color Pattern	CBCP	Stray Light	Other	Beam Edge	Falloff/Gradient	Brightness Pattern	Color Pattern	CBCP	Stray Light	Other
Spot Group															
A1	LED	11	10	8	8	2	3	0	2	0	1	3	2	1	0
A2	HIR	7	5	6	2	5	1	0	0	3	1	9	0	8	2
A3	LED	8	3	4	7	11	0	0	6	2	2	2	1	10	1
A4	Halogen	1	2	3	1	5	0	0	4	2	3	6	2	14	1
A5	LED	4	3	2	5	7	0	0	2	5	3	3	0	9	1
A6	LED	1	1	1	2	2	2	0	3	2	3	14	4	10	1
A7	LED	1	1	2	6	7	0	0	3	3	6	4	1	10	1
A8	Halogen	2	2	1	1	6	0	0	4	5	7	6	3	13	2
LED		19%	15%	16%	21%	24%	4%	0%	14%	7%	12%	22%	8%	35%	3%
Halogen		22%	17%	15%	14%	31%	2%	0%	7%	15%	13%	21%	3%	35%	6%
Narrow Flood Group															
B1	LED	0	2	2	1	4	0	0	7	1	8	6	0	14	1
B2	LED	8	10	6	9	6	1	0	3	1	2	1	4	1	0
B3	LED	3	1	4	4	3	0	0	7	3	5	2	2	12	1
B4	Halogen	2	4	4	2	2	0	0	2	3	5	5	4	12	3
B5	LED	11	7	6	7	3	4	0	1	1	2	5	9	0	2
B6	LED	9	6	8	7	12	6	1	1	1	0	0	5	0	0
LED		21%	17%	17%	19%	19%	7%	1%	18%	6%	16%	13%	19%	25%	4%
Halogen		14%	29%	29%	14%	14%	0%	0%	6%	9%	15%	15%	12%	35%	9%
Flood Group															
C1	LED	10	8	6	15	2	2	0	2	1	2	3	0	1	1
C2	LED	1	4	3	1	2	1	0	1	2	3	11	2	3	0
C3	LED	3	2	1	2	2	1	0	2	4	4	11	1	2	0
C4	Halogen	4	2	2	3	0	2	0	3	2	10	8	1	6	1
C5	LED	0	2	0	2	3	1	0	3	5	7	11	8	1	1
C6	LED	4	10	9	12	5	1	0	3	1	0	0	0	2	0
LED		16%	23%	17%	28%	12%	5%	0%	11%	13%	16%	37%	11%	9%	2%
Halogen		31%	15%	15%	23%	0%	15%	0%	10%	6%	32%	26%	3%	19%	3%

Appendix D: Individual Product Rankings

Table D1. Beam quality ranking and rating details for each observer. Some participants did not complete the questionnaire form in its entirety.

Responses of Individual Observers — Rank (Number) and Acceptability Rating (Color) for Beam Quality

Group	Code																					
Spot Group	A1	5	1	2	1	1	8	2	3	1	2	1	1	1	5	1	5	1	1	6	-	-
	A2[1]	3	2	6	5	3	7	6	6	4	5	4	2	3	4	6	6	2	3	3	-	-
	A3	1	3	4	6	2	1	1	1	2	1	7	6	6	1	2	7	5	2	1	-	-
	A4[2]	4	-	7	2	5	4	4	7	7	3	2	5	4	7	4	1	3	5	7	-	-
	A5	2	6	1	4	6	2	5	5	3	7	5	-	7	3	5	3	7	6	2	-	-
	A6	8	7	5	7	7	5	7	4	8	6	6	4	5	2	3	2	4	8	8	-	-
	A7	7	-	3	3	8	6	3	2	5	4	8	-	2	6	7	4	6	4	5	-	-
	A8[2]	6	8	8	8	4	3	8	8	6	8	3	7	8	8	8	8	8	7	4	-	-
Narrow Fl. Group	B1	5	5	6	5	5	6	5	6	6	3	6	6	6	6	6	4	6	6	**5**	6	-
	B2	3	3	2	1	2	2	1	1	1	2	1	2	3	3	2	3	3	1	4	2	-
	B3	2	4	4	6	6	5	6	5	4	4	5	5	4	2	4	5	5	5	2	5	-
	B4[2]	6	6	5	4	4	4	4	4	5	6	4	4	5	5	5	6	4	4	3	4	-
	B5	4	2	3	2	3	3	2	3	3	5	2	3	1	4	3	1	1	3	6	3	-
	B6	1	1	1	3	1	1	3	2	2	1	3	1	2	1	1	2	2	2	1	1	-
Flood Group	C1	1	2	2	1	3	2	1	5	2	1	5	2	1	2	2	2	2	2	3	1	-
	C2	4	3	5	4	5	3	5	2	5	5	1	3	2	3	4	4	3	4	5	5	-
	C3	3	5	4	2	4	4	4	6	3	3	6	5	4	4	3	3	4	5	2	2	-
	C4[2]	5	4	1	6	2	6	3	4	4	6	3	4	5	6	6	6	6	3	6	4	-
	C5	6	6	6	5	6	5	6	3	6	4	4	6	6	5	5	5	5	6	4	3	-
	C6	2	1	3	3	1	1	2	1	1	2	2	1	3	1	1	1	1	1	1	6	-

