



This is a repository copy of *The effect of a pre-trial range demonstration on subjective evaluations using category rating of discomfort due to glare.*

White Rose Research Online URL for this paper:  
<http://eprints.whiterose.ac.uk/148964/>

Version: Published Version

---

**Article:**

Kent, M.G. and Fotios, S. [orcid.org/0000-0002-2410-7641](https://orcid.org/0000-0002-2410-7641) (2020) The effect of a pre-trial range demonstration on subjective evaluations using category rating of discomfort due to glare. LEUKOS, 17 (1). pp. 43-58. ISSN 1550-2724

<https://doi.org/10.1080/15502724.2019.1631177>

---

**Reuse**

This article is distributed under the terms of the Creative Commons Attribution (CC BY) licence. This licence allows you to distribute, remix, tweak, and build upon the work, even commercially, as long as you credit the authors for the original work. More information and the full terms of the licence here:  
<https://creativecommons.org/licenses/>

**Takedown**

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing [eprints@whiterose.ac.uk](mailto:eprints@whiterose.ac.uk) including the URL of the record and the reason for the withdrawal request.



[eprints@whiterose.ac.uk](mailto:eprints@whiterose.ac.uk)  
<https://eprints.whiterose.ac.uk/>

# The Effect of a Pre-Trial Range Demonstration on Subjective Evaluations Using Category Rating of Discomfort Due to Glare

Michael G Kent & Steve Fotios

To cite this article: Michael G Kent & Steve Fotios (2021) The Effect of a Pre-Trial Range Demonstration on Subjective Evaluations Using Category Rating of Discomfort Due to Glare, LEUKOS, 17:1, 43-58, DOI: [10.1080/15502724.2019.1631177](https://doi.org/10.1080/15502724.2019.1631177)

To link to this article: <https://doi.org/10.1080/15502724.2019.1631177>



© 2019 The Author(s). Published with license by Taylor & Francis Group, LLC.



Published online: 23 Jul 2019.



Submit your article to this journal [↗](#)



Article views: 550



View related articles [↗](#)



View Crossmark data [↗](#)



Citing articles: 3 View citing articles [↗](#)

# The Effect of a Pre-Trial Range Demonstration on Subjective Evaluations Using Category Rating of Discomfort Due to Glare

Michael G Kent <sup>a,b</sup> and Steve Fotios <sup>c</sup>

<sup>a</sup>Department of Architecture and Built Environment, University of Nottingham, University Park, Nottingham, UK; <sup>b</sup>Berkeley Education Alliance for Research in Singapore; <sup>c</sup>School of Architecture, University of Sheffield, Sheffield, UK

## ABSTRACT

Category rating is a procedure commonly used to evaluate visual discomfort due to glare. One recommended step for good practise in a category rating procedure is to use a pre-trial demonstration (PTD) of the range of stimuli to be experienced. However, PTD have rarely, if at all, been used in past research on discomfort glare. In this study, two experiments were conducted to test the influence of the PTD on evaluations of discomfort due to glare. In the first experiment participants evaluated four glare source luminances with, and without, a PTD. The results suggest that using a PTD increased the reported degree of discomfort perceived for the same glare setting, although this may depend on the manner in which the PTD is presented. In the second experiment, participants evaluated four glare source luminances using PTD with three different luminance ranges: the results suggest this had significant effect on discomfort evaluations in that evaluations of discomfort were lower when a higher luminance range was used. Along with other recent studies, these findings suggest a need to derive a more robust procedure for measuring the discomfort due to glare.

## ARTICLE HISTORY

Received 29 August 2018  
Revised 6 June 2019  
Accepted 10 June 2019

## KEYWORDS

Discomfort glare; category rating; pre-trial demonstration; experimental bias

## 1. Introduction

Category rating is a widely used procedure for evaluating the effect of changes in lighting on occupants' evaluations of the visual environment. Observers use one or a series of rating scales to quantify the visual scene in terms of brightness, visual clarity, pleasantness and other subjective parameters (Boyce and Cuttle 1990; Flynn and Spencer 1977; Viénot et al. 2009; Vrabel et al. 1998). Like all quantitative subjective evaluations, category rating suffers from many sources of bias which affect the precision and accuracy of the responses that are given. Evaluating a visual environment using a rating scale is difficult, particularly early in an experiment when the observer has not yet seen the range of possibilities, and they must develop their own internal criteria. One suggestion for countering this is to demonstrate the range of stimuli before trials commence, thus to define to observers the meaning of the upper and lower limits of a rating scale, and anchoring the response range to the stimulus range (Adams et al. 2004;

Fotios and Houser 2009; Houser and Tiller 2003; Olkkonen et al. 2014; Tiller and Rea 1992). This is referred to here as pre-trial demonstration (PTD). Response range anchoring through PTD has the potential to reduce variance in the data, or at least to increase the internal consistency of each subject, and to reduce response contraction bias (Poulton 1989).

One topic where category rating is commonly used is the study of discomfort due to glare, the visual discomfort caused by exposure to excessive luminances or luminance contrasts within the field of view greater than that to which the eyes are able to adapt (Vos 2003). Observers assign the degree of discomfort experienced to one of several categories, these labelled with a defined degree of discomfort (Velds 2002).

PTDs were used rarely, if at all, in past studies of discomfort due to glare. Given that PTD is recommended to be good practise, it is desirable to investigate the likely influence of not doing so, because this may affect the recommended

**CONTACT** Michael G Kent  [michaelkent@berkeley.edu](mailto:michaelkent@berkeley.edu)  Department of Architecture and Built Environment, University of Nottingham, University Park, Nottingham, NG7 2RD, United Kingdom

Color versions of one or more of the figures in the article can be found online at [www.tandfonline.com/ulks](http://www.tandfonline.com/ulks).

© 2019 The Author(s). Published with license by Taylor & Francis Group, LLC.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

thresholds for the control of glare. Two experiments were carried out to examine this question. In the first experiment, discomfort due to glare was evaluated with and without PTD to determine whether this affected the subsequent evaluations of discomfort, with the PTD luminance range held constant for all trials. In the second experiment, the lower and upper limits of the PTD luminance ranges were varied.

## 2. Method

### 2.1. Experimental setting

The apparatus used in this work (Fig. 1) was a semi-hexagonal chamber as used in previous studies of discomfort glare (Kent et al. 2017; Tuaycharoen and Tregenza 2005). The interior surfaces (2.7 m in height) were matte-white. A desk with a matte-white surface was placed at the centre of the chamber, on which was placed a flat screen VDU monitor (17" viglen TS700 liquid crystal display, mean self-luminance =  $65 \text{ cd/m}^2$ ), just below the glare source in the rear wall of the chamber. The frame and mount of the VDU were both matte-white, thereby reducing contrast between the VDU and background partition walls. The connection cables from the VDU to the desktop computer were covered with matte-white tape as were the corners of the chamber where the edges met the rear wall.

Background lighting was produced from three 3W LEDs positioned above the visual scene. Luminance measurements were collected from the location of the test participant's eyes using a luminance meter (LS-100, Minolta, Japan – manufacturer's reported accuracy  $\pm 2\% \text{ cd/m}^2$ ) mounted on a tripod. From this position, the mean background luminance was calculated from 16 individual measurements taken on a regular grid symmetrical about a small diffusive screen and extended across the width and height of the cubicle. An additional measurement was taken to record the luminance of the VDU. The mean luminance was held at a constant  $65 \text{ cd/m}^2$  throughout the experiment, as this is within the range of values commonly found in interior spaces (CIBSE 1994). Both the VDU and background lighting produced a correlated colour temperature (CCT) of 4000 K, which was recorded from the viewing position with a calibrated illuminance chromameter (CL-200a, Minolta, Japan, manufacturer's reported accuracy  $\pm 2\% \text{ lux}$  and  $\pm 2\% \text{ CCT}$ ).

The glare source was a small diffusive screen ( $0.08 \times 0.04 \text{ m}^2$ ) made from three sheets of translucent paper and backlit by a computer project that was operated by the experimenter. The paper allowed direct light from the projector to be evenly spread across the area of the screen. The glare source subtended an angle at the participant's eye of 0.009 steradians and could be set to luminances in the range from  $229 \text{ cd/m}^2$  to  $32000 \text{ cd/m}^2$ .

