UC Berkeley
Indoor Environmental Quality (IEQ)

Title

An Examination of Range Effects When Evaluating Discomfort Due to Glare in Singaporean Buildings

Permalink

https://escholarship.org/uc/item/50t169w4

Authors

Kent, Michael
Jakubiec, JA

Publication Date

2021-12-01

DOI

10.1177/14771535211047220

Peer reviewed

An examination of range effects when evaluating discomfort due to glare in Singaporean buildings

MG Kent PhD[a], **JA Jakubiec** PhD[b]

[a] Berkeley Education Alliance for Research in Singapore, Singapore

[b] John H. Daniels Faculty of Architecture, Landscape, and Design, University of Toronto, Canada

Abstract

This article discusses ratings of visual discomfort from glare across different buildings located in Singapore. These data were used to determine if range effects influence the vertical illuminance values for the same ratings of visual discomfort when the category rating procedure is used. The effect occurs when maxima and minima vertical illuminance (i.e. the range) vary across buildings. Our analyses showed that with a higher vertical illuminance range in a building, the mean vertical illuminance value for the same criterion of visual discomfort also increased. The results suggest that the effect caused by different ranges of measured vertical illuminance present across the buildings biased the ratings of visual discomfort. Although these effects may be unavoidable in some buildings that have vastly different levels of light, the data suggest that the overall range of vertical illuminance must be carefully evaluated when predicting visual discomfort. Matching these conditions may enable vertical illuminance to provide more reliable evaluations of discomfort due to glare.

Email: michaelkent@berkeley.edu

1. Introduction

The category rating procedure is a widely used method when evaluating visual discomfort in buildings.[1–3] Observers are typically asked to select a rating that best describes their current level of discomfort due to glare from several available criteria.[4] Like most subjective assessment methods, this procedure can provide unreliable results due to experimental bias.[5,6] For example, when observers evaluate their visual environment using a rating scale, they are not always aware of the range of conditions that may lead to glare. This causes them to create their own impression of what each criterion on the rating scale means and how it corresponds to the visual environment. When this happens, our previous research demonstrated that observers provided lower subjective ratings of visual discomfort when they had not yet seen the full stimulus range of the glare source. While observers that had seen this provided higher ratings, it was not clear whether these assessments provided a better approximation to what would have been given in a natural setting (e.g. an office)). Nevertheless, these results showed that care is required when glare is evaluated using the category rating procedure.[7]

Subjective evaluations can also be influenced by the range bias effect.[5] Range bias is a phenomena where the range of stimuli available influence the threshold values at which subjective evaluations occur – as stimuli maxima increase, the values at which subjective criteria such as visual discomfort are assessed also increase. For example, range effects have been found to influence discomfort due to glare when using an adjustment procedure (i.e., the source luminance was varied to target thresholds of visual discomfort).[8] Table 1 shows that when the stimulus range increases, the final setting made to the same criterion of discomfort due to glare (i.e., the highest criterion (equivalent to intolerable glare)) also increases.

Table 1 Three stimulus ranges that present the minimum and maximum luminances, and the mean luminance when the glare source output was adjusted to a threshold of visual discomfort observers thought resembled "almost intolerable glare".[8]

Stimulus range		Mean luminance at final setting to highest glare criterion (cd/m^2)
Minimum luminance (cd/m^2)	Maximum luminance (cd/m^2)	
	5,106	4,112
441	7,288	5,009
	9,469	6,186

This study[8] also showed that settings were biased when different ranges were used in the first condition that was seen (i.e., when observers had no prior knowledge), and persisted when adjustments were performed across the other range conditions (i.e., after they had experienced the first range). Since the effect is caused when different ranges are used, its influence has also been found when evaluating preferred colour[9] and brightness.[10]

In daylit buildings, a recommended method that evaluates the risk of discomfort due to glare is to calculate the daylight glare probability (DGP).[11] The DGP was originally derived in a study,[3] which asked observers to evaluate daylit windows using the category rating

procedure. From 349 evaluations collected under a range of different conditions, equation (1) was proposed to calculate glare from daylight:

$$DGP = 5.87 \bullet 10^{-5} \bullet E_v + 9.18 \bullet 10^{-2} \log\left(1 + \sum_i \frac{L_{s,i}^2 \bullet \omega_{s,i}}{E_v^{1.87} \bullet P_i^2}\right) + 0.16$$

(1)

whereby, E_v is the vertical illuminance received at the eye (lx), $L_{s,i}$ is the luminance of contrast glare source(s) (cd/m²), $\omega_{s,i}$ is the solid angle of the source(s) (sr), and P_i is the position index of the source(s). Note: P_i is a unit-less factor that increases as a glare source is closer to the visual periphery.

Table 2 shows the categories were established to minimise the potential of causing discomfort due to glare in buildings from daylit windows.[12] There are four step-changes on the DGP scale, whereby the lowest criterion corresponds to calculated DGP values that are ≤0.33 and the highest criterion to the highest when DGP values are ≥0.52. Since thresholds in Table 2 were calculated from mean DGP at each of the four ratings given by observers,[12] it is difficult to discern across which range of DGP corresponds to which threshold. Thresholds presumably based around values in Table 2 containing similar semantics were later published in the EN 17037,[11] which overcame this issue. These were as follows: "Glare is mostly not perceived" (DGP≤0.35); "Glare is perceived but mostly not disturbing" (0.35<DGP≤0.40); "Glare is perceived and often disturbing" (0.40<DGP≤0.45); and "Glare is perceived and mostly intolerable" (DGP≥0.45).

Table 2 The ranges of calculated DGP and their corresponding categories of expected discomfort caused due to glare.[12]

Categories of discomfort due to glare	DGP
Imperceptible	0.33
Perceptible	0.38
Disturbing	0.42
Intolerable	0.52

To ensure observers experienced discomfort due to glare in the original DGP study,[3] the test-room conditions were setup to produce high vertical illuminances. The authors orientated the test-rooms so that their windows always faced the direction of the Sun and in some conditions, observers were seated 1.5 m away from a fully glazed façade. Although the windows contained a transmission of 54 % and had blinds to block direct sunlight, the daylight levels are much higher than what is expected in Singaporean buildings. In Singapore, the maximum transmission rate for windows is relatively lower (45 %),[13] and solar shading devices (e.g. fins, light-shelves, screens) are recommended in standards.[14]

Since up to 60 % of the electrical energy used in Singaporean office buildings are for space cooling purposes,[15] these criteria help reduce direct solar gains through windows all year around. But as an additional consequence, buildings in Singapore also may not receive high vertical illuminances similar to those found in Wienold and Christoffersen.[3]