1. HIR

Rating Scale
Not Provided
Unacceptable
Acceptable
Outstanding

Table D2. Shadow quality ranking and rating details for each observer. Some participants did not complete the questionnaire form in its entirety.

Responses of Individual Observers — Rank (Number) and Acceptability Rating (Color) for Shadow Quality

Code																					
B1	2	2	1	2	2	2	2	1	1	2	2	1	1	2	1	2	3	2	1	2	-
B2	3	3	3	3	3	3	3	3	3	4	3	3	3	5	3	3	2	3	3	3	-
B3	5	5	4	5	5	6	5	4	5	5	4	5	5	4	4	5	5	5	5	4	-
B4[1]	6	6	6	6	6	4	4	5	6	6	6	6	6	6	5	6	6	6	6	5	-
B5	4	4	5	4	4	5	6	6	4	3	5	4	4	3	6	4	4	4	4	6	-
B6	1	1	2	1	1	1	1	2	2	1	1	2	2	1	2	1	1	1	2	1	-

1. Halogen

Rating Scale
Not Provided
Unacceptable
Acceptable
Outstanding

Table D3. Color quality ranking and rating details for each observer. Some participants did not complete the questionnaire form in its entirety.

Code	**Responses of Individual Observers** Rank (Number) and Acceptability Rating (Color) for Color Quality																				
D1	4	6	5	6	1	1	5	6	5	1	2	5	5	6	4	1	6	2	5	5	-
D2	1	4	2	4	6	3	3	4	3	3	4	4	4	2	3	4	4	6	2	4	-
D3	2	2	3	3	3	4	1	2	4	5	5	1	1	4	2	5	1	3	1	1	-
D4[1]	5	1	4	2	2	2	2	3	1	4	1	3	2	3	5	2	3	5	4	3	-
D5	6	5	6	5	5	5	6	5	6	6	6	6	6	5	6	3	5	4	3	6	-
D6	3	3	1	1	4	6	4	1	2	2	3	2	3	1	1	6	2	1	6	2	-

1. HIR

Rating Scale
Not Provided (white)
Unacceptable (red)
Acceptable (yellow)
Outstanding (green)

Appendix E: Lamp Photographs (Beam Quality Evaluations)

Additional photographs are available in CALiPER Application Summary Report 20, or in the detailed reports for each CALiPER-tested lamp, which are available online at ssl.energy.gov/caliper.html.

