Visual Task Difficulty and Temporal Influences in Glare Response

Sergio Altomonte[a*], Michael G. Kent[a], Peter R. Tregenza[b], Robin Wilson[a]

[a] Department of Architecture and Built Environment, The University of Nottingham, UK

[b] Sheffield School of Architecture, The University of Sheffield, UK

*Corresponding author: sergio.altomonte@nottingham.ac.uk

Abstract

The literature suggests that glare sensation may be influenced by visual task difficulty. Previous research by the authors provided reasons to infer that the perceived level of visual discomfort may vary with time of day and be affected by temporal and personal factors. The study presented here explores the postulated relationships between visual task difficulty, temporal variables, and glare response as the day progresses. Under controlled laboratory conditions, twenty subjects were exposed to a constant artificial source luminance at four times of day and gave glare sensation votes while completing twelve visual tasks of various difficulties. Self-assessments of temporal variables (fatigue, food intake, caffeine ingestion, mood, previous daylight exposure and sky condition) were provided by test subjects together with their glare judgements. Statistical analysis of responses confirmed that the time interval between test sessions showed a direct relationship to the increased tolerance to artificial source luminance along the day. The temporal variation of glare response was found to be influenced by the difficulty in extracting information from the visual stimulus. Moreover, statistically significant and substantive evidence was detected of a direct effect of fatigue and caffeine ingestion, and an inverse influence of food intake, on reported glare sensation. Consideration of inferential results from all test sessions led to hypothesise that some temporal variables may interact with each other and significantly affect the variation of glare response at different times of day.

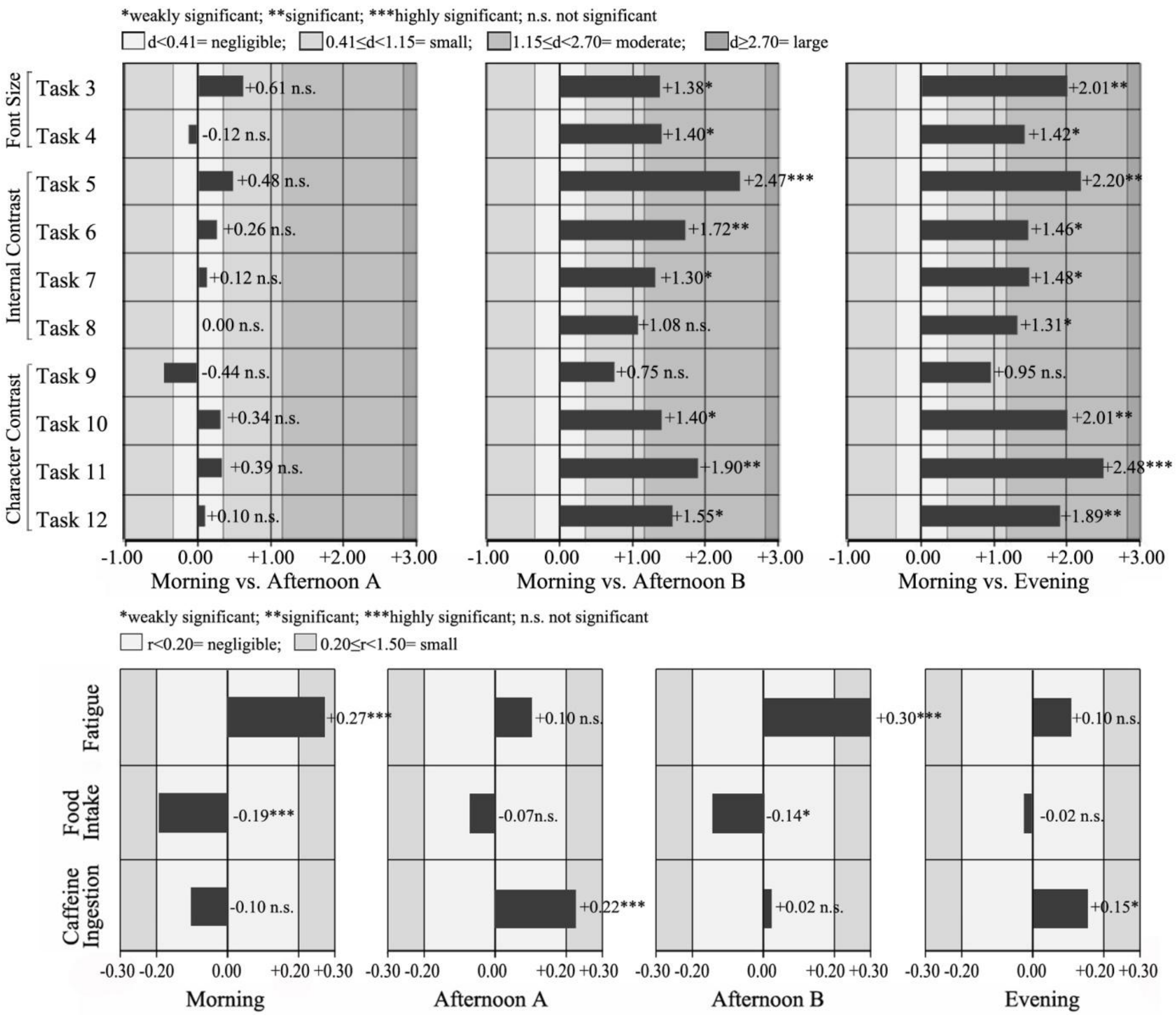
Keywords: Glare; Visual Task Difficulty; Time of Day; Fatigue; Caffeine Ingestion; Food Intake