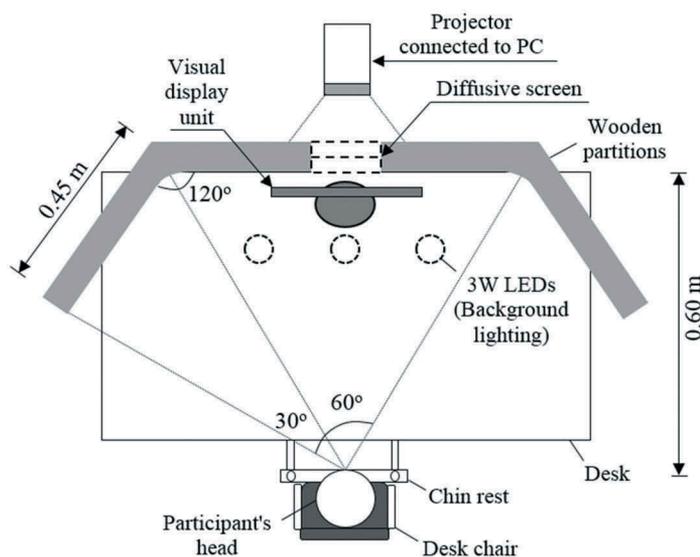


Fig. 1. Plan of the experimental setup and image of the lighting chamber used in this study.

In both experiments, visual fixation was directed towards the VDU, located below the glare source, and upon which was presented a small circle used to draw the participant's visual attention. This was located at centre of the computer screen, at a distance of 20° below the centre of the glare source. A chin rest was used to maintain a constant viewing location throughout all trials. The background luminance, glare source size, and the position of the glare source within the field of vision of the observer were held constant. Only the luminance of the glare source was varied.

In both experiments, discomfort was evaluated at four luminances associated with different levels of visual discomfort, referred here as glare settings (Table 1). These luminances were intended to provide the four levels of visual discomfort as used in Hopkinson's multiple criterion scale (Hopkinson 1960). The appropriate luminances for the current context were determined using the Illuminating Engineering Society Glare Index (IES-GI) (1), originally proposed by the luminance study panel of the IES technical committee (Robinson et al. 1962). Since the VDU was located below the glare source, the position index formula (2) proposed by Luckiesh and Guth (1949) was used to modify the glare index formula for glare sources located below the line of sight (IESNA 2011).

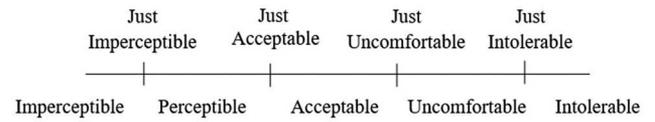
$$\text{IES - GI} = 10 \cdot \log_{10} \cdot 0.478 \sum_{i=1}^n \left( \frac{L_s^{1.6} \cdot \omega_s^{0.8}}{L_b \cdot P_s^{1.6}} \right) \quad (1)$$

$$P = \exp \left[ \left( 35.2 - 0.31889\alpha - 1.22e^{-\frac{2\alpha}{9}} \right) 10^{-3}\beta \right] + (21 + 0.2667\alpha - 0.0029663\alpha^2) 10^{-5}\beta^2 \quad (2)$$

$\alpha$  = the angle from the vertical plane containing the glare source and the line of sight (°)

**Table 1.** The four glare settings at which category rating evaluations were provided. These luminances were determined using the IES-GI and Hopkinson's discomfort glare criteria.

Hopkinson's Criteria	IES-GI	Luminance (cd/m <sup>2</sup> )
Just Imperceptible	10	762
Just Acceptable	16	1799
Just Uncomfortable	22	4122
Just Intolerable	28	9819



**Fig. 2.** Continuous scale used to evaluate the magnitude of discomfort glare sensation.

$\beta$  = the angle between the line of sight from the observer to the glare source (°)

To provide evaluations of discomfort, participants were instructed to place a mark on a 10 cm long continuous scale (Fig. 2) (Altomonte et al. 2016; Tuaycharoen and Tregenza 2007). The scale features Hopkinson's original borderline criteria (e.g., "just imperceptible") above the scale; these were previously used in the development of the IES-GI (Petherbridge and Hopkinson 1950). Underneath the scale are absolute criteria (e.g., "perceptible") as later proposed by Hopkinson when evaluating glare from daylight (Hopkinson 1972).

To evaluate the results given from the linear glare scale with conventional glare indices, the evaluations given on the continuous scale (i.e., the position along the response scale) were scaled to equivalent glare index values following the method proposed by Altomonte et al. (2016). Equivalent values of glare response vote scaled to the IES-GI (GRV (IES-GI)) were obtained, using the data obtained by Hopkinson (1960, 1972) that relate each of the four discomfort glare criteria to corresponding values of IES-GI. These values can be considered suitable to assessing the glare sensation from small artificial light sources. In this study, GRV (IES-GI) were calculated according to equation (3):

$$\text{GRV(IES - GI)} = 6 \cdot (0.39x - 0.39) + 10 \quad (3)$$

Whereby,  $x$  = the distance (cm) along the scale from the left corresponding to the discomfort experienced as indicated by the participant on the continuous glare scale.

## 2.2. Statistical analyses

Null hypothesis statistical significance testing (NHST) was performed to determine if the differences in mean GRV (IES-GI) were significantly different across the independent variable. Emphasis in was placed on the effect size, a standardised measure

of the magnitude of differences examined (Ellis 2010), and not solely on the statistical significance (which, in cases of small sizes, could confound the outcome) (Cohen 1994).

Parametric tests that relied on the assumptions of normality were used to analyse the data. Analyses using graphical (Q-Q plots) and statistical (one-sample Shapiro-Wilk tests) methods were used to check whether data were drawn from a normally distributed population (Field 2013). To test the assumption of sphericity, the Mauchly's test was used to test whether the variances of differences between all paired comparison of the within-subject variable (i.e., the independent variable) were equal (Mauchly 1940). When the assumption of sphericity was not met, the method of Greenhouse-Geisser was used to give a conservative  $F$ -test statistic protected against Type I errors (Greenhouse and Geisser 1959). To test the assumption of equal variance, the Levene's tests were used to check whether variances across groups were not statistically different (Field 2013).

When *Post-hoc* ( $t$ -tests) testing was performed, all combination between the independent variable were compared against each other. To avoid Type III errors (i.e., incorrectly specifying the directionality of the effect) (Shaffer 1995), the directionality of the hypothesis was carefully selected through inspection of central tendencies and graphical displays (Hauschke and Steinijans 1996). Therefore, only when consistent trends (i.e., direct or inverse relationships between variables) were identified, directional (one-tailed) tests were selected. If this was not the case, non-directional (two-tailed) tests were used (Ruxton 2006). Since multiple comparisons were carried out on the same data and with the same hypothesis, Bonferroni corrections were applied to control for the experiment-wise error rate caused by the significance level inflating (Cabin and Mitchell 2000; Shaffer 1995).

The effect size was calculated by making use of equivalence between the observed differences. In the analysis of variance (ANOVA) tests, the effect size was based on the  $\eta_p^2$ , and from the  $t$ -test the Pearson's coefficient,  $r$  (Field 2013). The interpretation of the outcome was derived from the benchmarks provided by Ferguson (2009), whereby values have been given for small, moderate, and strong

effect sizes ( $\eta_p^2 \geq 0.04, 0.25, 0.64$  and  $r \geq 0.20, 0.50, 0.80$ ), respectively. Values lower than the recommended minimum effect size ( $\eta_p^2 < 0.04$  and  $r < 0.20$ ) do not represent a practically significant effect.

### 3. Experiment 1

#### 3.1. Procedure

The aim of the first experiment was to compare discomfort evaluations with and without a luminance range PTD. Thirty-four postgraduate students were recruited, 20 females and 14 males with a mean age of 26 years ( $SD = 5$ ). Twenty-three wore their normal corrective lens during the tests, and all self-certified as having no other health or eye problems.

At the start of the experiment, participants were given a set of instructions, including a definition of discomfort glare, the meaning of the glare criteria anchored onto the continuous glare scale (see Appendix), and an overview of the experimental procedure. Participants adjusted the chair so that it was comfortable to sit with their head on the chin rest.

There were three blocks of trials. One block required evaluations to be given without first seeing demonstration of upper and lower luminances (no-PTD). Two blocks presented low and high luminance settings before trials (with-PTD), one block showing these luminances in ascending order of magnitude (low to high: L-H) and the other block showing them in descending order (high to low: H-L). A quasi-balanced order was adopted, in which the no-PTD block was either the first or third block to be observed. The L-H and H-L blocks were presented in a balanced order. Within each block, the four glare settings (Table 1) were evaluated in a random order.

The glare source luminances presented as the PTD were  $229 \text{ cd/m}^2$  (low setting) and  $12219 \text{ cd/m}^2$  (high settings). The demonstration started by presenting one of these, and then gradually increasing or decreasing the luminance to reach the other setting. As recommended by Poulton (1989) these are slightly beyond the range of glare settings presented in trials (i.e., 762 to  $9819 \text{ cd/m}^2$ ).

Whilst evaluating discomfort, participants were instructed to keep their visual focus on a small circle displayed at the centre of the computer screen, the visual focal point used in previous work (e.g., (Kent et al. 2018a; Petherbridge and Hopkinson 1950; Stone and Harker 1973)). The impact of different visual foci (Kent et al. 2019). For each setting, participants were required to wait for 10 seconds after the luminance was set before providing the evaluation by placing a mark on the glare scale.