Figure 1 plots data collected in 11 different Singaporean buildings from the study by Jakubiec *et al.*,[16]. When plotting calculated DGP from the workspace of occupants, all 147 evaluations collected would be denoted as "Imperceptible glare" (DGP= [0, 0.33]).[12] However, when analysing the percentages of visual discomfort, whereby occupants had indicated that the glare source had originated from a daylit window (as report on a secondary categorical scale),[17] this showed that a majority of the subjective ratings – 75 % (i.e. noticeable, disturbing and intolerable glare) – indicate that occupants still perceived glare.
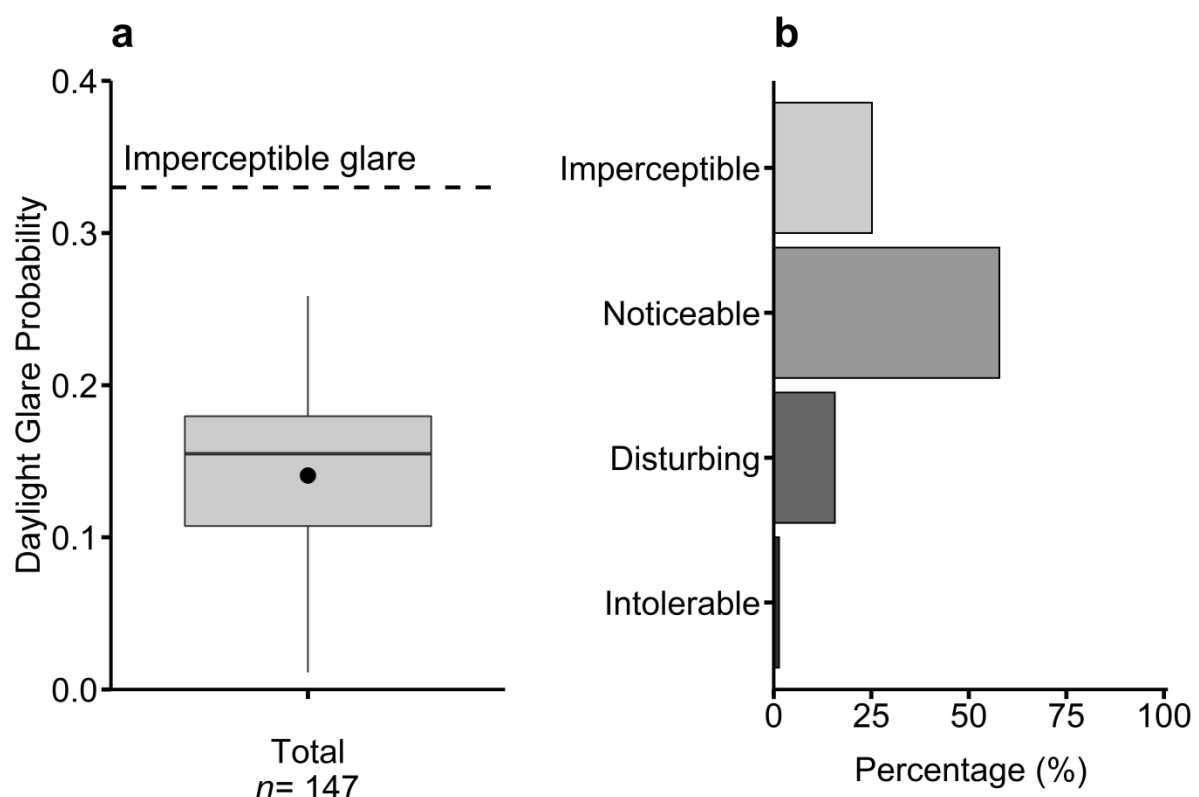

Figure 1 Data collected in the study by Jakubiec *et al.*[16] (a) Boxplot of DGP values calculated across 11 different buildings in Singapore. The circle indicates the mean value. (b) Percentages of the corresponding subjective ratings of glare given by occupants on a 4-point scale.

Figure 1 shows that the range of daylit conditions in the test rooms used by Wienold and Christoffersen[3] may not be applicable to buildings in Singapore. In fact, Pierson *et al.*[18] showed that test room settings generally produced higher vertical illuminances than field studies, which had collected measurements inside occupied buildings. The authors[18] concluded that stimulus range differences across different environments may have biased the ratings given by observers since in field studies with lower vertical illuminance levels, building occupants expressed discomfort at lower DGP thresholds than in test room studies.

Table 3 Studies of daylight glare showing the measured maximum illuminance, subjective criterion, and corresponding value of DGP.

Study	Max. illuminance (lx)	Rating	DGP
Wienold and Christoffersen[3] and Wienold[12]	~12,000	Imperceptible	0.33
Hirning *et al.*[19]	2,354	Glare reported	0.28*
Van Den Wymelenberg and Inanici[20]	4,816	Just Uncomfortable	0.23
Kent *et al.*[21]	10,643	Just Uncomfortable	0.24

*Approximately the highest value of DGP for observers that reported glare.
Note: DGP first appeared in Wienold and Christoffersen,[3] but its glare thresholds were based on the same data later published in Wienold.[12]

In the study by Wienold and Christoffersen,[3] the approximated maximum illuminance is generally much higher than the other three studies reported in Table 3. What is also apparent is that the mean DGP for the lowest possible threshold of glare (imperceptible) is also much larger than the DGP corresponding to higher ratings of glare in the other studies (i.e., although observers in the three studies[19–21] reported glare, the corresponding DGP values are still lower than what had been calculated for observers that did not reported any glare in the study by Wienold and Christoffersen).[3] While it appears that the DGP in the studies found in Table 3 underestimate the degree of discomfort due to glare, we suspect that the range of illuminance levels during the period where observers provided their ratings were not comparable. In other words, when the illuminances are generally higher in one study, the calculated DGP was also larger for the same rating of glare given by the observers. Although we believe range bias is the most candidate reason behind the inconsistencies shown in Table 3, it effects are brought on by dissimilar conditions (e.g. climate, window size, shading systems, etc.), which influence the range of vertical illuminance.

In the study by Kent *et al.*,[21] the maximum reported illuminance is 10,643 lx. Although this value is much higher compared to what had been reported in Van Den Wymelenberg and Inanci,[20] the DGP for the same glare rating in both studies are relatively similar. When examining the mean illuminances in both studies, these were 1,074 lx[21] and 1,467 lx,[20] thereby demonstrating that the maximum illuminance found in Kent *et al.*[21] is an outlier and may not be a reliable representation of the upper range of illuminances found in this study. In general, it difficult to ascertain the presence of a range bias effect from the literature alone and warrants further investigation using more comprehensive data to confirm its presence amongst the studies found in Table 3.

Since past studies tend to utilise different glare scales, quantities (e.g. illuminances, DGP, etc.), and/or statistical values, it is difficult to evaluate these method biases using only the available literature. We propose to evaluate the range bias effect using data collected using uniform procedures across different buildings in our own study. We used the data measured in the study by Jakubiec *et al.*,[16] whereby ratings of discomfort due to glare and vertical illuminances (not consistently reported in the aforementioned study) collected in different buildings found in Singapore were analysed.

