A new subjective-objective approach to evaluating lighting quality: A case study of concert lighting for Cambridge King’s College Chapel

Wing Lam Lo Koen Steemers
University of Cambridge, The Martin Centre for Architectural and Urban Studies, Department of Architecture, Cambridge, United Kingdom.

ABSTRACT
Drawing on theoretical insights from acoustic research, this paper presents a subjective-objective method that complements and unifies existing approaches to assessing lighting quality. In essence, the method establishes numerical relationships between subjective attributes and objective lighting measures, the applicability of which is demonstrated through a comparative study of four different concert light settings in Cambridge King’s College Chapel. A total of 624 subjective responses were collected from a full-scale experiment. These focused on seven subjective aspects: visual clarity, visual balance, visual uniformity, brightness, spatial intimacy, appropriateness/comfort, and overall impression of the luminous environment. Variations in the response were analysed systematically and correlated with 22 objective attributes, which were computed using luminance and photometric data extracted from High Dynamic Range images. This study reveals significant connections between the selected subjective parameters and the objective attributes that are used to describe a visual scene mathematically and spatially.

Keywords: Lighting Quality, Objective Measure, Subjective Judgment, High Dynamic Range Imaging

INTRODUCTION
Despite the growing awareness of the need to create a sensible lit environment and to develop appropriate lighting systems to reduce energy use, there is currently little consensus on the way we evaluate lighting quality in relation to its quantity. This is a complex subject, in part because lighting perception is often influenced by non-quantifiable human factors. To be able to characterise lighting quality accurately, it is important to gain an understanding of the connection between seeing and perceiving. In the past, several attempts (Flynn et al., 1979; Loe et al., 1994, Moore et al., 2003) have been made to explore this subjective-objective connection, yet little has been established about its framework and applicability. In this paper, drawing an analogy between acoustics (Beranek, 1962; Barron, 1988; Barron, 1993) and lighting research, we develop an analytical framework based on the acoustics’ numerical framework — one that studies what perceptual attributes contribute to the overall listening experience in concert halls — and we test this framework in order to analyse concert lighting for Cambridge King’s College Chapel. In the literature, the majority of lighting studies have focused on contemporary spaces such as offices, schools and museums (Newsham et al., 2001; Wymelenberg et al., 2014). In this sense, our study is unique in that it examines the effects of light in a historic context. Our aim is to present and apply the subjective-objective framework to study lighting quality. First, we describe the experimental set-up, including the procedure through which our subjective and objective data were collected. We then identify the key attributes that influence lighting impression and the relationships among them, followed by the key findings. Finally, we discuss the strengths and limitations of this study.
EXPERIMENTAL PROCEDURE AND DATA COLLECTION

Working with historic buildings poses practical and logistical challenges, especially when it comes to collecting objective data and conducting field experiments. For fidelity, we reconstructed a concert environment with the use of dummies (Lo & Steemers, 2014) (Figures 1a and 1b). While peripheral lighting for the walls and the wooden screen were switched on in all tests, luminance, spot luminance and illuminance were measured under four artificial light settings (Figure 1c): Setting I (Interim Lighting), Setting II (Interim Plus), Setting III (All Lighting) and Setting IV (Rig Lighting). The visual fields of occupants were captured with a fish-eye lens at six locations (Spots O and A = audience members; Spot B = conductor; Spots C, D and E = musicians) (Figure 1a), resulting in a total of 24 visual scenarios.

All the images were generated using High Dynamic Range (HDR) photography. Each HDR image was calibrated against physical luminance measurements taken with a Minolta LS-100 Luminance meter and was generated in Radiance by merging 15 multiple-exposure RAW images. With a fixed aperture of f/5, these images were taken with exposure times ranging from 1/1000s to 15s. The photometric data and images were then processed in Matlab and analysed in relation to the structure of our visual field, providing a comprehensive mathematical and spatial analysis of light for the antechapel (Lo & Steemers, 2014). Through the use of these techniques and classical principles of quantifying light, 22 equations were derived from the images to evaluate visual acuity, luminance contrast, uniformity, variation, perceived brightness and contrast, relative luminance, visual boundary and light patches for all the scenarios. Subjective data were collected through a highly structured field experiment and questionnaires.