Reference: Altomonte, S., Kent, M.G., Tregenza, P.R. and Wilson, R., 2016. Visual Task Difficulty and Temporal Influences in Glare Response. ***Building and Environment*** 95, 209-226

Highlights

- Influence of task difficulty and temporal variables on glare response was studied
- Reported tolerance to source luminance was found to increase as the day progresses
- The effect of time of day on glare response varied with task type and difficulty
- A direct relation was found between glare sensation, fatigue and caffeine ingestion
- Temporal variables interact with each other affecting glare response along the day

Graphical Abstract



https://doi.org/10.1016/j.buildenv.2015.09.021
https://escholarship.org/uc/item/5g20q4dg

1. Introduction

Discomfort glare is a phenomenon that has not yet been completely characterised [1, 2]. Other than the factors typically embedded in glare formulas (e.g., source luminance, background luminance, solid angle, position index, etc.), several additional parameters have been associated with the occurrence and perceived magnitude of visual discomfort. Among others, various studies have investigated the potential relationship between task difficulty, visual performance, and subjective glare sensation.

Conventionally, the ability to extract information from a visual stimulus has been considered independently from *discomfort glare* [3]. In fact, in the literature, visual impairment and reduced performance have been predominantly linked to *disability glare*, whereas in the presence of discomfort glare the observer may not experience any immediate direct effect on task visibility [4]. Boyce [1] and Sivak *et al*. [3] stated that disability and discomfort glare may be regarded as part of the same phenomenon, and that the mechanisms behind these two types of glare may not be as different as commonly assumed, although they are generally discerned by the ranges of luminance in the visual field. Hitherto, in most cases, the magnitude of the perceived impairment resulting from discomfort glare is considered to be lower than disability glare [5].

According to Boyce [1], visual discomfort can result either from a combination of photometric conditions present in the environment or from the visual task itself. The variation of glare sensation has been associated to the size and contrast of the task; as these augment, visual performance (i.e., speed and accuracy) increases and discomfort reduces [6, 7].

Ostberg *et al.* [8] provided evidence of a dependence of discomfort glare on the difficulty of the task that is being executed. In their study, the same luminous source was reported as more discomforting if the concurrent task was relatively difficult. In a subsequent study, Gunnarson and Ostberg [9] described an interaction whereby discomfort glare was rated greater at higher task difficulty, with objective lighting conditions held constant. Likewise, ratings of task difficulty changed with variation in the perception of glare.

Sivak *et al.* [3] demonstrated an effect of task difficulty on discomfort glare by varying the size of a gap that subjects had to detect in a stimulus that was presented simultaneously with a glare source. More discomfort was reported when the gap was smaller and, thus, harder to locate. In their study, however, the gap location task was always required, hence making it difficult to infer whether, or to what extent, the effect of gap size on discomfort ratings depended on the gap stimulus being made explicitly relevant. In a follow-up work [10], gap size was varied in factorial combination with whether the gap location task was required, also including changes in the luminance of the gap stimulus as a second way of varying task difficulty. The results confirmed that glare evaluation was affected by the presence and nature of a concurrent visual task. In addition, the findings suggested that stimulus luminance influenced discomfort ratings even when the stimulus was not relevant to task performance.