Our aim was to determine whether the range of vertical illuminances inside buildings biases the evaluations given to the same criteria of discomfort due to glare when the category rating method is used. In other words, if vertical illuminances are higher at the point in which glare ratings are given in a building, this would increase the vertical illuminance for the same rating of visual discomfort compared to another building with lower vertical illuminances. This study aimed to determine whether the range effect could generally be found when visual discomfort was evaluated by occupants (i.e., we do not aim to establish if the glare source originated from artificial or daylit sources).

2. Method

2.1. Data collection

Although we initially considered both DGP and vertical illuminance as the dependent variable, they produced similar results, and to simplify the interpretation of the results, we reported only the vertical illuminance (results of our DGP analysis can be found in the Supplemental Materials). Since we used the vertical illuminance, we did not differentiate where the glare source had originated from (e.g. artificial light or window) within the occupants' visual field.

Ten Singaporean office spaces were evaluated in this study, and a rotatable office-like test-room located at the top of a building. These are the same buildings that were measured in the study by Jakubiec *et al.*,[16] but our study features ratings of discomfort due to glare and measurements of DGP that were not previously used. The test-room was classified as a building. Since it was expected that the test-room would produce higher vertical illuminances, this was included in our analyses. All the spaces that were evaluated were lit by both natural and artificial light. In total, 592 occupants from the 11 different building spaces gave subjective assessments using the category rating method, and photometric measurements were collected at the same time.[2] The sample size (*n*) (i.e., the number of occupants that gave glare evaluations) and percentage distribution of age, gender and eyewear are showed in Table 4 for the occupants that evaluated the 11 different buildings.

Table 4 Total number of occupants and their percentage distribution according to age, gender and eyewear.

Demographic feature	*n*	%
Age		
Less than 20	5	1
21-30	186	31
31-40	227	38
41-50	111	19
51-65	58	10
Greater than 65	4	1
Gender		

Female	286	48
Male	305	52
Eyewear		
Contacts	72	12
Glasses	299	51
None	220	37

*One occupant did not provide their demographic information

Evaluations were provided using subjective rating scales, whereby occupants were asked to rate the visual scene using a 4-point scale containing semantics that described in – an ascending order – the degree of discomfort due to glare: Imperceptible, noticeable, disturbing and intolerable.[3] Before providing their evaluations, occupants were given a succinct description of the visual phenomenon that was to be evaluated: "*Glare is physical discomfort caused by excessive light, bright reflections or contrast*". Once a glare rating had been given by an occupant, a high-dynamic range (HDR) image was captured using a camera with a fisheye lens, and vertical illuminance values were calculated using the software *Evalglare*.[22] Occupants were only required to provide a single rating of discomfort due to glare. To verify the integrity of images captured, an independent vertical illuminance measurement was taken and later compared to the vertical illuminance calculated from the calibrated HDR images. As formerly reported by Quek and Jakubiec,[23] the differences – for the same data we have analysed in this study – show that across these two sets of vertical illuminance measurements show a slight negative bias (logRMSE= 9.6% and mean bias error= -21.9 lx).

Images were collected using the method described in Jakubiec *et al.*[16] utilising recommendations,[24,25] which can improve the accuracy in which photometric measurements are collected through the use of cameras. For example, 16 low dynamic range photographs were captured at exposures ranging from 4 s to 1/8000 s to compose each HDR photograph in order to minimise luminous overflow or underflow; luminance and illuminance were measured before and after each HDR sequence to identify rapidly changing lighting conditions; and a moderate aperture (f/11) was selected to reduce lens flare while maintaining a sufficient dynamic range. Luminance measurements were used to calibrate each HDR image, and illuminance measurements were used to validate the calibrated accuracy. The measurements (i.e. glare ratings and photometric measurements) correspond to the degree of discomfort glare occupants experienced within their visual scene at their workstation.
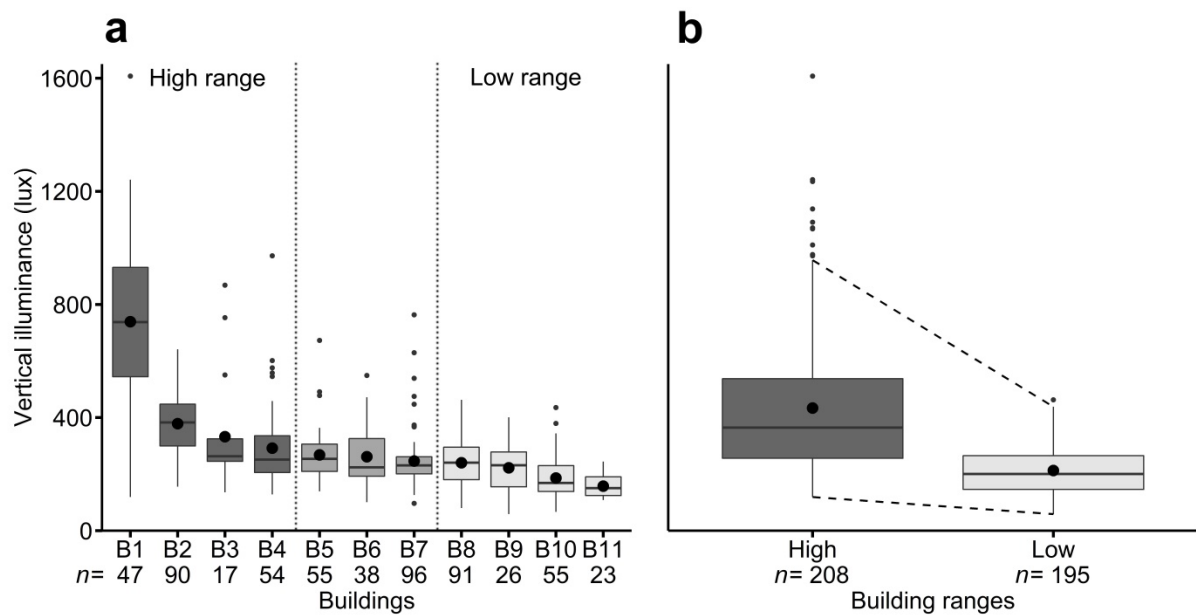
Figure 2 Boxplots displaying the vertical illuminance values calculated from the viewing position of the occupants showing: (a) the vertical illuminances in each of the 11 buildings (and the number of occupants in each) ordered according their means, and (b) the two new formed groups "high" and "low" (and their sample sizes). Note: in plot (b), the Tukey upper and lower whiskers[26] have been interpolated (using dashed lines) across the building ranges.