### 3.2. Experiment 1: Results

Fig. 3 presents the results of experiment 1. This shows, on the y-axis, the mean GRV (IES-GI) values as calculated from evaluations given using the continuous discomfort response scale. The x-axis groups the mean ratings according to the test condition (four glare settings and the no-PTD and with-PTD conditions). When using a PTD to show the lower and upper limits, this could be done lower limit first (ascending sequence) or upper limit first (descending sequence). In the absence of any recommended practise for PTD, advice for the method of limits (Gescheider 1985) and luminance adjustment (Logadóttir et al. 2011) was followed: both sequences (ascending and descending) were used and the mean of the subsequent results was used as the best estimate. We

refer to the average of the two conditions as with-PTD. For the with-PTD results, these data are thus the average of trials in the L-H and H-L blocks. Fig. 3 shows that for all four glare settings, the with-PTD trials led to higher average GRV (IES-GI) than did the no-PTD trials.

Paired-samples *t*-tests were used to compare the difference in mean GRV (IES-GI) values across the no-PTD and with-PTD conditions for each glare setting. Table 2 presents, for each glare setting, the mean and standard deviations (SD), the mean difference ( $\Delta M$ ) and statistical significance (*p*-value), and the effect size (*r*). Table 2 suggest there are consistent differences in that higher mean GRV (IES-GI) values are seen under the with-PTD condition. That is, participants expressed a greater degree of discomfort on the continuous response scale for the same glare setting. The differences are statistically significant ( $p \leq 0.01$ ) for the Just Imperceptible setting and weakly significant ( $p \leq 0.05$ ) for the Just Intolerable setting but not significant ( $p > .05$ ) for the Just Acceptable and Just Uncomfortable glare settings. The differences in three glare settings demonstrate small effect sizes ( $0.20 \leq r < 0.50$ ) and negligible ( $r < 0.20$ ) for the Just Uncomfortable glare setting. The standard deviations are smaller for the with-PTD evaluations than the no-PTD trials.

These data suggest that showing a PTD had a statistically significant effect on the discomfort evaluations at the lowest (Just Imperceptible) and

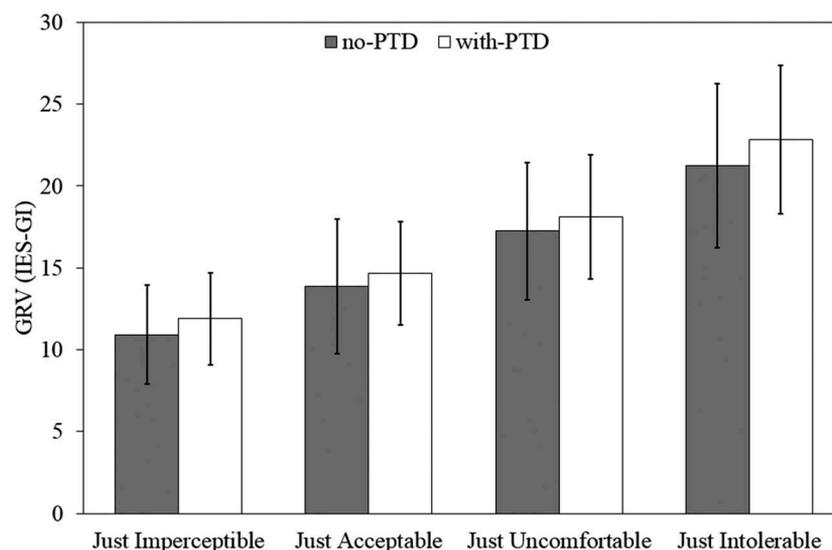


Fig. 3. Results of experiment 1: Mean GRV (IES-GI) for the no-PTD and with-PTD conditions across the four glare settings. Error bars show standard deviation.

**Table 2.** Analysis of experiment 1 results: Paired-samples *t*-tests used to compare no-PTD and with-PTD for each glare setting.

Glare Setting	Mean GRV (standard deviation)		$\Delta M^{\text{NHST}}$	<i>r</i>
	No-PTD	With-PTD		
Just Imperceptible	10.81 (3.00)	11.89 (2.80)	-1.09**	0.40
Just Acceptable	13.80 (4.11)	14.57 (3.14)	-0.77 n.s.	0.19
Just Uncomfortable	17.41 (4.19)	18.04 (3.78)	-0.63 n.s.	0.20
Just Intolerable	21.49 (5.02)	22.74 (4.53)	-1.24*	0.31

\* $p \leq 0.05$ ; \*\* $p \leq 0.01$ ; \*\*\* $p \leq 0.001$ ; n.s. = not significant ( $p > 0.05$ )  
 $r < 0.20$  = negligible;  $0.20 \leq r < 0.50$  = small;  $0.50 \leq r < 0.80$  = moderate;  
 $r \geq 0.80$  = large

highest (Just Intolerable) glare settings. Although the other two glare settings (Just Acceptable and Just Uncomfortable) did not suggest a significant difference, higher mean GRV (IES-GI) values can be seen in all four glare settings (i.e., a greater degree of discomfort for the same glare settings).

Fig. 4 shows the with-PTD trials broken down according to the ascending (L-H) or descending (H-L) sequence. This shows that the sequence had an effect: the mean GRV (IES-GI) is higher (a higher degree of discomfort for the same luminance) for the H-L sequence and lower (a lesser degree of discomfort for the same luminance) for the L-H sequence, with the mean GRV for no-PTD trials lying in between the two (other than for Just Intolerable).

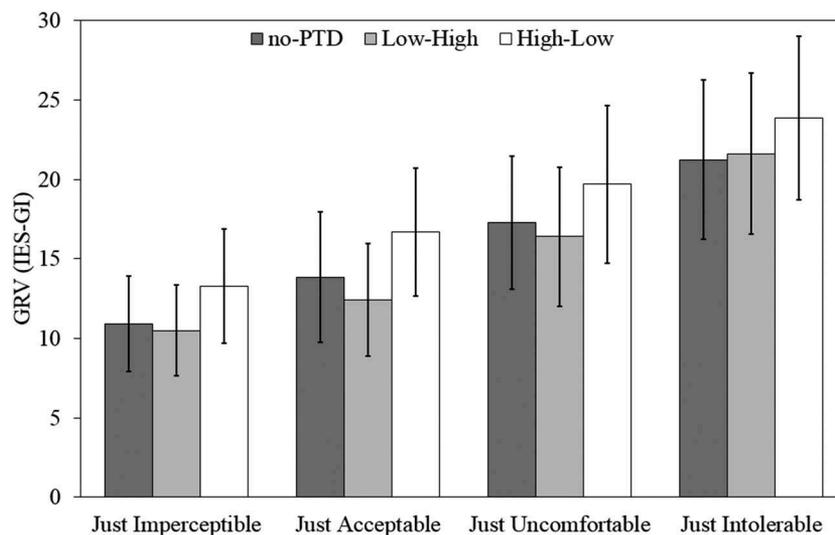
To test the effect of the sequence on the mean GRV (IES-GI) values, comparisons were made between the no-PTD, L-H and H-L. Tables 3 and 4 present the results of the Repeated-Measures

(RM)-ANOVA and paired-sample *t*-tests, providing each glare setting, the comparisons between the no-PTD, L-H and H-L sequences.

Analyses of the data using graphical (Q-Q plots) and statistical (one-sample Shapiro-Wilk tests) methods suggest that the differences in GRV (IES-GI) between all within-subject comparisons (e.g., the luminance ranges) were normally distributed about the mean. Mauchly's test of sphericity revealed that, the variances of the differences across the three luminance ranges were not significant in all four luminance settings: Just Imperceptible,  $\chi^2(2) = 0.98$ ,  $p > .05$ ; Just Acceptable,  $\chi^2(2) = 0.90$ ,  $p > .05$ ; Just Uncomfortable,  $\chi^2(2) = 0.75$ ,  $p \leq 0.01$ ; Just Intolerable,  $\chi^2(2) = 0.82$ ,  $p \leq 0.05$ . Therefore, to meet the assumption of sphericity, Greenhouse-Geisser corrections were applied.

Table 3 presents the results of the RM-ANOVA, providing the glare settings, the test statistic (*F*), the statistical significance (*p*-value), and the effect size ( $\eta_p^2$ ). The interpretation of the outcome was again derived from Ferguson's tables (Ferguson 2009).

The results of the RM-ANOVA suggest that, the differences in means across the independent variable are highly significant for all four glare settings, i.e., the PTD design (no-PTD, L-H or H-L) resulted in significantly different settings. The differences have substantive effect sizes, ranging from moderate ( $0.25 \leq \eta_p^2 < 0.64$ ) for Just Imperceptible and Just Acceptable and small ( $0.04 \leq \eta_p^2 < 0.25$ )

**Fig. 4.** Results of experiment 1: Mean GRV (IES-GI) for the three PTD conditions across the four glare settings. Error bars show standard deviation.