Seventy-eight university students and staff members volunteered to participate in the experiment. Twenty-six responses were collected at each spot. The participants were asked to imagine themselves attending or performing in an orchestral and chorus concert of religious music, and each was assigned to a specific spot. Upon completion, they were assigned to another spot with a different visual field and were asked to repeat the task. They were also instructed to compare Settings I, II and III with Setting IV (i.e. the control) and to make objective and subjective judgements accordingly. To avoid possible bias, we randomised the sequence in which the settings were presented. For each setting, the participants had to fill in a one-page questionnaire, which was organised into three main sections — i) Visual clarity, ii) Distribution of Light and iii) Spatiality — and concluded with a question concerning the overall lighting impression. Additional space was provided for further comments. The subjective questions focused on 1) visual clarity, 2) visual uniformity, 3) visual balance, 4) brightness, 5) spatial intimacy, and 6) appropriateness/comfort. The ratings for these were given on a seven-point Likert scale. Unlike other
lighting studies, which relied on semantic differential scales (Flynn et al., 1973; Hawkes et al., 1979; Loe et al. 1994), our questionnaire was a combination of factual, semi-subjective and subjective questions, which enabled us to justify the responses. For visual clarity, for example, we asked the audience participants ‘how well can you read the programme and see the facial expressions of the musicians’ along with the Hazy/Clear scale. Other semantic differential scales used in this study were Non-uniform/Uniform, Inappropriate/Appropriate, Dim/Bright, Confined/Public and Uncomfortable/Comfortable. More details on the experimental set-up and questionnaire design are available in Lo & Steemers (2014).

ANALYTICAL METHOD, FINDINGS AND DISCUSSION

Validation of the experimental technique

The experimental technique was validated i) by comparing the overall impression rating obtained from the concluding question — ‘How satisfied are you with the overall lighting experience?’ — with ratings computed by averaging the sum of the average scores derived from the three main sections; ii) by examining whether the structure of the experimental design had an effect on the overall impression; and iii) by identifying whether there were significant differences among the scores of each question.

i. The changes in the rated and computed impression scores follow a similar pattern, except in the case of Setting III where a divergence is observed (Figure 2). Results from a T-test, performed without assuming equal variances, indicate that there is a significant difference between the scores at Spot O (t (31.40) = -3.39, \( \rho < .01 \)), Spot A (t (29.69) = -4.51, \( \rho < .01 \)), Spot B (t (30.87) = -3.45, \( \rho < .01 \)) and Spot E (t (30.84) = -3.57, \( \rho < .01 \)). Their mean values show that the computed scores (\( \bar{O} = 5.54; \bar{A} = 5.42; \bar{B} = 5.73; \bar{E} = 5.63 \)) were significantly higher than the rated scores (\( \bar{O} = 4.27; \bar{A} = 3.81; \bar{B} = 4.63; \bar{E} = 4.31 \)). This implies that although the factual questions and Dim/Bright scale received higher scores, the participants felt uncomfortable and tense in some cases because of the brightly lit environment, and had an undesired impression as a result. This is also evident in their comments, as one said, ‘By far the least comfortable lighting of all. It’s very stressful for the eyes.’ (Spot A: Rated score = 1; Computed score = 5.22)

ii. Results from ordinal regression indicate that the order of the experimental session was not significantly associated with the tendency of response (Sig. =.279). In addition, there is no consistent agreement in the standard deviation of the scores observed between the first and second sessions. We therefore reject the hypothesis that the experimental structure affected the overall impression.

iii. A one-way Kruskal-Wallis analysis of variance test was used to analyse the ratings for all the scenarios, and post-hoc tests were performed to pinpoint the differences. The broadest consensus among the responses is observed at all the positions in Setting III. Significant differences, however, were detected between the factual and subjective questions in the section on ‘Distribution of Light’, as well as in the section on ‘Spatiality’ between the scales of Inappropriate/Appropriate, Dim/Bright and Uncomfortable/Comfortable. Somewhat surprisingly, the participants tended to give lower ratings to the subjective questions when details of architectural features could be clearly seen.
Correlation between subjective and objective attributes

Next, we grouped the factual and subjective questions in accordance with our subjective attributes, whose internal reliabilities were assessed by Cronbach’s alpha. Pearson correlation coefficients were calculated to examine the independent relationship between the attributes. The results are in good agreement with significant results obtained for all the attributes ($\rho < .001$). The correlation matrix reveals that visual clarity, visual uniformity, visual balance and appropriateness/comfort account, respectively, for 10.24%, 26.73%, 46.10%, and 72.76% of the variability in the overall impression, while brightness and spatial intimacy account for 4.12% and 2.02%.