Figure 2 shows the 11 buildings in Table 4 ordered (plot (a)) according their mean vertical illuminances. Although the building illuminances could have been ranked by their maximum values, this process is not consistent with the ordering of the buildings according to their minimum values and would have also been influenced by outlier and extreme points. For these reasons, we believed that ordering the buildings by their mean was a more straightforward approach. Because the sample sizes across the buildings were vastly different and to ensure there were a sufficient number of ratings on the 4-points scale across comparisons, we combined the buildings into two groups. The two groups were denoted as "high" consisting of the buildings B1, B2, B3 and B4, and "low" containing the buildings B8, B9, B10 and B11. Plot (b) shows the vertical illuminances in the two groups and their sample sizes. To visualise the difference in illuminance range across the building groups, we interpolated the Tukey upper and lower whiskers. Since the whiskers extend no further than 1.5 multiplied by the interquartile range from the upper (to the largest value) and lower (to the smallest value) hinges (i.e. 75th and 25th percentiles respectively),[26] outliers and extreme values were excluded. To help ensure reliable inferential analyses, buildings B5, B6 and B7 were not included to help provide a balanced sample sizes (i.e., approximately the same occupants) and buildings (i.e. four) across the two groups. Since Table 5 showed that there were only four occupants that reported intolerable glare, analyses across the two range groups were not considered for this criterion of visual discomfort.

Table 5 presents the descriptive statistics and number of glare ratings provided on the 4-point scale across the two building range groups. The mean, maximum and minimum values show that vertical illuminances under the high range are larger than the low range group. Across the two building range groups, the number of evaluations given to each criterion on the 4-point scale was relatively similar. When analysing the ratings given across the 4-point scale, the percentages also appear to be relatively equal across the two groups, whereby the

most utilised criterion being noticeable glare and the least used criterion being intolerable glare.

Table 5 Descriptive statistics for the vertical illuminances presenting the mean, standard deviation, maximum and minimum (range), and the total number of glare ratings given to each criteria of discomfort due to glare (and percentage across the 4-point scale) contained in the high and low building range groups.

Building range group	High	Low
Illuminance (lx)		
Mean	434	213
Std dev	250	84
Maximum	1,608	463
Minimum	119	59
Glare rating (n (%))		
Imperceptible	71 (34%)	79 (41%)
Noticeable	100 (48%)	99 (52%)
Disturbing	35 (17%)	13 (7%)
Intolerable	3 (1%)	1 (0%)

2.2. Statistical analyses

To compare the vertical illuminances across the building range groups for the same criterion of discomfort due to glare, we used frequentist approaches (i.e. *p*-values)[27] to determine whether the differences across the high and low range conditions were statistically significant. Because the occupants in each building space gave a single evaluation of discomfort using the two glare scales, between-subject analyses were utilised in this study.[28] To interpret the outcome of the analyses (i.e. *p*-values) derived from the tests, we used the threshold recommended by Benjamin *et al.*,[29] which recommends that values below an alpha-level of 0.005 are denoted as statistically significant. Values above 0.005 were considered not statistically significant (n.s.). Since we had reasons to believe that when a higher stimulus range is available, this increases the photometric values given to the same criterion of discomfort due to glare (Table 1),[8] we used one-tailed (directional) hypotheses.[30,31]

To test the assumption of normality we used Quantile-Quantile plots[32] using the standardised residuals derived from a linear model, which were used to determine if the assumption of normality has been met.[28] In these plots, we overlayed Kolomogorov-Smirnov[33] confidence bands about the datapoints. Beside Quantile-Quantile plots, we also used Lillifers test (Kolmogorov-Smirnov)[34] to determine if the distribution derived from our data was significantly different from a distribution which is normal about its mean. If the assumption of normality was met, we used the independent samples *t*-test.[28] To ensure that the spread across the building range groups was approximately equal, we used the Levene's test of homogeneity of variance.[35] When the assumption of normality was not met, we used

the Wilcoxon-Mann-Whitney test.[36,37] The non-parametric Ansari-Bradley test[38] was also used to ensure the dispersion of the data across the two building range groups were similar. Since frequentist tests are dependent on both the size of the effect and sample,[39] emphasis in the analyses was placed on the effect size.[40] For the independent samples t-test and Wilcoxon-Mann-Whitney test, we used the Pearson's, r, which was calculated according to either equation (2) or (3):[41]

$$r = \frac{Z}{\sqrt{n}}$$

(2)

$$r = \sqrt{\frac{t^2}{t^2 + df}}$$

(3)

whereby, t is the test statistic and df is the degrees of freedom from the independent samples t-test, and z is the test statistic and n is the number of occupants from the Wilcoxon-Manny-Whitney test. To avoid abstract self-interpretation of the outcome, we used the tables given by Ferguson,[42] which provides benchmarks for "small", "moderate" and "large" effect sizes ($r \geq 0.20$, 0.50 and 0.80, respectively). Values below 0.20 were denoted as not representing any practical relevance ("negligible").

3. Results

3.1. Range effects

Figure 3 displays the results of the Quantile-Quanitle plots, whereby these show that some of the datapoints depart from the diagonal line that represents a normal distribution, which suggests the distribution derived from our data is not normal about its mean. This can be seen in plots for: (a) imperceptible glare, (b) noticeable glare, and (c) disturbing glare. Since too few evaluations for intolerable glare were given (Table 5), this criterion was not included. To verify these observations, Table 6 shows that the Lillifers test was statistically significant ($p < 0.005$) for each glare rating. Based on these analyses, this showed that the sample distributions were significantly different from a distribution that was normal about its mean.
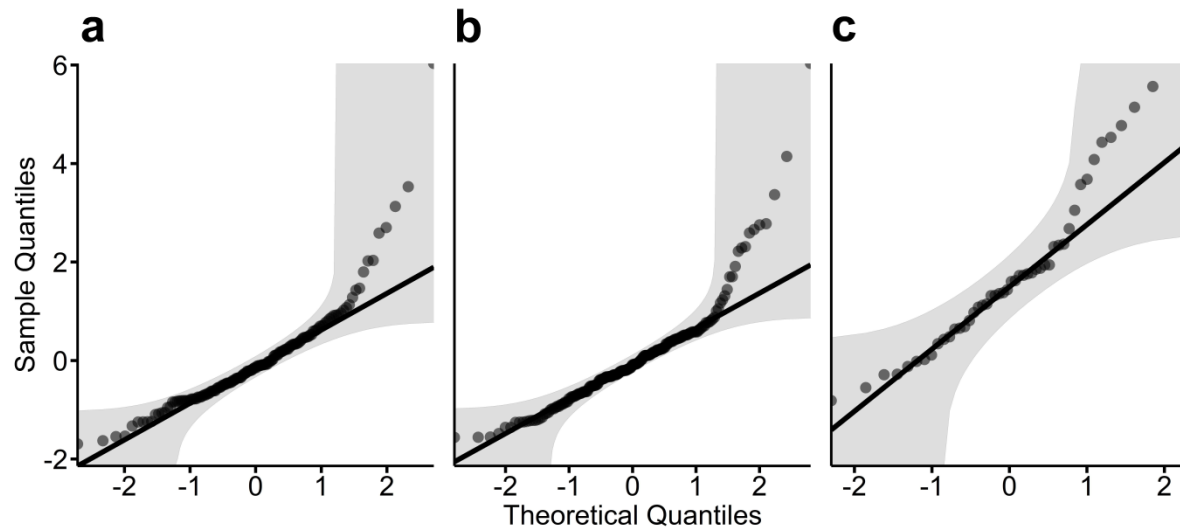
Figure 3 Quantile-quantile plots comparing the theoretical quantiles against the sample quantiles for: (a) imperceptible glare, (b) noticeable glare, and (c) disturbing glare. Note: the diagonal line represents a sample distribution that is normal about its mean.[28]