Dugas and Wierwille [11] studied whether measures of reading performance are affected by short exposure text on a visual display terminal (VDT) when glare is present. In a preliminary experiment, reading passages were ranked according to subjective difficulty. Their findings suggested the presence of a reliable interaction between glare and task difficulty, but also led to hypothesise that, when faced with glare on a VDT, subjects may choose some method of compensation for its effects, such as reducing their time of exposure to the stimulus.

In a study conducted under artificial lighting conditions and utilising a visual display terminal containing paragraphs of randomly generated pseudo-words, Osterhaus and Bailey [12] found reduced visual performance under high levels of glare sensation. Their study also emphasised that decreases in task performance were likely to be expected with longer exposure to the glare source due to fatigue and potential distraction. Lynes [13] also contemplated a relationship between glare and distraction, which could affect the processing of visual information. In this context, however, Boyce [14] stated that human factors may be more important than physical factors in affecting visual performance, and suggested that discomfort glare studies based on task difficulty should remove the effects of distraction and motivation since they may have an indirect influence on visual task efficiency.

Rodriguez and Pattini [15] used a Reading Span Task (RST) displayed on a visual display terminal to show that measurable dependent differences in glare sensation could be detected upon consideration of task-related and behavioural factors, and changes in visual fixation point between the VDT and the glare source. Coherent with previous work [16], the size of the glare source was found to have statistically significant effects on the RST.

Similarly, Ko et al. [17] studied the effect of age, font size, and reflected glare from bright LED task light on performance for visually demanding text-based tasks on a computer screen. The VDT location was fixed, but subjects were allowed to move their posture. The results indicated that, as font size increased, so did performance, accuracy, and viewing distance, while perceived task difficulty decreased regardless of subjects' age. Adding reflective glare on the VDT led to a reduced viewing distance but had no effect on performance or accuracy.

In essence, a review of the literature suggests likely connections between visual performance, perceived task difficulty, and luminous conditions that could lead to discomfort glare [18]. However, the variation of such relationships over the time of day has not yet, to the authors' knowledge, been explored in detail.

Previous research by the authors [19, 20], conducted under a controlled laboratory setting, provided significant and substantive evidence of growing tolerance to luminance increases in artificial lighting as the day progresses. This trend was found to be particularly apparent for earlier chronotypes (a personal attribute reflecting individual circadian phases that indicates at what time of day physiological functions are activated [21]) and for subjects not having ingested caffeine. In interpreting these findings, it was hypothesised that the abstraction caused by the artificial lighting glare source, and the request to report visual discomfort in terms of Glare Sensation Votes (GSVs, i.e. benchmarks corresponding to the level of glare sensation experienced: 'Just Perceptible', 'Just Noticeable', 'Just Uncomfortable', and 'Just Intolerable'), could be among the causes for the large scattering detected when individual glare responses were regressed against the source luminance. In fact, although in the tests the GSV criteria were linked to time-span descriptors to aid participants giving more meaningful judgements [22], subjects had no task-related stimulus to associate their visual perception to.

On the basis of the literature and of previous findings, the study presented here sought to investigate the influence of inclusive features and difficulty of the visual task, and the potential effect of several temporal variables, on the subjective evaluation of glare sensation as the day progresses.

2. Method

2.1 Experimental Design

A systematic experimental design approach was adopted to respond to three research aims:

1. The first aim consisted in searching for temporal variations in the perceived level of glare sensation when subjects performed twelve visual tasks at distinct times of day. Individually for each visual task, differences between glare responses along the day were analysed so as to substantiate (or challenge) the previously detected increase in tolerance to artificial source luminance as the day progresses [19].
2. The second aim involved comparing temporal differences in glare response across groups of visual tasks at various times of day. Thus, the influence of task manipulation and difficulty over the postulated effect of time of day on glare sensation was analysed.

3. The third aim intended to study the influence of several temporal variables on the glare response provided by test subjects while engaging with visual tasks. This was to deepen the exploration of the role of temporal factors on glare sensation along the day [20].

The experimental method developed for this investigation derived from the procedures that Tuayacharoen and Tregenza [23] and Flannagan *et al.* [10] adopted to analyse the influence, respectively, of view interest and task difficulty on the perception of discomfort glare. An apparatus similar to earlier experiments was retained for this study [19].