**Table 3.** Analysis of experiment 1 results: RM-ANOVA tests (Greenhouse-Geisser corrections).

Glare Setting	$F(2, 66)$	Statistical Significance	Effect Size ( $\eta_p^2$ )
Just Imperceptible	16.29	<b>0.00***</b>	0.33
Just Acceptable	16.73	<b>0.00***</b>	0.32
Just Uncomfortable	8.97	<b>0.00***</b>	0.21
Just Intolerable	5.57	<b>0.00***</b>	0.14

\* $p \leq 0.05$ ; \*\* $p \leq 0.01$ ; \*\*\* $p \leq 0.001$ ; n.s. = not significant ( $p > 0.05$ )  
 $\eta_p^2 < 0.04$  = negligible;  $0.04 \leq \eta_p^2 < 0.25$  = small;  $0.25 \leq \eta_p^2 < 0.64$  = moderate;  $\eta_p^2 \geq 0.64$  = large

for Just Uncomfortable and Just Intolerable. The magnitude of the effect sizes increases at higher settings, which suggests that the difference in the evaluations across the independent variable increase at higher luminance settings.

Table 4 presents the results of the paired-samples  $t$ -tests, providing the four luminance settings, the comparison under examination (condition) denoted by group 1 and 2, the mean and standard deviations (SD) for each group, the mean difference ( $\Delta M$ ) and the interpretation of the statistical significance (NHST), and the effect size ( $r$ ). The differences are highly significant ( $p \leq 0.01$ ) in eight of the 12 cases and not significant ( $p > .05$ ) in the remaining four. The differences are of substantive effect sizes, ranging from moderate ( $0.50 \leq r < 0.80$ ) in seven cases, small ( $0.20 \leq r < 0.50$ ) in four cases, and negligible ( $r < 0.20$ ) in one case.

Table 4 shows that discomfort evaluations made when using the H-L PTD are significantly different from either the L-H and no-PTD settings, but does not suggest a difference between the L-H and no-PTD settings. The negative differences found in Table 4 suggests an increased level of discomfort for the same glare setting. Compared with a no-

PTD design, showing a PTD affected the subsequent evaluations of discomfort when the H-L sequence was used but not when the L-H sequence was used. For the with-PTD trials, the order in which the upper and lower settings were presented had an effect on the evaluation of discomfort.

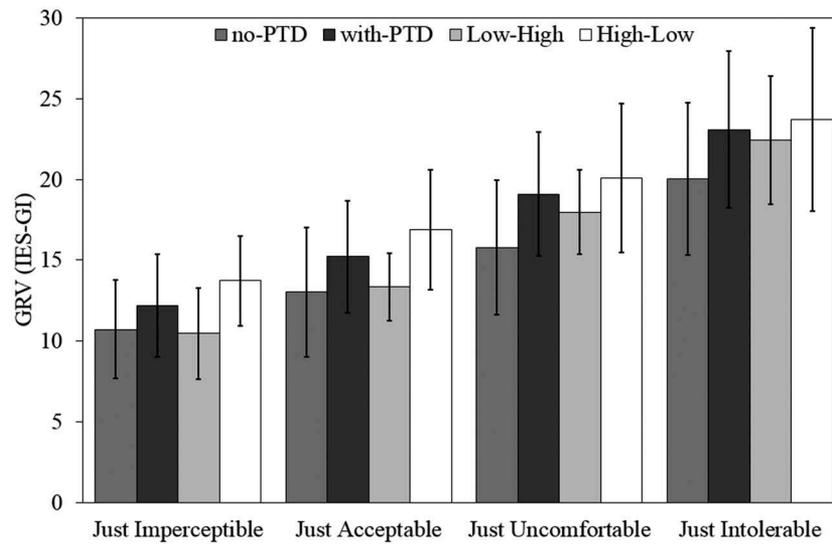
In experiment 1, the block order meant that for some test participants, the no-PTD block was carried out after the with-PTD blocks, and therefore they had seen some demonstration of the likely range of conditions. In other words, while the range was not demonstrated at the start of the no-PTD block, test participants had already seen the range demonstrated (and experienced all four glare settings) in both with-PTD trials (L-H and H-L) before the no-PTD block. To overcome this, the data were reanalysed by considering only the first block of trials conducted by each test participant. This better represents the settings made in an experiment which did not show a PTD and provides a better comparison of the no-PTD and with-PTD conditions, although this is achieved at the expense of a smaller sample size. Of the 34 test participants, 17 made their first evaluations with no-PTD and 17 made evaluations to glare settings with-PTD, of which there were 8 to the L-H sequence and 9 to the H-L sequence.

Fig. 5 presents, on the y-axis, the mean GRV (IES-GI) values. On the x-axis, the plots are organised by the no-PTD, with-PTD: L-H and with-PTD: H-L conditions, with-PTD combined both L-H and H-L sequences. These data are those evaluations made in the first block only. The error bars show the standard deviations.

**Table 4.** Analysis of experiment 1 results: Paired-samples  $t$ -tests and effect sizes.

Glare Setting	Condition <sub>(1 vs. 2)</sub>	Mean(SD) <sub>1</sub>	Mean(SD) <sub>2</sub>	$\Delta M^{\text{NHST}}$	$r$
Just Imperceptible	no-PTD vs. L-H	10.81 (3.00)	10.50 (2.87)	0.30 n.s.	0.10
	no-PTD vs. H-L	10.81 (3.00)	13.28 (3.59)	<b>-2.48***</b>	0.64
	L-H vs. H-L	10.50 (2.87)	13.28 (3.59)	<b>-2.78***</b>	0.65
Just Acceptable	no-PTD vs. L-H	13.81 (4.12)	12.43 (3.52)	1.37 n.s.	0.27
	no-PTD vs. H-L	13.81 (4.12)	16.71 (4.03)	<b>-2.91***</b>	0.60
	L-H vs. H-L	12.43 (3.52)	16.71 (4.03)	<b>-4.27***</b>	0.72
Just Uncomfortable	no-PTD vs. L-H	17.42 (4.19)	16.40 (4.37)	1.02 n.s.	0.20
	no-PTD vs. H-L	17.42 (4.19)	19.69 (4.96)	<b>-2.27***</b>	0.60
	L-H vs. H-L	16.40 (4.37)	19.69 (4.96)	<b>-3.29***</b>	0.67
Just Intolerable	no-PTD vs. L-H	21.49 (5.02)	21.62 (5.09)	-0.12 n.s.	0.23
	no-PTD vs. H-L	21.49 (5.02)	23.85 (5.16)	<b>-2.36***</b>	0.55
	L-H vs. H-L	21.62 (5.09)	23.85 (5.16)	<b>-2.24***</b>	0.43

With Bonferroni corrections: \* $p \leq 0.05$ ; \*\* $p \leq 0.01$ ; \*\*\* $p \leq 0.001$ ; n.s. = not significant ( $p > 0.05$ )  
 $r < 0.20$  = negligible;  $0.20 \leq r < 0.50$  = small;  $0.50 \leq r < 0.80$  = moderate;  $r \geq 0.80$  = large



**Fig. 5.** Results of experiment 1 for the first block of trials only: Mean GRV (IES-GI) for the no-PTD, L-H and H-L and with-PTD conditions across the four glare settings. Error bars show standard deviations. Note that with-PTD is the average of the L-H and H-L trials.

Compared with the mean GRV (IES-GI) established in no-PTD trials, the mean GRV (IES-GI) found with the H-L trials is higher, as was found from the results of all blocks (Fig. 4). However, the mean GRV (IES-GI) established in the L-H with-PTD trials now appears to be equal to or greater than that of the no-PTD trials: in Fig. 4 the mean GRV (IES-GI) was lower than that for the no-PTD trials.

Independent-samples *t*-tests were used to analyse the mean differences in GRV (IES-GI) between: the no-PTD and average-PTD conditions and also the with-PTD: L-H and H-L sequences. Table 5 shows the results of the analyses, presenting for each glare setting, the mean and standard deviations (SD), the difference ( $\Delta M$ ) and statistical significance (*p*-value), and the effect size (*r*).

Since the GRV (IES-GI) values are larger under the average-PTD, negative differences appear across the four glare settings. This shows a higher degree of discomfort for the same glare setting. These differences are statistically significant ( $p \leq 0.01$ ) for the Just Uncomfortable glare settings and weakly significant ( $p \leq 0.05$ ) for glare settings: Just Imperceptible, Just Acceptable, and Just Intolerable. The effect sizes are all of a substantive magnitude, corresponding to small ( $0.20 \leq r < 0.50$ ) in all four cases. The differences in GRV (IES-GI) values between the with-PTD conditions (L-H and H-L) are all negative, similar to the findings derived in Table 4 when considering all block trials. The differences are statistically significant ( $p \leq 0.01$ ) for the setting: Just

**Table 5.** Analysis of experiment 1 results from the first block of trials: Independent-samples *t*-tests used to analyse the with-PTD conditions. Independent-samples *t*-tests used to analyse the with-PTD: L-H and H-L sequences.