A multiple regression analysis was performed to test the dependence of the attributes on the overall impression. The results show that the regression model is statistically significant ($\rho < .001$), with approximately 74.5% of the variability of the overall impression accounted for by the subjective attributes ($r^2 = 0.745$). The unexplained variation may be attributed to the occupant’s role, viewing position, expectation and experience (Lo & Steemers, 2014). The regression coefficients suggest that visual balance, visual uniformity and appropriateness/comfort are significant predictors of the overall lighting impression ($\rho < .01$). However, there is no statistically significant linear dependence of visual clarity, brightness and spatial intimacy on the overall impression. This suggests that the significant predictors have direct influences on the overall impression, while the insignificant predictors have indirect influences.

The predictors were entered into the model in the following sequence: visual clarity, visual uniformity, visual balance, brightness, spatial intimacy and appropriateness/comfort. The change in $R^2$ shows that visual uniformity explains 19.8% more of the variance in the overall impression than visual clarity (10.2%) does alone, and that visual balance (19.5%), brightness (0.3%), spatial intimacy (0.8%) and appropriateness/comfort (23.9%) further explain the variability. The analyses indicate that appropriateness/comfort was the most important subjective attribute (Figure 3), while brightness was the least important, which suggests that the participants prioritised their needs subconsciously such that the exact brightness level was not pertinent to the overall impression.

The five most significant objective measures: Univariate Linear Regression Analysis

The Pearson correlation of the subjective attributes and objective measures were computed. Because of the large number of correlation coefficients, the following steps were used to categorise the objective measures: i) ranked the five most significant measures; ii) identified the core measures that are common to all the subjective attributes; and iii) identified additional measures that were selected in step one but were not common to all the subjective attributes. Among 154 correlation tests, 84 were reported as highly significant ($\rho < .01$) and 13 were regarded as statistically significant ($\rho < .05$) (Lo & Steemers, 2014). As Figure 4 shows, $L_{\text{avg}}$, $\text{RIM}_{\text{whole}}$, $\text{RISD}_{\text{whole}}$, $\text{RL}_{\text{std}}$, $\text{Light to Dark ratio (L:D)}$, $\text{RL}_{\text{avg}}$ and Total area of light patches (Area Light patches) are the core measures, while $L_{\text{std}}$, Perimeter and VASheets are the additional measures. It is interesting to find that $L_{\text{avg}}$ is the strongest parameter with which to evaluate visual clarity,
visual uniformity, visual balance, brightness and spatial intimacy, while AreaLight patches and VASheets are the strongest ones for appropriateness/comfort and overall impression, respectively. Lavg is the only measure that was ranked as one of the top five for all the subjective attributes except for the overall impression. Lavg only accounts for the physical measurements of light, which does not reflect changes in acuity within the visual field, eye movements and the selective nature of the eye. Considering Lavg to be a fundamental measure along with other key attributes, however, would be a more robust and thorough approach to making predictions of subjective responses.

To better understand the relationships between the significant measures and subjective attributes, we plotted the subjective ratings against the objective measures in two different ways: a) first as a jittered scatterplot, which was generated with the raw data (N=624), enabling analysis of changes in response in relation to the magnitude of the objective attributes; b) second as a non-jittered scatterplot, which was generated with the total mean score given at each position under each lighting condition (N=24), enabling a more general comparison among different visual scenarios. Furthermore, all the data points were categorised on the basis of the occupant’s role in order to highlight differences in response between the audience members, conductor and musicians.

1) Visual Clarity. The correlation between visual clarity and Lavg (r = 0.46) is the strongest (Figure 5). An increase in Lavg led to an improvement in visual clarity. For low levels of Lavg, as is the case for Setting IV, there was a greater variation in the responses regardless of the occupant’s role. Interestingly, there is an appreciable difference in the visual clarity ratings between Spot B and Spot E in Setting I, with mean ratings of 4.09 and 5.40, respectively, although the Lavg levels are similar. This can be attributed to the apparent differences in RLavg, RLstd, L:D and AreaLight patches. Halving the levels of RLavg and RLstd as well as quartering the L:D ratio and the area of light patches could potentially lower the total mean rating for visual clarity (Figure 7). The two-level greyscale images (Figure 7) illustrate that there is a striking contrast of spatial hierarchy between Spot B and Spot E. For Spot B, white pixels are mostly clustered in the background. For Spot E, however, white pixels are distributed evenly across the work plane and background, highlighting the essential features of a musician’s visual scene.