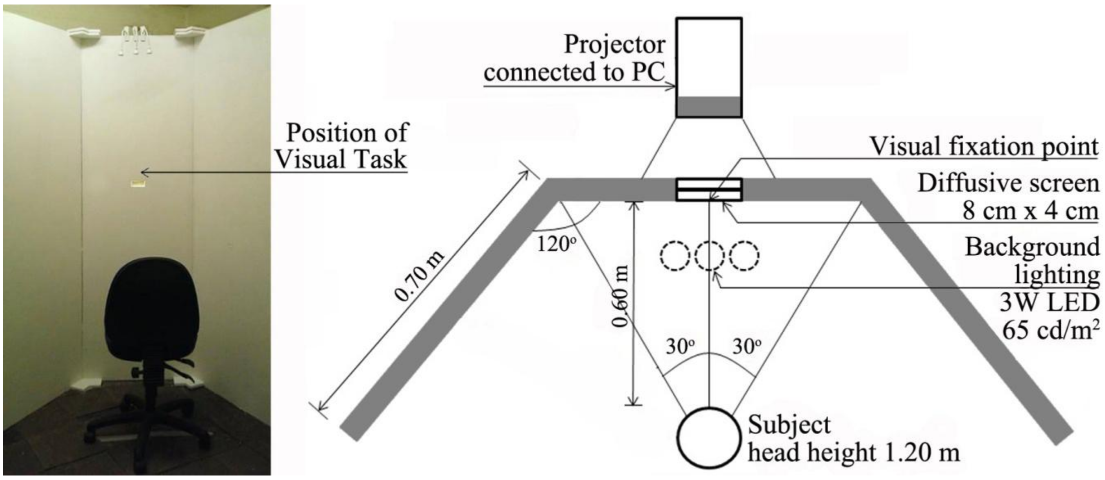


Figure 1. Experimental lighting chamber

The lighting chamber was semi-hexagonal in plan, with interior surfaces (2.70m in height) painted matte white (Fig. 1). Three 3W LED lamps, mounted from above, produced a background luminance of 65cd/m^2 with a warm white light (2,700 Kelvin). The subject's eye position was located at a height of 1.20m from the floor, facing a glare source represented by a small diffusive screen (8 cm x 4 cm) made from two sheets of tracing paper; this was mounted in front of a projector connected to a computer controlled by the experimenter. The source of glare subtended an angle at the eye of 0.009 steradians. The visual stimulus comprised reading tasks of various difficulty that subjects were requested to perform at different times of day when providing votes of glare sensation. To refine the setup, a series of pilot tests (N= 3) was completed to determine the position of the visual task and the set luminance of the diffusive screen. A study by Iwata and Tokura [24] showed that a source luminance placed above the fixation point triggered a lower level of visual discomfort than the same source positioned below the line of sight. Coherent with this finding, the visual task was placed above the glare source, with a displacement angle of 2.87^o from the subject's viewpoint. Indeed, a series of trials performed at lower displacement angles revealed that the differences between the votes of perceived glare sensation reported by subjects when varying visual tasks did not allow robust statistical analysis. The pilot tests also contributed to define the set luminance of the diffusive screen that was kept at a constant value of 10541.31cd/m^2 throughout the experimental procedure. This corresponds to an IES-GI (Illuminating Engineering Society Glare Index) of 22 that, on the Hopkinson scale [25], describes the source as being 'Just Uncomfortable'. In fact, no substantial variations in glare response between different visual tasks were reported by pilot test subjects when the source luminance was set at a constant level corresponding to a GSV of 'Just Perceptible' or 'Just Noticeable'.

2.2 Visual Task Difficulty

Visual task difficulty can be intended as a hindrance to the ability of an observer to extract information from a visual stimulus. According to Boyce [1], increased task difficulty can cause excessive pressure on the lens of the eye for it to increase its optical power and keep the retinal image sharp, this resulting in muscular fatigue and symptoms of visual discomfort. Coherent with the literature [6, 26, 27], the measureable dependencies of visual task difficulty were represented in this study by the size and contrast of the characters contained within the stimulus. Abstract numerical detection methods were adopted in the visual tasks [28, 29, 30], since they provide the advantages of being relatively simple and holding no real meaning for participants to interpret. This helped masking the experiment from confounding variables (e.g., learning, interest, boredom) as compared with their reading counterparts.