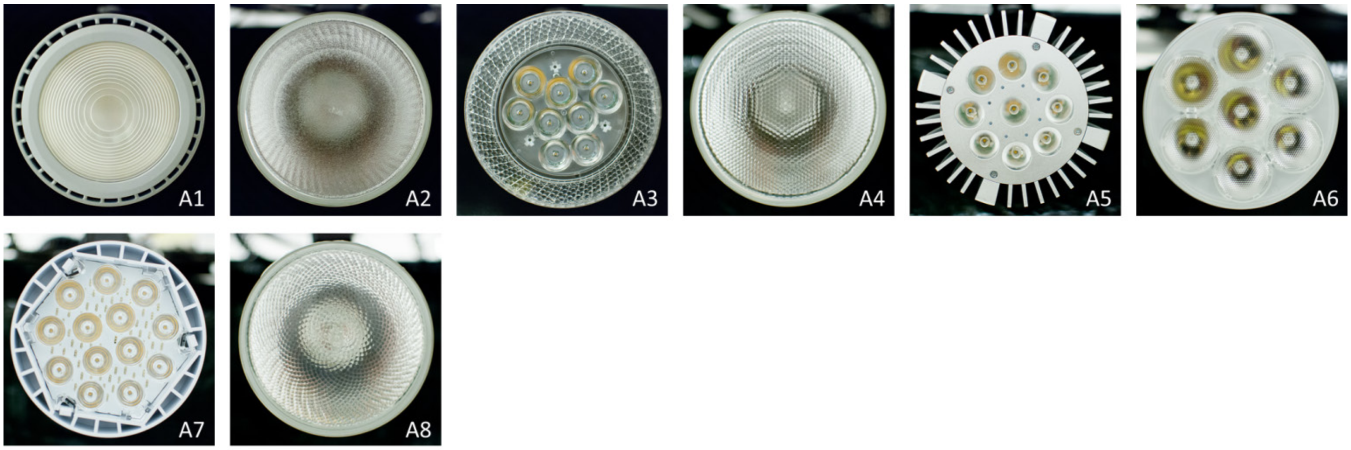


Figure E1. The lamps evaluated for beam quality in the spot group.

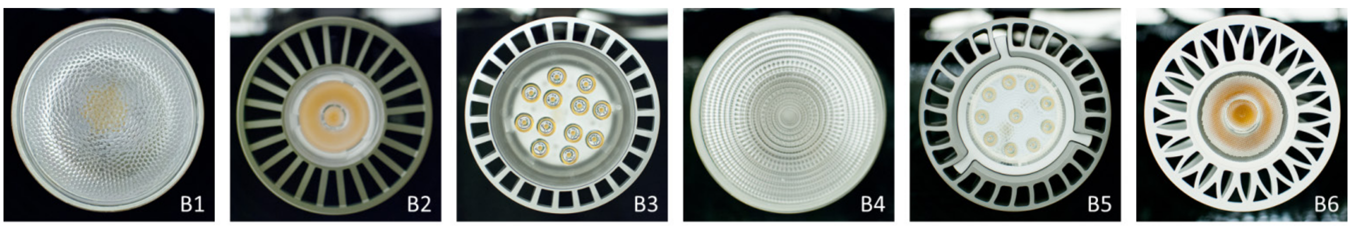


Figure E2. The lamps evaluated for beam quality in the narrow flood group.

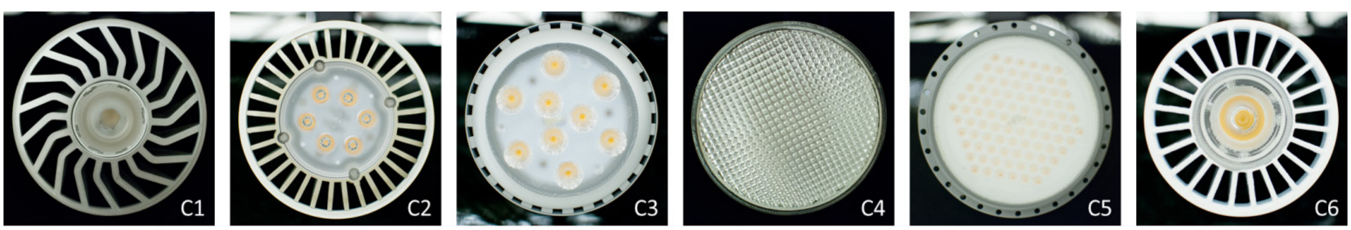


Figure E3. The lamps evaluated for beam quality in the flood group.