Glare Setting	Mean GRV (standard deviation)		$\Delta M^{\text{NHST}}$	<i>r</i>
	<i>no-PTD</i>	<i>with-PTD</i>		
Just Imperceptible	10.73 (3.02)	12.19 (3.17)	<b>-1.46*</b>	0.24
Just Acceptable	13.03 (3.98)	15.21 (3.46)	<b>-2.18*</b>	0.29
Just Uncomfortable	15.79 (4.17)	19.09 (3.84)	<b>-3.30**</b>	0.39
Just Intolerable	20.02 (4.70)	23.09 (4.84)	<b>-3.07*</b>	0.32
Glare Setting	L-H Sequence	H-L Sequence	$\Delta M^{\text{NHST}}$	<i>r</i>
Just Imperceptible	10.47 (2.82)	13.73 (2.75)	<b>-3.25**</b>	0.53
Just Acceptable	13.35 (2.06)	16.86 (3.70)	<b>-3.51*</b>	0.52
Just Uncomfortable	17.97 (2.60)	20.09 (4.61)	<b>-2.13*</b>	0.28
Just Intolerable	22.42 (3.98)	23.68 (5.67)	-1.27 n.s.	0.14

\* $p \leq 0.05$ ; \*\* $p \leq 0.01$ ; \*\*\* $p \leq 0.001$ ; n.s. = not significant ( $p > 0.05$ )

$r < 0.20$  = negligible;  $0.20 \leq r < 0.50$  = small;  $0.50 \leq r < 0.80$  = moderate;  $r \geq 0.80$  = large

Imperceptible, Just Acceptable and Just Uncomfortable and weakly not significant ( $p > .05$ ) for Just Intolerable. The effect sizes are of a substantive magnitude, corresponding to moderate ( $0.50 \leq r < 0.80$ ) in two cases, small ( $0.20 \leq r < 0.50$ ) for Just Uncomfortable, and negligible ( $r < 0.20$ ) for Just Intolerable.

These results provided evidence that using a PTD influenced evaluations of discomfort made in the first block of trials. Since the same influence could be detected when analysing the results from all blocks, this shows the effect of PTD was not confounded when the block trials were balanced across the participants.

### 3.3. Experiment 1: Summary

Although it has been recommended that a PTD is used before seeking evaluations using category rating (Adams et al. 2004; Fotios and Houser 2009; Houser and Tiller 2003; Olkkonen et al. 2014; Tiller and Rea 1992), this has been rarely done in past studies of discomfort due to glare. A PTD means showing low and high stimulus magnitudes and this could follow an ascending (lower limit first) or descending (upper limit first) sequence. One approach is to use both orders and take the average of the setting made with each order. Fig. 4 indicates that showing a PTD leads to significant effect on the discomfort evaluation: The glare evaluation tends to be higher (a greater degree of discomfort for the same glare setting) when a PTD (with-PTD trials) is shown than when it is not shown (no-PTD trials). Analysis of the L-H and H-L PTD sequences separately shows, however, that while evaluations made after the H-L PTD are significantly different to the no-PTD trial (and significantly different from the L-H trials), those evaluations made following the L-H PTD were not suggested to be different from the no-PTD trials.

One question not addressed by experiment 1 is whether the PTD itself led to the higher discomfort ratings or whether it was the specific upper and lower limits of the PTD that were used. To explore this, a second experiment was conducted in which three different PTD ranges were employed. If it is the use of the PTD that matters, then the discomfort evaluations would not vary

with PTD range. If, however, it is the range that matters, then different PTD ranges would lead to different evaluations of discomfort. Exploring evaluations made with different PTD ranges may also explain why the H-L but not the L-H trials in experiment 1 led to significant differences compared with the no-PTD trials.

## 4. Experiment 2

### 4.1. Procedure

Twenty-one postgraduate students were recruited, and these were independent from those used in experiment 1. The sample comprised 11 female and 10 male, with a mean age of 28 years ( $SD = 4$ ). Twelve participants wore their normal glasses or corrective lens during the tests, and all self-certified as having no other health or eye problems.

The procedure was identical to that used for experiment 1 other than variation in the luminances used for the with-PTD trials and that the no-PTD block was omitted. Each of three blocks presented a different range of luminances in the PTD (Table 6) these being overlapping sub-sections of the PTD range used in experiment 1. For any one of these three PTD ranges this means that the upper and lower luminances do not both extend beyond the range of glare settings as would otherwise be recommended by Poulton (1989). For example, consider the low range: while the lower limit ( $229 \text{ cd/m}^2$ ) is below the just imperceptible glare setting ( $762 \text{ cd/m}^2$ , see Table 1) the upper limit ( $1799 \text{ cd/m}^2$ ) is identical to the just acceptable glare setting. This means that test participants were asked to evaluate glare settings of higher luminance than the upper limit of the low range PTD. This was done to test Poulton's advice. An alternative approach would be to select, for each PTD range, luminances that were beyond the glare settings.

In experiment 1, it was shown that the PTD sequence (ascending or descending) influenced the

**Table 6.** Definition of the glare source luminances used to demonstrate the PTD in the three blocks.

PTD Luminance Range	Lower Limit ( $\text{cd/m}^2$ )	Upper Limit ( $\text{cd/m}^2$ )
Low	229	1799
Middle	762	4122
High	2354	12219

discomfort evaluations: specifically, the H-L sequence led to significantly different evaluations than the no-PTD trials, but the L-H sequence did not. The objective of experiment 2 was to compare different PTD ranges, not to compare PTD sequences, and therefore only one sequence was used. The PTD started from the lower limit, was then adjusted to the upper limit, and then adjusted back to lower limit (i.e. L-H-L). This sequence was used in each block. The three blocks were used in a random order, and within each block, the four glare settings were evaluated in a randomised order.

#### 4.2. Results

Figure 6 presents the results of experiment 2. The y-axis is the mean GRV (IES-GI) value calculated from evaluations provided on the continuous glare scale. Along the x-axis, the figure presents the four glare settings. The mean plots are organised according to the PTD luminance ranges of the three blocks. Error bars show the standard deviations about the means.

Figure 6 shows that higher glare settings are associated with a higher mean GRV (IES-GI), indicating as expected that higher luminance leads to greater discomfort. Within each of the four glare settings, there is a consistent effect of PTD range in that the higher PTD range led to lower GRV (IES-GI), signalling less discomfort for the same luminance.

Analyses of the data using graphical (Q-Q plots) and statistical (one-sample Shapiro-Wilk tests) methods suggest that the differences in GRV (IES-GI) between all within-subject comparisons (e.g., the luminance ranges) were normally distributed about the mean. Mauchly's test of sphericity revealed that, the variances of the differences across the three luminance ranges were not significant in all four luminance settings: Just Imperceptible,  $\chi^2(2) = 2.04, p > .05$ ; Just Acceptable,  $\chi^2(2) = 1.45, p > .05$ ; Just Uncomfortable,  $\chi^2(2) = 0.79, p > .05$ ; and Just Intolerable,  $\chi^2(2) = 1.90, p > .05$ . Since the assumption of sphericity had been met, the RM-ANOVA was suitable for this analysis (Field 2013).

Table 7 presents the results of the RM-ANOVA (for the data in experiment 2), providing the glare settings, the test statistic ( $F$ ), the statistical significance ( $p$ -value), and the effect size ( $\eta_p^2$ ). The interpretation of the outcome was again derived from Ferguson's tables (Ferguson 2009).

The results of the RM-ANOVA suggest that, the differences in means across the independent variable are significant ( $p \leq 0.01$ ) for the Just Imperceptible and Just Acceptable settings, weakly significant ( $p \leq 0.05$ ) for the Just Uncomfortable setting, and not significant ( $p > .05$ ) for the Just Intolerable setting. The differences all have small, but substantive effect sizes ( $0.04 \leq \eta_p^2 < 0.25$ ). When participants gave evaluations of to higher glare settings, the magnitude of effect size decreases. Therefore, the differences in mean GRV

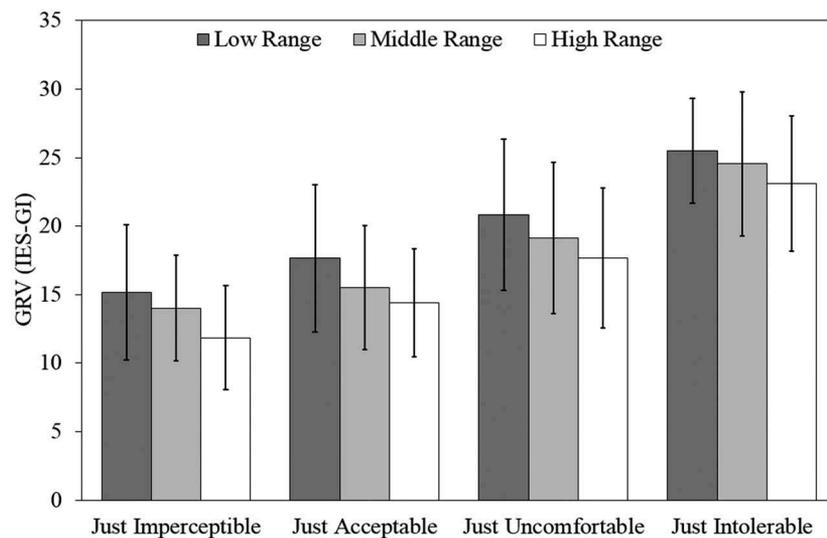


Fig. 6. Results of experiment 2: Mean GRV (IES-GI) for the three luminance ranges for the four glare settings. Error bars present the standard deviations.