Since the Quantile-Quantile plots (Figure 3) and inferential (Lilliefors test (Table 6)) analyses showed that the sample distributions were not normal about their mean, we used the Wilcoxon-Mann-Whitney test to compare the differences in vertical illuminance for the same glare rating (i.e., imperceptible, noticeable and disturbing glare) across the two range groups. Table 6 shows that the Ansari-Bradley was not statistically significant ($p > 0.005$) for each glare rating, which suggests that the spread in vertical illuminance were approximately equal across the two building ranges. This showed that the assumption of homogeneity of variance associated with the Wilcoxon-Mann-Whitney test had been met.

Table 6 Results of the Lilliefors test used to examine the assumption of normality and Ansari-Bradley test used to verify the assumption of homogeneity of variance for the four different glare ratings.

Glare rating	Lilliefors test		Ansari-Bradley test	
	D	p-value	AB	p-value
Imperceptible	0.11	<0.005*	2,641	0.67 n.s.
Noticeable	0.14	<0.005*	4,687	0.16 n.s.
Disturbing	0.17	<0.005*	485	0.03 n.s.

Figure 4 presents boxplots comparing the vertical illuminances across the two building range groups for the same glare rating: imperceptible (a), noticeable (b) and disturbing (c) glare. For each plot, we also present the median different (ΔM_{dn}), statistical significance (p-value), and effect size (r) and its associated interpretation.
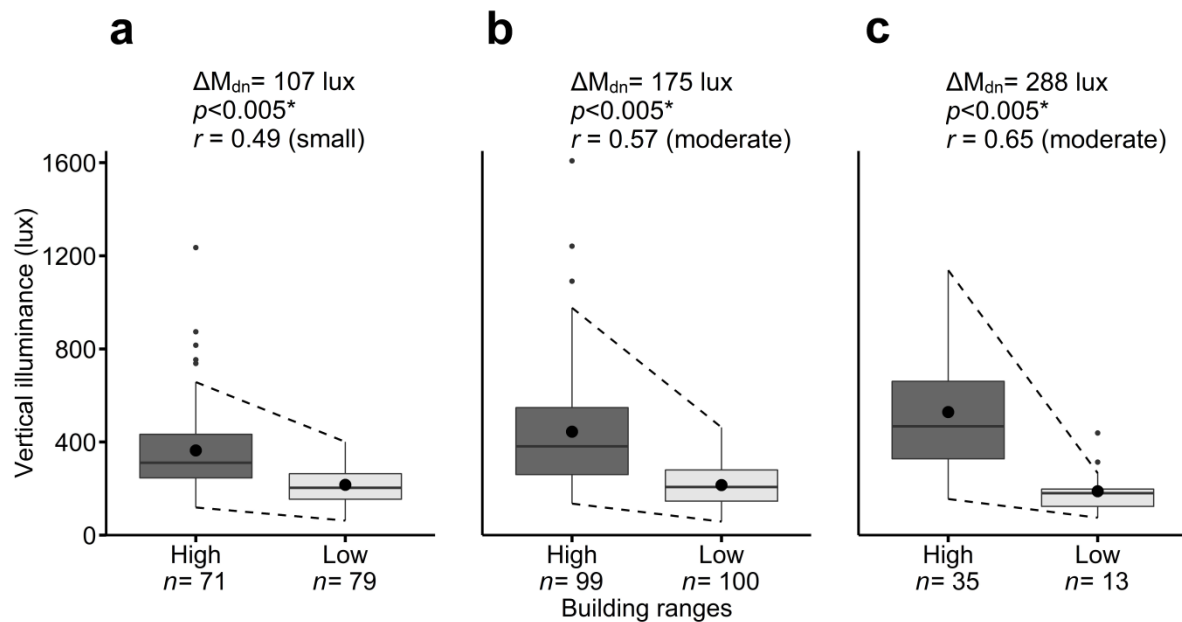
Figure 4 Boxplots comparing the vertical illuminance values across the two building ranges (high and low) with their associated sample sizes for the same glare rating: (a) imperceptible, (b) noticeable, and (c) disturbing glare. Above each plot, we present the descriptive and inferential statistical values.

The results showed that the vertical illuminances are consistently larger for the same rating of glare when considering the high range condition. The differences according to the median difference and effect sizes across the range groups also appear to increase (i.e., the differences in vertical illuminance increases) when occupants reported a higher criterion of glare. These differences were statistically significant ($p<0.005$) across the two building range groups for all three glare ratings shown in Figure 4, and the effect sizes ranged from "small" for imperceptible glare to "moderate" for noticeable and disturbing glare. These results generally showed that when the range in vertical illuminance across the two groups is unequal, this biased the vertical illuminance for the same glare rating.

3.2. High and low building ranges

While Figure 4 shows evidence that range bias influences the vertical illuminance for the same criteria of glare, this analysis compared different groups of buildings against each other. To verify if the same range bias effect persisted across two different buildings only (i.e. conditions that did not contain more than one building), we repeated our analyses – albeit at the expense of a smaller sample size and was restricted to one criterion of glare. We compared B2 and B8, since they had a relatively even number of occupants in each building and had the highest sample sizes in the original high and low range groups (Figure 2), respectively. We denoted these two buildings according to their ranges: namely, B2 (high range) and B8 (low range).

We compared the vertical illuminance for the rating of noticeable glare, because this rating received the highest number of glare ratings compared to the other categories (i.e., imperceptible, disturbing and intolerable glare did not receive sufficient glare ratings to perform inferential tests), and had approximately the same number of evaluations given by

the occupants (Table 7). Since the overall vertical illuminance range (i.e., without considering the subjective ratings (Table 5)) was larger in the building denoted as B2, for the rating of noticeable glare given by occupants, the vertical illuminances are also larger for this building when compared to B8 (Figure 5). For ratings given to noticeable glare, Table 7 shows that the mean vertical illuminances are larger when the building has a higher range.

Table 7 Descriptive statistics for the vertical illuminances presenting the mean, standard deviation, maximum and minimum (range) when occupants provided a rating of noticeable glare, and the total number of ratings contained in the high and low building ranges.