2) Visual Uniformity and Visual Balance. The ranking of the five most significant measures for visual balance is exactly the same as that for visual clarity, although the correlation coefficients are lower. Using other physical measures is more appropriate for describing these relationships. For visual uniformity and visual balance, an increase in RLavg (r = 0.20) and RIMwhole (r = 0.16) respectively resulted in a moderate increase in the mean ratings. The mean ratings of these two attributes were at their highest when RLavg and RIMwhole were among the highest of all the settings. In contrast, the attributes received the lowest scores in Setting IV at Spots A, C and E, and in Setting I at Spot B. For Spots O and D, the lowest ratings were observed in Setting I. These are also reflected in the overall rating specifically for Setting I at Spots O and B and for Setting IV at Spot C (Figure 2). This, however, was not the case for Setting IV at Spots A, D
Jittered scatterplot (left) and non-jittered scatterplot (right) of visual clarity plotted against Lavg based on the raw data (N = 624) and average ratings (N = 24) respectively.

Figure 6  Non-jittered scatterplot: Spatial intimacy plotted against L:D

Setting I - Spot B (Conductor)
Lavg  = 15.015 cdm⁻²                 RLavg  = 9.2% L:D ratio = 0.148; Area = 37782

Setting I - Spot E (Musician)
Lavg  = 15.386 cdm⁻²                 RLavg  = 18.0% L:D ratio = 0.591; Area = 109120

Figure 7  Spatial analysis of light for the antechapel and E, where the ratings for appropriateness/comfort were comparatively higher. This suggests that by enhancing the appropriateness of the visual environment and visual comfort, it is possible to improve the overall impression despite lower ratings for visual balance and visual uniformity.

3) Brightness and Spatial Intimacy. There was a good correlation (r = 0.36) between brightness and RISDwhole — the second most significant measure. With bright sources isolated, this measure refers to the relative brightness contrast of the visual field. Analysis of scatterplots revealed that the audience members’ ratings exhibited more scatter than those of the conductor and musicians, suggesting that the agreement for brightness was to some extent related to the occupant’s role. This observation also applies to spatial intimacy. As Figure 6 shows, despite its positive correlation (r = 0.26) with L:D, the responses given by the audience members at Spots O and A, where the objective values were very similar, yielded a broader range of scores (min. = 3.23, max. = 5.65) than that given by the conductor and musicians. The lowest and highest mean ratings were observed in Setting IV and Setting III, respectively. In fact, RLavg and RLstd Spots O and A were the same in Setting III (RLavg = 4.2%, RLstd = 8.9%) and were nearly double those in Setting IV (O: RLavg = 1.9%, RLstd = 4.5%; A: RLavg = 1.8%, RLstd = 4.6%). Taken together, these findings support the view that the core objective measures are pertinent to the subjective evaluation.

4) Appropriateness/Comfort. Although this attribute contributes the most variability to the overall impression, as discussed above, its correlations with the core objective measures are comparatively weak, indicating that this subjective attribute is difficult to be quantified with physical measures alone. AreaLight patches is found to be the most significant measure for this subjective attribute (r = 0.214), followed by RLstd-L:D, RLavg and Perimeter. These measures have similar correlation coefficients, indicating that spatial analysis offers a more complete understanding of a visual scene than does merely relying on physical light levels. A further comparison of the objective variables reveals another interesting observation. Despite
similar values for AreaLight patches, as in the case of Spot E in Setting I (area = 109120 pixels, mean rating = 4.17) and II (area = 113489 pixels, mean rating = 4.83), there is a discrepancy in the ratings for appropriateness/comfort. An increase in L:D (I = 0.59, II = 0.63) and a decrease in Perimeter (I = 7624.69 pixels, II = 7353.52 pixels) led the participants to prefer Setting II over Setting I.

Combination of the significant measures: Multivariate Regression Analysis

It is worth reminding that the main objectives of this study are to test the analytical framework and the experimental methods, as well as to establish relationships between the subjective attributes and the significant measures ($\rho < .01$ and $\rho < .05$), rather than simply to find best-fit regression lines. To derive a prediction equation — through multivariate regression analysis — for each subjective attribute, we checked for multicollinearity with reference to the tolerance value and variance inflation factor (VIF) of each objective measure. The analysis involved two steps: i) an overall analysis based on all 624 responses; and ii) a group analysis based on the type of occupants. The first analysis shows that there are statistically significant correlations for all the equations, where visual balance ($r = 0.21$, Sig. = .003) and visual clarity ($r = 0.57$, Sig. = .000) are the least and the most correlated with the objective measures, respectively. The second analysis, however, reveals that the objective measures are the least correlated with appropriateness/comfort in the cases involving the conductor ($r = 0.44$, Sig. = .000), but it returns null results for the audience members and the musicians. The weak correlations are in agreement with that obtained from the univariate linear regression analysis.