**Table 7.** Analysis of experiment 2 results: RM-ANOVA tests (sphericity assumed).

Glare Setting	$F(2, 42)$	Statistical Significance	Effect Size ( $\eta_p^2$ )
Just Imperceptible	5.48	<b>0.01**</b>	0.22
Just Acceptable	5.02	<b>0.01**</b>	0.20
Just Uncomfortable	3.17	<b>0.05*</b>	0.14
Just Intolerable	2.68	0.08 n.s.	0.12

\* $p \leq 0.05$ ; \*\* $p \leq 0.01$ ; \*\*\* $p \leq 0.001$ ; n.s. = not significant ( $p > 0.05$ )  
 $\eta_p^2 < 0.04$  = negligible;  $0.04 \leq \eta_p^2 < 0.25$  = small;  $0.25 \leq \eta_p^2 < 0.64$  = moderate;  $\eta_p^2 \geq 0.64$  = large

(IES-GI) across the luminance ranges become smaller. Similar findings were also detected in experiment 1, which showed the range equalizing bias became smaller at higher luminances.

Table 8 presents the results of the paired-samples  $t$ -tests, providing the four luminance settings, the comparison under examination (range) denoted by group 1 and 2, the mean and standard deviations (SD) for each group, the mean difference ( $\Delta M$ ) and the interpretation of the statistical significance (NHST), and the effect size ( $r$ ).

The differences are positive across the paired comparisons, signalling lower values of GRV (IES-GI) when the PTD used a higher luminance range. The differences are significant ( $p \leq 0.01$ ) in four cases, weakly significant ( $p \leq 0.05$ ) in five cases, and not significant ( $p > .05$ ) in three out of 12 cases. The differences are of substantive effect sizes, ranging from moderate ( $0.50 \leq r < 0.80$ ) in two cases and small ( $0.20 \leq r < 0.50$ ) in 10 out of 12 cases.

The inferential analysis of the data hence confirms that, when observers experienced the high luminance range in the PTD, they gave lower glare evaluations to the same settings. This suggests that: (1) the

luminance range used in the PTD influences the evaluations given to the four glare settings, and (2) the differences in the evaluations made on the continuous scale are larger when considering comparisons made between the low and high with-PTD luminance ranges. This is apparent for all four glare settings.

Since all evaluations made to the glare settings in each block of trials commenced after a PTD with different luminance ranges was shown, participants were always presented a condition with-PTD in the first block. However, to determine whether the luminance ranges used in other PTD blocks effected the evaluations in Fig. 6, a further analysis was conducted using only results from the first block of trials from each participant.

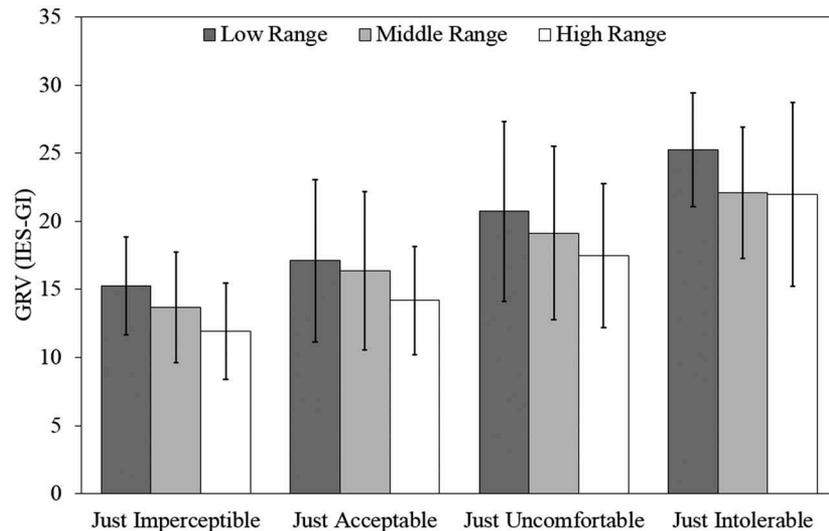
When considering only the evaluations made in the first block of trials, participants were equally assigned ( $n = 7$ ) to one of the three PTD luminance range conditions. The results of the first block trial are shown in Fig. 7. On the y-axis Fig. 7 shows the mean GRV (IES-GI) values. On the x-axis, the plots are organised by the with-PTD: low, middle and high luminance ranges. The error bars show the standard deviations.

Similar to the findings seen in Fig. 6, the analysis of the first block shows the evaluations are biased by the luminance range used in the PTD. Lower GRV (IES-GI) (less discomfort on the continuous glare scale) are seen when higher luminances were used in the PTD range. This can be seen for all four luminance settings. Statistical analyses using a one-way ANOVA did not suggest differences between any of the

**Table 8.** Analysis of experiment 2 results: Paired-samples  $t$ -tests and effect sizes.

Glare Setting	Range <sub>(1 vs. 2)</sub>	Mean(SD) <sub>1</sub>	Mean(SD) <sub>2</sub>	$\Delta M^{NHST}$	$r$
Just Imperceptible	Low vs. Medium	15.16 (4.91)	14.02 (3.86)	1.14 n.s.	0.21
	Low vs. High	15.16 (4.91)	11.86 (3.82)	<b>3.30**</b>	0.58
Just Acceptable	Medium vs. High	14.02 (3.86)	11.86 (3.82)	<b>2.16**</b>	0.48
	Low vs. Medium	17.66 (5.38)	15.52 (4.52)	<b>2.15*</b>	0.38
Just Uncomfortable	Low vs. High	17.66 (5.38)	14.41 (3.96)	<b>3.26**</b>	0.56
	Medium vs. High	15.52 (4.52)	14.41 (3.96)	<b>1.11*</b>	0.26
Just Intolerable	Low vs. Medium	20.83 (5.50)	19.14 (5.54)	<b>1.69*</b>	0.26
	Low vs. High	20.83 (5.50)	17.69 (5.12)	<b>3.14**</b>	0.49
Just Intolerable	Medium vs. High	19.14 (5.54)	17.69 (5.12)	<b>1.45*</b>	0.26
	Low vs. Medium	25.49 (3.81)	24.54 (5.24)	0.95 n.s.	0.20
Just Intolerable	Low vs. High	25.49 (3.81)	23.09 (4.94)	<b>2.40*</b>	0.46
	Medium vs. High	24.54 (5.24)	23.09 (4.94)	1.45 n.s.	0.28

With Bonferroni corrections: \* $p \leq 0.05$ ; \*\* $p \leq 0.01$ ; \*\*\* $p \leq 0.001$ ; n.s. = not significant ( $p > 0.05$ )  
 $r < 0.20$  = negligible;  $0.20 \leq r < 0.50$  = small;  $0.50 \leq r < 0.80$  = moderate;  $r \geq 0.80$  = large



**Fig. 7.** Results of experiment 2 for the first block of trials only: Mean GRV (IES-GI) for the three luminance ranges for the four glare settings. Error bars present the standard deviations.

ranges to be significant, which may be due to the small sample size. However, results of the first trials (Fig. 7) and all trials (Fig. 6) are similar, which suggests the same evaluations occurred in both cases.