Twelve visual tasks were used in this study, divided in three groups of four tasks each (Table 1). The font of the characters was set at Verdana [15]. In the first group, Tasks 1 to 4 varied in *font size*, which was set respectively at size 10 (Task 1), 8 (Task 2), 6 (Task 3) and 4 (Task 4). In the second group, Tasks 5 to 8 varied in *internal contrast* [6, 26], saturating the background of the visual stimulus by 20% (Task 5) and 40% (Task 6), and distorting task information with vertical (Task 7) and horizontal (Task 8) lines [31]. In the third group, Tasks 9 to 12 varied in *character contrast*, which was set respectively at 50% (Task 9), 35% (Task 10), 25% (Task 11) and 15% (Task 12). Each visual task was printed on a strip of white paper, with dimensions of 13.7 cm x 1.7 cm, subtending a solid angle at the eye of 0.006 steradians. This allowed subjects to fully accommodate the visual task, reducing the likelihood of their changing head position while completing the test. Characters were arranged in one row and nine columns, and were spaced at 1.3 cm intervals so as to give clarity to the task.

Coherent with the procedure adopted by Tuayacharoen and Tregenza [23], before the tests the twelve visual tasks were presented in a randomised order to an independent group of observers (N= 20, a different sample of subjects from those used in the experiment) that provided scores of perceived task difficulty on an ascending four-point rating scale (from 1= least difficult to 4= most difficult). The visual tasks were presented to subjects with the same size, font, and contrast manipulations as those used in the experiment. The independent group of observers was selected by convenience sampling among University doctoral students, and varied in nationality and cultural background. The mean and standard deviation (SD) scores from these independent assessments are given in Table 1.

Table 1. Visual Tasks

Task No.	Perceived Difficulty Mean (SD)	Visual Task
1	1.05 (0.22)	4 6 1 7 9 8 2 5 3
2	1.95 (0.22)	6 9 3 4 5 1 2 8 7
3	3.00 (0.00)	9 2 3 7 5 6 1 4 8
4	4.00 (0.00)	2 7 1 9 4 8 6 3 5
5	1.00 (0.00)	6 7 5 4 9 1 8 3 2
6	2.75 (0.79)	6 9 4 7 3 8 1 5 2
7	2.80 (0.70)	4 9 5 6 8 3 2 1 7
8	3.45 (0.83)	5 7 3 6 2 4 9 1 8
9	1.00 (0.00)	6 5 9 1 7 8 3 2 4
10	2.00 (0.00)	7 2 9 3 4 8 1 6 5
11	3.00 (0.00)	3 1 2 6 9 8 7 5 4
12	4.00 (0.00)	3 9 5 4 1 7 6 2 8

2.3 Experimental Procedure

Subjects were requested to participate to four test sessions, each held on a different day and time, under a randomised sequence. The sessions were scheduled at: 09:00-09:30 (Morning); 12:00-12:30 (Afternoon A); 15:00-15:30 (Afternoon B); 18:00-18:30 (Evening). The random sequence was assigned by the Randomizer software, which shuffled all independent variables (day and time of sessions, and order of tasks) to create unsystematic behavioural effects on subjects. The randomisation sought to counterbalance possible influences of confounding variables and make the study more sensitive to detect the effects of interest [32].

Before the start of their first test session, each subject was asked to perform a *pre-test procedure*. As part of this, a clear set of instructions was given on a participant information sheet that included a definition of discomfort glare, the meaning of the Glare Sensation Votes (GSV) benchmarks linked to time-span descriptors, and an explanation of how the experiment would run. Subjects were requested to sign a consent form and to provide general demographic information (e.g., age, gender, ethnicity) on a paper questionnaire. In the pre-test procedure, a reference visual task was also presented to the subject. This contained nine numbers ranging from 1-9 in ascending order from left to right – the font was black Verdana and its size was 10 – placed against a white background. The subject was instructed to look at the visual task positioned above the glare source – whose luminance was kept at a constant 10541.31cd/m^2 – and to place a mark corresponding to their glare sensation on a 10 cm long continuous GSV visual analogue scale (VAS) that was handed out by the experimenter. In addition, the participant was asked to indicate their self-assessed level of fatigue on another visual analogue scale that was calibrated to the size of the GSV one (Fig. 2). Visual analogue scales have already been used in glare studies [5] as a method to measure variables that, from the subject's perspective, range across a continuum of values. Operationally, a VAS is presented as a horizontal line 'anchored' by word descriptors at each end. The subject is requested to mark on the line the point that they feel best represents their perception of their current state. After the subject had provided their assessments of glare sensation and fatigue on the visual analogue scales, a relaxation period of two minutes was allowed.