Combination of the significant measures: Multivariate Regression Analysis

<table>
<thead>
<tr>
<th></th>
<th>$r$</th>
<th>$r^2$</th>
<th>Sig.</th>
<th>Average VIF</th>
</tr>
</thead>
<tbody>
<tr>
<td>$Brightness_{All}$</td>
<td>0.50</td>
<td>0.25</td>
<td>.000</td>
<td>4.78</td>
</tr>
<tr>
<td>$Brightness_{Conductor}$</td>
<td>0.68</td>
<td>0.46</td>
<td>.000</td>
<td>1.63</td>
</tr>
</tbody>
</table>

Figure 8 Prediction equations for the subjective attribute of brightness

As Figure 8 shows, although the objective measures account for 26% of the variance ($r = 0.51$) in brightness ($Brightness_{All}$), the percentage becomes almost doubled ($r = 0.68$) when only the conductor’s responses ($Brightness_{Conductor}$) are taken into account. This suggests that different occupants’ luminance requirements and judging criteria led to the lower percentage of variance in $Brightness_{All}$. A closer examination of the equation for $Brightness_{Conductor}$ reveals that the subjective judgement of brightness is strongly associated with the relative luminance contrast as seen in the left ($LLM_{std}$) and right ($LRM_{std}$) monocular crescents of the visual field.

CONCLUSION

The subjective-objective method, borrowed from acoustic research, is proven to be effective and reliable through an analysis of concert lighting in King’s College Chapel. With this method, it is possible to extract hidden information regarding perceptual judgments, alongside factual parameters. Another virtue is that it gives an overview of the effects of light on the occupants’ impression, the variances of which could be examined in depth through six subjective aspects: visual clarity, visual balance, visual uniformity, brightness, spatial intimacy, and appropriateness/comfort. This method can also make possible the creation of a positive lighting impression by manipulating the way space, objects, surfaces and people are illuminated. Because this study only tested a limited number of lighting conditions, further research involving a wider range of stimuli and wider range of the objective magnitude of the measures is needed in order to formulate universal relationships between the variables. The significant results obtained from
this study demonstrate that the experimental techniques and the analytical framework can be applied successfully to a highly sophisticated context. With easier access to HDR imaging, this method is readily applicable for analysing the quality of light in other complex buildings. Meticulous attention to detail of the luminous environment and its relationship with people is essential. Such an understanding needs to be reflected in the experimental design and setting, for example by reconstructing the visual scene through the use of dummies. Further, it is important to address lighting criteria for different types of occupants when acquiring subjective judgments. This would help us better appreciate the diverse nature not only of the luminous environment, but also of the occupants’ perception.

ACKNOWLEDGMENTS

This paper is based on the findings of Wing Lam Lo’s on-going PhD research. She would like to thank the Cambridge Overseas Trust, Emmanuel College, King’s College and Dr. Tom White for their support, and the volunteers for donating their time to participate in this study.

REFERENCES


Acronyms and abbreviations

Lavg : Average luminance of the full visual field (cdm\(^2\))
Lstd : Standard deviation of luminance of the full visual field (cdm\(^2\))
Lvar : Variation of luminance of the full visual field (cdm\(^2\))
RIMwhole : Relative mean of pixel intensity of a full visual image as resolved by the eyes. The pixel intensity value is multiplied by the relative visual acuity as derived from the visual angles subtended at the eye
RISDwhole : Relative standard deviation of pixel intensity of a full visual image as resolved by the eyes
RLavg : Average of RL value with bright sources isolated from an image, where Relative Luminance (RL) is a ratio of spot luminance to maximum luminance, i.e. \(\text{RL} = (\text{Luminance}_{\text{spot}} / \text{Luminance}_{\text{max}}) \times 100\%\)
RLstd : Standard deviation of RL value
RLarea : Percentage of area for bright sources as isolated from the calculation of RLavg in relation to the full visual field
L:D : Light to Dark ratio: By converting a coloured image into a two-level greyscale image, the ratio is derived by calculating the total number of white and black pixels.
Total area of light patches (Area\(_{\text{light patches}}\)): Total number of white pixels of a two-level greyscale image
Perimeter: Total length of outer edges around white area of a two-level greyscale image
VA: Ability to discern fine detail of music sheets or programmes, i.e. \(\text{TargetLuminance} : \text{BackgroundLuminance}\)