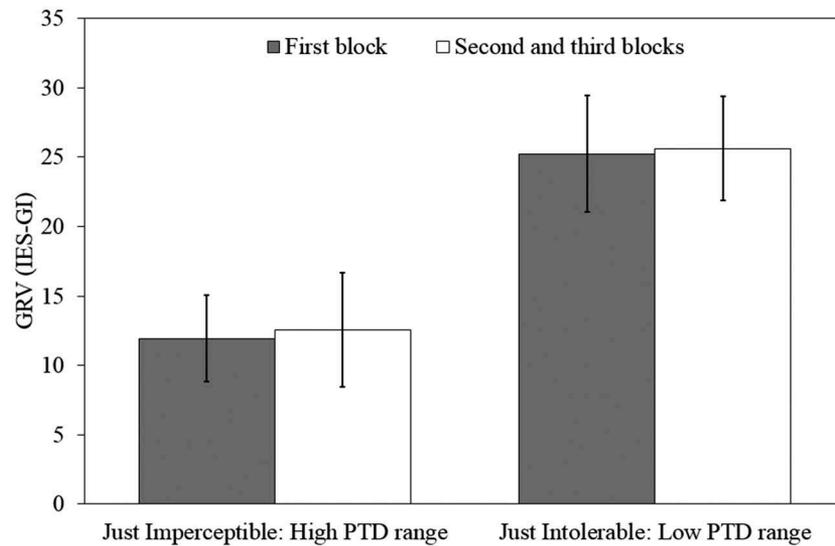
#### 4.3. Experiment 2: Summary

When using PTDs with different ranges, the results suggest this will influence the degree of discomfort evaluated by test participants. Specifically, when higher luminances are used in the PTD, lower evaluations are given (a lesser degree of discomfort is reported for the same glare source luminance). This shows that the PTD range matters and different PTD limits used will lead to different outcomes. The influence of PTD range may be explained as an adaptation effect: with lower PTD luminances, subsequent glare settings were relatively brighter, leading to greater sensation of discomfort.

In experiment 2, participants evaluated glare settings with luminances beyond (i.e., higher and/or lower, depending upon the PTD range) those used in the PTD range. For example, in the low PTD range, the Just Intolerable glare setting ( $9819 \text{ cd/m}^2$ ) is greater than the upper end of the PTD ( $1799 \text{ cd/m}^2$ ), and in the high PTD range the Just Imperceptible glare setting ( $762 \text{ cd/m}^2$ ) is lower than the lower end of the PTD ( $2354 \text{ cd/}$

$\text{m}^2$ ). The three PTD ranges were separate blocks, with the blocks used in a randomised order. For those test participants where the low PTD range was the first block, the Just Intolerable glare setting, being greater than the upper end of the PTD, may have appeared brighter than in those trials when the low PTD range was the second or third block and thus influenced by exposure to higher glare source luminances. For these two situations, Fig. 8 compares the mean GRV (IES-GI) for the first block of trials with the second and third blocks of trials. Since the order of the three blocks (PTD ranges) was balanced across the 21 test participants, the sample size is reduced to  $n = 7$  for the first block and  $n = 14$  for the second and third blocks combined.

Consider the low PTD range. When this was experienced as the first block the evaluations were not influenced by the higher PTDs of the other two blocks. If the upper PTD limit being within rather than beyond the range of glare settings was of impact this would result in different glare settings for the just intolerable setting when the low range was the first block than when it was a second or third block. A similar prediction can be made for just imperceptible settings with the high PTD range. Independent samples  $t$ -tests were used to analyse the data. For both comparisons seen in Fig. 8 the differences were not statistically significant or practically relevant: Just Imperceptible: High range,  $p = .74$ ,  $r = 0.10$

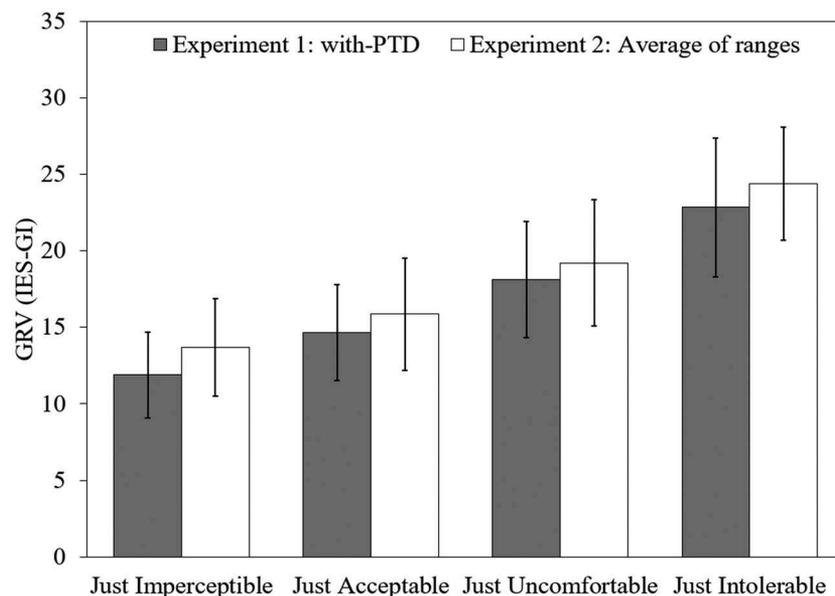


**Fig. 8.** Results of experiment 2 for the low and high ranges only with the first block of trials with the second and third block trials: Mean GRV (IES-GI) for Just Imperceptible and Intolerable glare settings. Error bars present the standard deviations.

(negligible) and Just Intolerable: Low range,  $p = .84$ ,  $r = 0.11$  (negligible). These results suggest that the evaluation of glare settings with luminances beyond those of limits used in the PTD range did not influence the evaluation.

Trials with a PTD were the L-H and H-L trials in experiment 1 and all three ranges in experiment 2. Figure 9 shows the average of the trials within each experiment. Table 9 presents the results of the independent-samples  $t$ -tests, providing the four glare settings, the mean and standard

deviations (SD), the differences between the means ( $\Delta M$ ) and the interpretation of the statistical significance (NHST), and the effect size ( $r$ ). For three glare settings the differences are not suggested to be statistically significant and the effect size is negligible: for just imperceptible glare setting the difference is statistically significant ( $p < .05$ ) and the effect size is small. Overall this suggests that when settings are averaged across multiple trials using different PTD designs, the differences were negligible.



**Fig. 9.** Comparison of with-PTD (L-H and H-L) in experiment 1 and average of the three ranges (low, middle and high) in experiment 2. Error bars present the standard deviations.

**Table 9.** Analysis of with-PTD in experiment 1 and average of the three ranges in experiment 2: Independent-samples *t*-tests and effect sizes.

Glare Setting	Experiment 1	Experiment 2	$\Delta M^{\text{NHST}}$	<i>r</i>
	Mean(SD) <sub>1</sub>	Mean(SD) <sub>2</sub>		
Just Imperceptible	11.89 (2.80)	13.68 (3.19)	-1.79*	0.28
Just Acceptable	14.57 (3.14)	15.86 (3.66)	-1.19 n.s.	0.19
Just Uncomfortable	18.04 (3.78)	19.22 (4.14)	-1.09 n.s.	0.15
Just Intolerable	22.73 (4.53)	24.37 (3.70)	-1.53 n.s.	0.19

\* $p \leq 0.05$ ; \*\* $p \leq 0.01$ ; \*\*\* $p \leq 0.001$ ; n.s. = not significant ( $p > 0.05$ )  
 $r < 0.20$  = negligible;  $0.20 \leq r < 0.50$  = small;  $0.50 \leq r < 0.80$  = moderate;  
 $r \geq 0.80$  = large

## 5. Conclusion

This article has discussed the category rating procedure when used to evaluate discomfort due to glare. It has been suggested that the range of stimuli should be demonstrated to test participants before asking them to provide evaluations using category rating (Adams et al. 2004; Fotios and Houser 2009; Houser and Tiller 2003; Olkkonen et al. 2014), here labelled a pre-trial demonstration (PTD). Two experiments were carried out to test the influence of the PTD. The results of the first experiment indicate that a PTD influenced subsequent evaluations of discomfort, specifically that with-PTD trials led to ratings of higher discomfort for the same glare source luminance than did no-PTD trials (Fig. 5, Table 5). Results of the second experiment indicate that the range of luminances used in the PTD affected the evaluations given, with a higher degree of discomfort reported when using the lower range of PTD luminances. The results of experiment 2 also suggest that it did not matter whether the PTD limits were within, or beyond, the range of glare source luminances that were evaluated.

While these results suggest significant effects, it is not clear how this should be implemented in practise. Specifically, which condition (i.e., with-PTD or no-PTD) provides a closer approximation to the degree of discomfort experienced in a natural setting. Assuming that the with-PTD procedure is deemed more relevant, it is not clear which luminance range is more appropriate and further investigation is required. The findings suggest that a PTD order following a H-L PTD sequence had a significant influence on the evaluations, while a reversed L-H PTD sequence did not. When lower luminances limits are used, this increases the evaluations of discomfort for the same glare settings. In future experiments, the use of luminance limits need to be

carefully considered, otherwise the final evaluation would be underestimated if the luminances used are too high or overestimated if they are too low.

One other aspect to consider is the evaluation of the first-block of trials with independent-subjects analyses together with all block trials using the full dataset and repeated-measures analyses. While this was used here to determine the influence of unwanted procedure effects due to the randomised demonstration order of the PTD conditions, a caveat to this approach is that independent-subjects analyses reduce the sample size by the number of conditions that are considered. Nonetheless, a comparison of the first-block and all block trials in studies could provide a useful indication to whether confounding influences of repeated-measures designs are present – regardless of when randomisation or Latin-squared approaches have been implemented in the experimental design.