Building range	High (B2)	Low (B8)
Illuminance (lx)		
Mean	377	236
Std. dev.	118	84
Maximum	642	463
Minimum	195	80
Glare ratings (n)		
Noticeable	46	53

Since the Lilliefors test showed that the standardised residuals were not significantly different from a distribution that was normal about its mean (D= 0.08, p-value= 0.08 n.s.), we used parametric tests. The Levene's test of homogeneity of variance revealed that the spread across the two building ranges was also not statistically significant (F= 5.70, p= 0.02 n.s.). Because this provided evidence that the variances across the two building ranges were approximately equal, we used the independent samples *t*-test to compare the mean difference in vertical illuminance for the rating of noticeable glare.

The results of *t*-test show that the differences in vertical illuminance across the two range groups for the rating of noticeable glare were statistically significant (p<0.005). The difference according to the effect size indictor was "moderate". When comparing the vertical illuminances across the two buildings containing different ranges for the rating of noticeable glare, this shows that when a building contains a higher range, the vertical illuminance for the same criterion of visual discomfort increases. This provided the same interpretation when a larger group of buildings were compared across two building range groups (Figure 4).
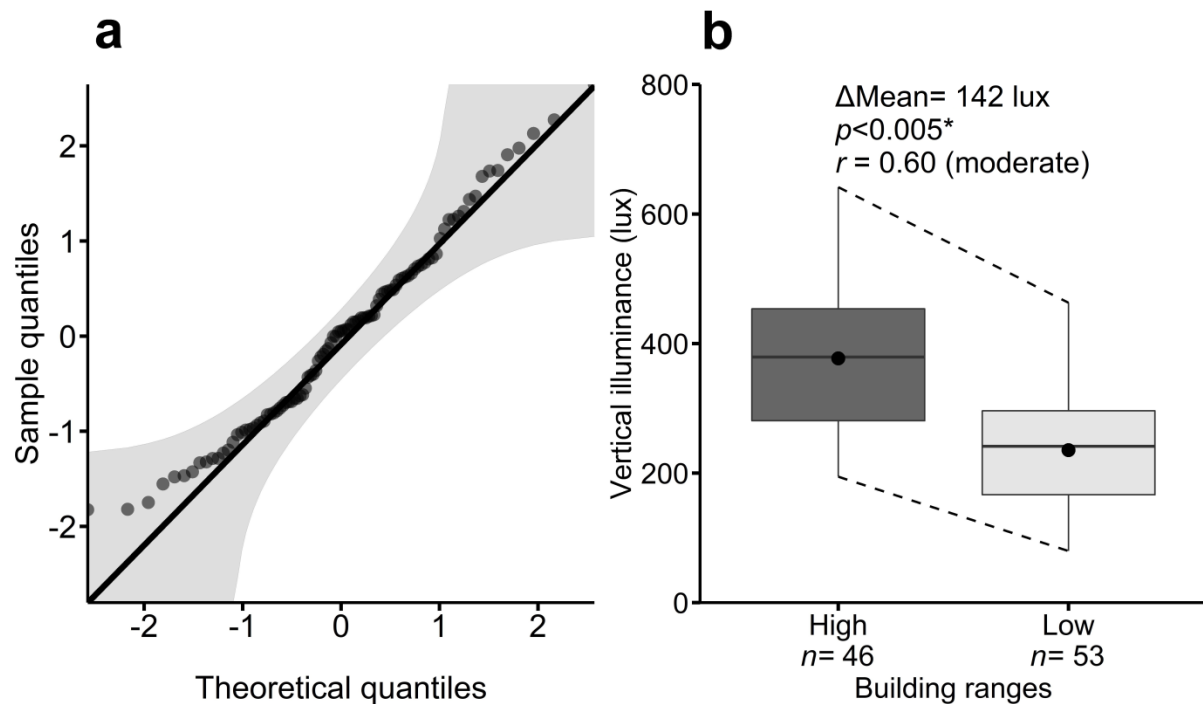
Figure 5 Results when comparing two building with different ranges of vertical illuminance showing: (a) quantile-quantile plots comparing the theoretical quantiles against the sample quantiles for noticeable glare, and (b) boxplots comparing the vertical illuminance values across the two building ranges (high and low) for the rating noticeable glare.

4. Discussion

The results from this study generally show supportive evidence that there is range bias when the category rating procedure is used to evaluate discomfort due to glare across different Singaporean buildings. When the range of vertical illuminance increases inside buildings, the vertical illuminance for the same rating of discomfort due to glare also increases. The results generally support previous research, which had identified that the luminance range can bias settings when using an adjustment procedure to evaluate discomfort due to glare.[8] The influence of the effect is brought on when the illuminance range (i.e., both maxima and minima values) across different groups and individual buildings vary, which biases the rating to the same degree of discomfort due to glare given by occupants.

These findings may have important implications on how the category rating procedure has been used to develop existing glare prediction models. For example, when DGP is used to evaluate daylit glare across different buildings, we believe that the range of photometric conditions must be approximate to each other otherwise incomparable estimates of the visual discomfort will be given. In the original study used to develop DGP,[3] high levels of vertical illuminance were produced in two European countries (i.e. Denmark and Germany) with the windows always facing toward the Sun. In our study, such high levels of vertical illuminance were not available. When DGP is used to evaluate glare in Singaporean buildings, we found DGP incorrectly classified 75 % the glare ratings given by building occupants that perceived from daylit windows (Figure 1). This result may also explain why DGP values for the ratings of discomfort due to glare found in Table 3 appear to show that glare was underestimated. In

other words, the range of vertical illuminance used to develop DGP was significantly higher than the range of daylight conditions found in these buildings.

In a review of different lighting studies that use the category rating procedure, Fotios[4] also highlighted issues of range bias. While our work generally shows similar results to those that had been reviewed (i.e., subjective ratings given by observers are dependent on the stimuli range), Fotios[4] also explains the wider impacts this bias can have on studies using different ranges. In the context of our work, DGP thresholds recommended by Wienold and Christoffersen[3] (Table 2) are based predominantly upon the range of vertical illuminances present in their test-rooms. If a second study uses higher vertical illuminances, DGPs for the same thresholds would also be higher and recommendations will then be given to ensure that glare is kept below those new limits. As explained by Fotios,[4] the findings from these studies will not converge and a likely cause is range bias.