Appendix F: Histogram of Responses (Color Quality Evaluation)

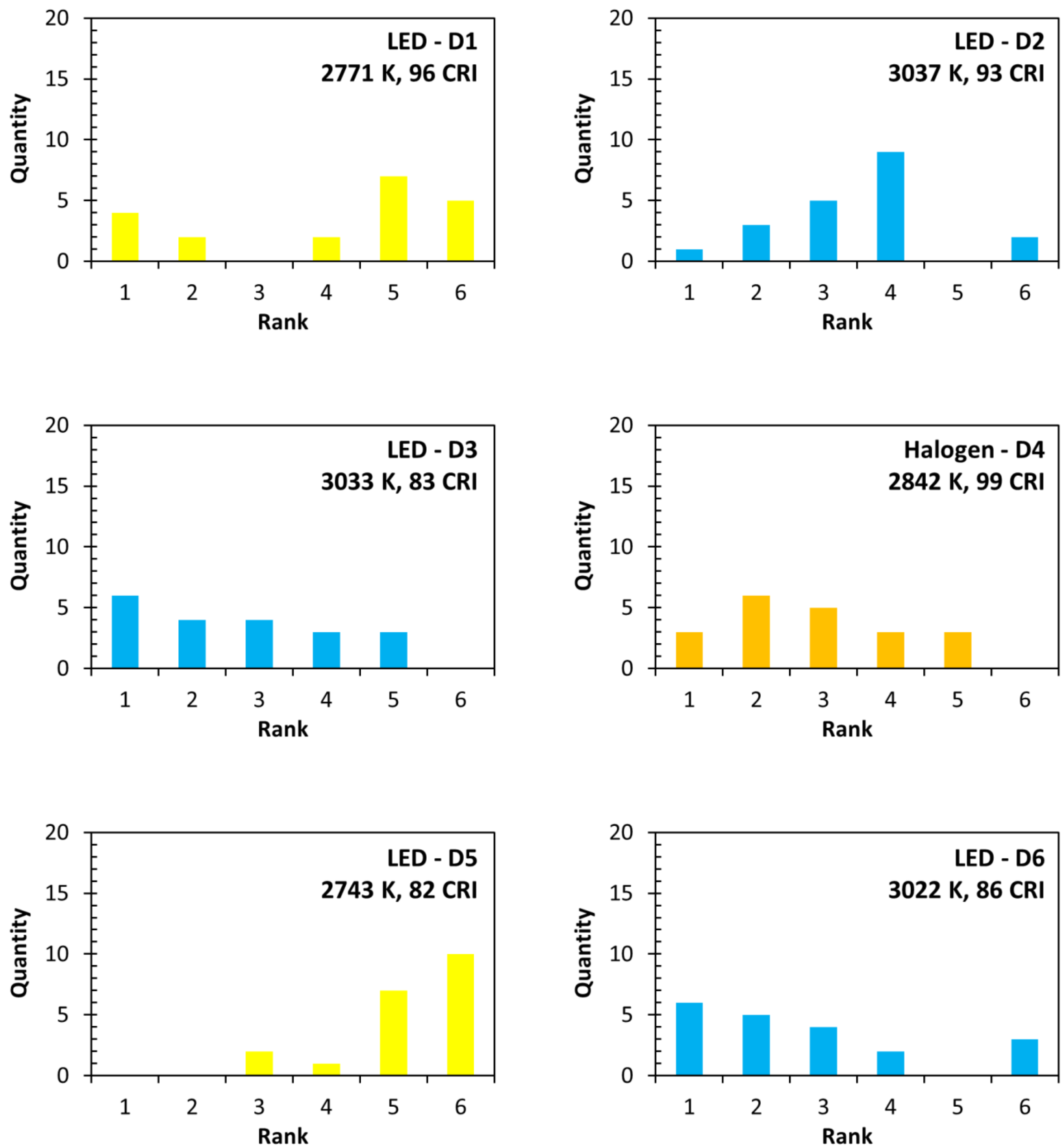


Figure F1. Histogram of individual responses for the color quality evaluation. For most of the lamp types, the responses were fairly evenly divided among the ranks. Unlike all of the other groups, no lamp in the color quality evaluation received 10 or more first place ranks—two lamps (D3 and D6) received six votes for first place. (Lower ranks were better; a rank of one was the most favorable, whereas a rank of six was the least favorable.)