Nevertheless, category rating is a widely used procedure in research of the discomfort due to glare. The current results, along with those from previous work (Gellatly and Weintraub 1990, Kent et al. 2017, 2018a, 2018b, 2019; Lulla and Bennett 1981) show that test results are influenced by a range of experimental parameters. We should therefore be cautious about the results gained from such studies, and thus cautious about recommendations based on their conclusions.

## Acknowledgment

Acknowledgement is given to Ayesha Batool (University of Nottingham) for the contributions given in the study.

## Disclosure statement

The authors have no financial interests to declare.

## Funding

This work was supported by the Engineering and Physical Sciences Research Council [Grant Number EP/N50970X/1]; and the Republic of Singapore's National Research Foundation through a grant to the Berkeley Education Alliance for Research in Singapore (BEARS) for the Singapore-Berkeley Building Efficiency and Sustainability in the Tropics (SinBerBEST) Program. BEARS has been established by the University of California, Berkeley as a centre for intellectual excellence in research and education in Singapore.

## ORCID

Michael G Kent  <http://orcid.org/0000-0002-4430-3893>

Steve Fotios  <http://orcid.org/0000-0002-2410-7641>

## References

- [CIBSE]. Chartered Institution of Building Services Engineers. 1994. Code for interior lighting. London (UK): CIBSE.
- [IESNA]. Illuminating Engineering Society of North America. 2011. The Lighting Handbook, Tenth ed. Reference and Application. New York (NY): IESNA.
- Adams WJ, Graf EW, Ernst MO. 2004. Experience can change the “light-from-above” prior. *Nat Neurosci*. 7(10):1057–1958.
- Altomonte S, Kent MG, Tregenza PR, Wilson R. 2016. Visual task difficulty and temporal influences in glare response. *Build Environ*. 95:209–226.
- Boyce PR, Cuttle C. 1990. Effect of correlated colour temperature on the perception of interiors and colour discrimination performance. *Light Res Technol*. 22(1):19–36.
- Cabin RJ, Mitchell RJ. 2000. To Bonferroni or not to Bonferroni: when and how are the questions. *Bull Ecol Soc Am*. 81(3):246–248.
- Cohen J. 1994. The Earth is round ( $p < .05$ ). *Am Psychol*. 49(12):997–1003.
- Ellis PD. 2010. The essential guide to effect sizes: statistical power, meta-analysis, and the interpretation of research results. Cambridge (UK): Cambridge University Press.
- Ferguson CJ. 2009. An Effect Size Primer: A Guide for Clinicians and Researchers. *Prof Psychol Res Pr*. 40(5):532–538.
- Field A. 2013. Discovering statistics using IBM SPSS statistics. 4th ed. London (UK): Sage.
- Flynn JE, Spencer TJ. 1977. The effects of light source color on user impression and satisfaction. *Illumin Eng*. 6(3):167–179.
- Fotios SA, Houser KW. 2009. Research methods to avoid bias in categorical ratings of brightness. *LEUKOS*. 5(3):167–181.
- Gellatly AW, Weintraub DJ. 1990. User reconfigurations of the de Boer rating scale for discomfort glare. Ann Arbor: The University of Michigan: Transportation Research Institute. UMTRI-90-20.
- Gescheider GA. 1985. Psychophysics: method, theory, and application. New York, NY: L. Erlbaum Associates.
- Greenhouse SW, Geisser S. 1959. On methods in the analysis of profile data. *Psychometrika*. 24(2):95–112.
- Hauschke D, Steinijans VW. 1996. Directional decision for a two-tailed alternative. *J Biopharm Stat*. 6(2):211–218.
- Hopkinson RG. 1960. A note on the use of indices of glare discomfort for a code of lighting. *IESTrans*. 25(3):135–138.
- Hopkinson RG. 1972. Glare from daylighting in buildings. *Appl Ergon*. 3(4):206–215.
- Houser KW, Tiller DK. 2003. Measuring the subjective response to interior lighting: paired comparisons and semantic differential scaling. *Light Res Technol*. 35(3):183–195.
- Kent MG, Fotios S, Altomonte S. 2017. Discomfort glare evaluation: the influence of anchor bias in luminance adjustments. *Light Res Technol*. 51(1):131–146.
- Kent MG, Fotios S, Altomonte S. 2018a. Order effects when using Hopkinson’s multiple criterion scale of discomfort due to glare. *Build Environ*. 136:54–61.
- Kent MG, Fotios S, Altomonte S. 2018b. An experimental study on the effects of visual tasks on discomfort due to peripheral glare. *LEUKOS*. 15(1):17–28.
- Kent MG, Fotios S, Cheung T. 2019. Stimulus range bias leads to different settings when using luminance adjustment to evaluate discomfort due to glare. *Build Environ*. 153:281–287.
- Logadóttir Á, Christoffersen J, Fotios SA. 2011. Investigating the use of an adjustment task to set preferred illuminance in a workplace environment. *Light Res Technol*. 43(4):403–422.
- Luckiesh M, Guth SK. 1949. Brightnesses in Visual Field at Borderline Between Comfort and Discomfort (BCD). *Illumin Eng*. 44:650–670.
- Lulla AB, Bennett CA. 1981. Discomfort glare: range effects. *Illumin Eng*. 10(2):74–80.
- Mauchly JW. 1940. Significant test for sphericity of a normal n-variate distribution. *Ann Math Stat*. 11(2):204–209.
- Olkkonen M, McCarthy PF, Allred SR. 2014. The central tendency bias in color perception: effects of internal and external noise. *J Vis*. 14(11):1–15.
- Petherbridge P, Hopkinson RG. 1950. Discomfort Glare and the Lighting of Buildings. *IESTrans*. XV(2):39–79.
- Poulton EC. 1989. Bias in quantifying judgements. Milton Park (UK): Taylor & Francis.
- Robinson W, Bellchambers HE, Grundy JT, Longmore J, Burt W, Hewitt H, Petherbridge P, Frith D, Hopkinson RG, Rowlands E. 1962. The development of the IES glare index system: contribution by the luminance study panel of the IES technical committee. *IESTrans*. 27(1):9–26.
- Ruxton GD. 2006. The unequal variance t-test is an under-used alternative to student’s t-test and the Mann-Whitney U test. *Behav Ecol*. 17(4):688–690.
- Shaffer JP. 1995. Multiple hypothesis testing. *Annu Rev Psychol*. 46(1):561–584.
- Stone PT, Harker SDP. 1973. Individual and group differences in discomfort glare responses. *Light Res Technol*. 5(1):41–49.
- Tiller DK, Rea MS. 1992. Semantic differential scaling: prospects in lighting research. *Light Res Technol*. 24(1):43–51.
- Tuaycharoen N, Tregenza PR. 2005. Discomfort glare from interesting images. *Light Res Technol*. 37(4):329–341.
- Tuaycharoen N, Tregenza PR. 2007. View and discomfort glare from windows. *Light Res Technol*. 39(2):185–200.
- Velds M. 2002. User acceptance studies to evaluate discomfort glare in daylight rooms. *Sol Energy*. 73(2):95–103.
- Viénot F, Durand ML, Mahler E. 2009. Kruithof’s rule revisited using LED illumination. *J Mod Opt*. 56(13):1433–1446.
- Vos JJ. 2003. Reflections on glare. *Light Res Technol*. 35(2):163–175.
- Vrabel PL, Bernecker CA, Mistrick RG. 1998. Visual performance and visual clarity under electric light sources: part II – visual clarity. *Illumin Eng*. 27(1):29–41.

## Appendix: Definitions of discomfort as given to test participants

In this experiment, you will be asked to express your own perceived level of discomfort glare when presented to the glare source, using four threshold criteria of glare sensation votes (GSVs): 'Just Imperceptible', 'Just Acceptable', 'Just Uncomfortable' and 'Just Intolerable'.

These are described below:

- *Just Imperceptible*: when the source of the light becomes quite bright without necessarily giving a sensation of glare. As the light source is being adjusted, for a moment while performing the visual task, the source would be something that attracts your attention.
- *Just Acceptable*: this corresponds to a glare sensation that could be tolerated for approximately one day when working in this room. If you had to work under this lighting condition at your own workstation, you may want to use blinds or other measures to decrease the perceived discomfort.
- *Just Uncomfortable*: this corresponds to a glare sensation that could be tolerated for approximately 15 to 30 minutes, for example if finishing a certain task would take this amount of time. After this, adjustments to the lighting conditions would be made, if the same degree of discomfort would be present over time.
- *Just Intolerable*: this corresponds to the point where you would no longer be able to work under these lighting conditions for any amount of time and would immediately intervene to change them.