In our study, we did not distinguish between glare sources originating from artificial or daylit sources. Although this allowed us to evaluate a more general effect of range bias across different buildings, when we compared the magnitude of glare for increasing levels of visual discomfort (Figure 4), the vertical illuminance did not increase, or did not substantially increase with each criterion step-change on the 4-point scale. An unexplored facet that may explain this could have been the relationship between window view and glare. In general, views from the windows predominantly showed urban landscapes (i.e. surrounding buildings, roads and pavements). However, due to differences in office layout and furniture arrangement in the buildings we assessed, this unavoidably meant that viewing proximity to the window varied. Although Tuaycharoen and Tregenza[43] and our previous work[44] both showed that more "interesting" view content increases tolerance to window glare, ratings of visual discomfort were given under fixed viewing positions. Since occupants in our study sat at different distances from the window, it is unclear how the window view may have influenced our results. Nevertheless, occupants sat closer to the window and had more access to a view may react differently to an equivalent amount of physical glare seen further away from the window. These differences may help explain why vertical illuminances did not substantially (or at all) increase when occupants gave higher ratings of visual discomfort.

Our study also questions whether ratings of discomfort due to glare are influenced by prior knowledge of the building space (i.e., expected (day)lighting levels) or if they are dependent on exposures to recent stimuli.[45] Studies of vision and colour hue have revealed that prior knowledge of a visual stimulus can influence the final ratings given by the observer.[46,47] Such experiences in a field study could also occur when occupants have spent enough time inside a building, and this could influence their expectations of what are reasonable levels of (day)light and what conditions cause discomfort. Since this behaviour is heuristic, this biases the ratings when the category rating is used to evaluate discomfort due to glare in different buildings.

Since DGP and vertical illuminance are correlated to each other, this meant that both variables generally produced the same results. In fact, Wienold *et al.*[48] showed that among 22 indicators of glare, DGP (ranked first) and vertical illuminance (ranked second) also produced similar results. This highlights a discussion point previously raised by Fotios and Wienold,[49] which questions whether the additional complexity of DGP is necessary when its precision may not be significantly greater than the vertical illuminance. Therefore, designers

may opt to use a much simpler measurement, and reinforces the idea raised by Kent and Fotios,[50] who questioned whether the use models containing precise values are necessarily to evaluate glare.

4.1. Recommendations

To overcome such range effects in future research, we propose some recommendations. When comparing results across different glare studies, the range of conditions that were available when ratings were given by the occupants should be compared. If the ranges are considerably different (e.g., the maximum illuminance value in one building does not overlap with the minimum limit in the other), this would cause an unavoidable bias, which would lead to different estimates for the same criterion of visual discomfort. In situations when the ranges for a parameter (e.g. vertical illuminance) are different but may still overlap, the range bias could be minimised by matching specific parts of the dataset.

To minimise the range effect in glare studies, a pre-trial demonstration could also be used. A study by Kent and Fotios[51] showed that when observers were asked to provide ratings to the same glare source, evaluations of discomfort were higher when a pre-trial demonstration of the minimum and maximum luminances is shown. While the use of a pre-trial demonstration is more practical when the source is artificial (i.e., its luminance can be consistently varied to the same minimum and maximum luminances across different conditions), it may help improve the consistency of ratings when using the category rating.

4.2. Limitations

While we believe that we utilised a sufficient sample size to perform a range of analyses to confirm our aim, it was not feasible to do this for all conditions that we would have ideally liked to have evaluated. In the data we used, there were too few ratings given the criterion of "Intolerable" glare to perform inferential analyses. Therefore, our tests were mostly restricted to evaluations made to the lower ratings of discomfort glare in Figures 4. Although we think that the effects would have been present for highest degree of discomfort glare – assuming a sufficient number of ratings had been given, it was not possible to confirm this due to uneven amount of evaluations given on the 4-point scale.

We think some questions may also need to be raised to whether the criteria used on the 4-point scale[3] accurately describe typically levels of discomfort due to glare found in office buildings. Since occupants mostly used the lower thresholds of discomfort, higher threshold of glare might rarely occur due to buildings that are able to better control excessive levels of daylight that enter through the windows and have other means of glare protection.

5. Conclusions

In our study, we found that the range bias can effect evaluations of visual discomfort when the ranges of vertical illuminance present across different buildings in Singapore vary. The

range effect persisted when evaluating different magnitude ratings of discomfort due to glare. This showed that the range effect was larger when the magnitude of discomfort due to glare increased. When the different groups and individual buildings were compared against each other, the range effect could still be found. The results of this study raise important questions when the category rating procedure is used to evaluate discomfort due to glare.

We think that the range effect can be used to understand the discrepancies found when glare is evaluated across different buildings using the category rating procedure. In Singapore, we found that DGP could not correctly classify a large majority (75 %) of the glare ratings given by building occupants that perceived visual discomfort from daylit windows. We think one of the reasons for this is that the range of vertical illuminances used in the original DGP study was much higher than those present in Singaporean buildings.

As part of ongoing research, we aim to develop new approaches that could minimise the effects of range bias in future glare studies. Using recognised statistical procedures, this will match the closest values (e.g. vertical illuminance) in one dataset with the range of values found in other dataset(s). Assuming that the datasets are large enough and have some common values that approximate each other, we think this may help control for unwanted range bias effects.

Declaration of conflicting interests

The authors declare that there is no conflict of interest.

Funding

References

1. Van Den Wymelenberg KG. Visual comfort, discomfort glare, and occupant fenestration control: Developing a research agenda. *LEUKOS* 2014; 10: 207–221.

2. Velds M. User acceptance studies to evaluate discomfort glare in daylit rooms. *Solar Energy* 2002; 73: 95–103.

3. Wienold J, Christoffersen J. Evaluation methods and development of a new glare prediction model for daylight environments with the use of CCD cameras. *Energy and Buildings* 2006; 38: 743–757.

4. Fotios S. Using category rating to evaluate the lit environment: Is a meaningful opinion captured? *LEUKOS* 2019; 15: 127–142.

5. Poulton EC. *Bias in quantifying judgements*. Hove and London, United Kingdom: Lawrence Erlbaum Associates, 1989.

6. Poulton EC. Quantitative subjective assessments are almost always biased, sometimes completely misleading. *British Journal of Psychology* 1977; 68: 409–425.

7. Fotios S, Kent M. Measuring discomfort from glare: Recommendations for good practice. *LEUKOS* 2020; 1–21.

8. Kent MG, Fotios S, Cheung T. Stimulus range bias leads to different settings when using luminance adjustment to evaluate discomfort due to glare. *Building and Environment* 2019; 153: 281–287.

9. Logadóttir Á, Christoffersen J, Fotios SA. Investigating the use of an adjustment task to set the preferred illuminance in a workplace environment. *Lighting Research and Technology* 2011; 43: 403–422.

10. Fotios SA, Cheal C. Stimulus range bias explains the outcome of preferred-illuminance adjustments. *Lighting Research and Technology* 2010; 42: 433–447.

11. EN 17037. *Daylight in buildings*. Brussels, Belgium: European Committee for Standardization, 2018.

12. Wienold J. *Daylight glare in offices*. Germany: Fraunhofer-Verlag, 2010.

13. BCA Green Mark. *Green Mark: Non-residential buildings NRB: 2015. Technical guide and requirements.* Singapore: Building and Construction Authority, 2016.

14. Singapore, Building and Construction Authority. *Building planning and massing*. Singapore: Centre for Sustainable Buildings and Construction, Building and Construction Authority, 2010.

15. BCA. *Super low energy building: Technology roadmap*. Singapore: Building Construction Authority, https://www.bca.gov.sg/GreenMark/others/SLE_Tech_Roadmap.pdf (2018, accessed 1 June 2019).

16. Jakubiec J, Quek MArch G, Srisamranrungruang T. Long-term visual quality evaluations correlate with climate-based daylighting metrics in tropical offices – A field study. *Lighting Research and Technology* 2021; 53: 5–29.

17. APA. Categorical scale – APA dictionary of psychology. *APA dictionary of psychology*, https://dictionary.apa.org/categorical-scale (2020, accessed 8 December 2020).

18. Pierson C, Sarey Khanie M, Bodart M, Wienold J. Discomfort glare cut-off values from field and laboratory studies. In: *Proceedings of the 29th Quadrennial Session of the CIE*. CONF, Washington DC, USA: Commission Internationale de L'Eclairage, pp. 295–305.

19. Hirning MB, Isoardi GL, Cowling I. Discomfort glare in open plan green buildings. *Energy and Buildings* 2014; 70: 427–440.

20. Wymelenberg KVD, Inanici M. A critical investigation of common lighting design metrics for predicting human visual comfort in offices with daylight. *LEUKOS* 2014; 10: 145–164.

21. Kent MG, Altomonte S, Wilson R, Tregenza PR. Temporal effects on glare response from daylight. *Building and Environment* 2017; 113: 49–64.

22. Wienold J. *Evalglare*. Freiburg, Germany: Fraunhofer Institute for Solar Energy Systems, https://www.radiance-online.org/learning/documentation/manual-pages/pdfs/evalglare.pdf/view (2016).

23. Quek G, Jakubiec JA. Calibration and validation of climate-based daylighting models based on one-time field measurements: Office buildings in the tropics. *LEUKOS* 2021; 17: 75–90.

24. Jakubiec JA, Van Den Wymelenberg K, Inanici M, Mahic A. Accurate measurement of daylit interior scenes using high dynamic range photography. In: *Lighting Quality & Energy Efficiency*. Melbourne, Australia: International Commission on Illumination, 2016.

25. Jakubiec JA, Van Den Wymelenberg K, Inanici M, Mahic A. Improving the accuracy of measurements in daylit interior scenes using high dynamic range photography. In: *36th International Conference on Passive and Low Energy Architecture. Cities, Buildings, People: Towards Regenerative Environments*. Los Angeles, CA, USA, 11-13 July, 2016.

26. Tukey JW. *Exploratory data analysis*. Addison-Wesley Publishing Company, http://archive.org/details/exploratorydataa00tuke_0 (1977, accessed 11 December 2020).

27. Fisher RA. 048: The arrangement of field experiments. *Journal of the Ministry of Agriculture of Great Britain* 1926; 33: 503–513.

28. Field A, Miles J, Field Z. *Discovering statistics using R*. London, United Kingdom: SAGE Publications, 2012.

29. Benjamin DJ, Berger JO, Johannesson M, Nosek BA, Wagenmakers E-J, Berk R, et al. Redefine statistical significance. *Nature Human Behaviour* 2018; 2: 6–10.

30. Hauschke D, Steinijans VW. Directional decision for a two-tailed alternative. *Journal of Biopharmaceutical Statistics* 1996; 6: 211–218.

31. Ruxton GD, Neuhäuser M. When should we use one-tailed hypothesis testing? *Methods in Ecology and Evolution* 2010; 1: 114–117.

32. Hyndman RJ, Fan Y. Sample quantiles in statistical packages. *The American Statistician* 1996; 50: 361–365.

33. Smirnov N. Table for estimating the goodness of fit of empirical distributions. *The Annals of Mathematical Statistics* 1948; 19: 279–281.

34. Lilliefors HW. On the Kolmogorov-Smirnov test for normality with mean and variance Unknown. *Journal of the American Statistical Association* 1967; 62: 399–402.

35. Olkin I. *Contributions to probability and statistics*. Stanford University Press, 1960.

36. Mann HB, Whitney DR. On a test of whether one of two random variables is stochastically larger than the other. *The Annals of Mathematical Statistics* 1947; 18: 50–60.

37. Wilcoxon F. Individual comparisons by ranking methods. *Biometrics Bulletin* 1945; 1: 80–83.

38. Ansari AR, Bradley RA. Rank-sum tests for dispersions. *The Annals of Mathematical Statistics* 1960; 31: 1174–1189.

39. Ellis PD. *The essential guide to effect sizes: Statistical power, meta-analysis, and the interpretation of research results*. Cambridge University Press, 2010.

40.   Cohen J. The earth is round (p $<$ .05). *American Psychologist* 1994; 49: 997–1003.

41.   Field A, Hole GJ. *How to design and report experiments*. London, United Kingdom: SAGE Publications, 2011.

42.   Ferguson CJ. An effect size primer: A guide for clinicians and researchers. *Professional Psychology: Research and Practice* 2009; 40: 532–538.

43.   Tuaycharoen N, Tregenza PR. View and discomfort glare from windows. *Lighting Research and Technology* 2007; 39: 185–200.

44.   Kent MG. *Temporal effects in glare response*. Doctoral Thesis, University of Nottingham, http://eprints.nottingham.ac.uk/35450/ (2016).

45.   Jakubiec JA, Reinhart CF. A concept for predicting occupants' long-term visual comfort within daylit spaces. *LEUKOS* 2016; 12: 185–202.

46.   Adams WJ, Graf EW, Ernst MO. Experience can change the 'light-from-above' prior. *Nature Neuroscience* 2004; 7: 1057–1058.

47.   Olkkonen M, McCarthy PF, Allred SR. The central tendency bias in color perception: effects of internal and external noise. *Journal of Vision* 2014; 14: 1–15.

48.   Wienold J, Iwata T, Sarey Khanie M, Erell E, Kaftan E, Rodriguez R, et al. Cross-validation and robustness of daylight glare metrics. *Lighting Research and Technology* 2019; 51: 983–1013.

49.   Fotios S, Wienold J. Correspondence: Discussion of 'The cross validation and robustness of daylight glare metrics'. *Lighting Research and Technology* 2020; 52: 314–317.

50.   Kent MG, Fotios S, Altomonte S. Discomfort glare evaluation: The influence of anchor bias in luminance adjustments. *Lighting Research and Technology* 2017; 51: 131–146.

51.   Kent MG, Fotios SA. The effect of a pre-trial range demonstration on subjective evaluations using category rating of discomfort due to glare. *LEUKOS* 2021; 17: 43–58.