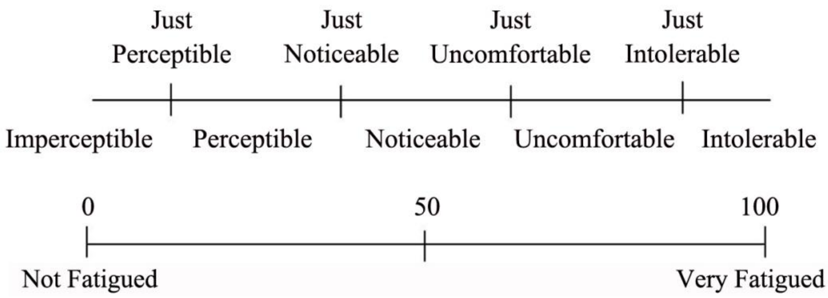


Figure 2. Continuous GSV and Fatigue visual analogue scales

The pre-test procedure allowed subjects to familiarise themselves with the experimental setup, provide a proper understanding of discomfort glare and the GSV thresholds by immediate experience [23], introduce the use of visual analogue scales, and clarify directly with the experimenter any doubt or questions they may have before any actual data were recorded.

At the start of each *experimental test session*, participants were requested to adjust the stool so that their head was properly located at the viewing position. The subject was then asked to direct their gaze towards a fixation point marked above the glare source. After a short time provided for the subject to adapt to the luminous conditions, the 12 visual tasks were presented, one after the other, in a randomised sequence. For each task, the subject was asked by the experimenter to locate three random numbers and to mark a vertical line on a GSV visual analogue scale to describe their glare sensation. At the end of each test session, participants were also requested to provide self-assessments of several temporal variables (i.e., factors that may vary with the time of day [20]) measured on continuous visual analogue scales presented to the subjects with the same size of the GSV one. The length of the scales used to measure all the variables considered in this study was calibrated in order to allow more robust statistical analysis. The visual analogue scales and relative descriptors for the temporal variables that were developed for this study asked subjects to indicate their self-assessment in terms of perceived level of fatigue (ranging from *not fatigued* to *very fatigued*), their food intake (from *no food* to *a lot of food*) and caffeine ingestion (from *no caffeine* to *a lot of caffeine*) before each test, their mood (from *bad mood* to *good mood*), the prior daylight exposure (from *no exposure* to *a lot of exposure*) and the prevailing sky condition (from *fully overcast* to *clear sky*) that they had experienced in the hours preceding each session.

Before the tests, photometric measures were taken from the participant's eye position using a calibrated Minolta LS-100 luminance meter mounted on a tripod. The mean background luminance was calculated from 17 measurements taken on a regular grid symmetrical about the central fixation point and extending across the width of the experimental apparatus. The luminance of the diffusive screen was measured averaging three direct point readings [19].

In order to allow an easy comparison between votes of glare sensation and conventional glare indexes, Tuaycharoen and Tregenza [33] proposed that Glare Sensation Votes (GSVs) given by test subjects could be 'scaled' to equivalent Glare Response Votes (GRVs). The index resulting from their work – GRV(DGI) scaled to the DGI (Daylight Glare Index) [34] – is applicable to large luminance sources such as daylight from windows. However, since the study here described was conducted in an artificial laboratory setting, the data obtained by Hopkinson [25, 35] and by Tokura *et al*. [36] – which allow to relate GSVs to corresponding values of IES-GI – were used to convert the scale of GSVs into a GRV(IES-GI) score, thereby creating a Glare Response Vote formula scaled to the IES-GI glare index that is suitable for assessing perceived glare sensation from small artificial lighting sources (1):

$$\textbf{(1)}\ \mathbf{GRV(IES - GI) = 6 \cdot GSV + 10}$$

A total of 20 subjects (all postgraduate students, 6 male and 14 female) volunteered to take part to the experiment. Subjects varied in nationality and cultural background, the mean age was 28.05 (SD= 3.10), 10 wore glasses or corrective lenses, and all were self-certified as having no other eye problems. This sample size is consistent with the literature, since statistically robust results using visual tasks have been obtained in studies based on an analogous number of participants [e.g., 3, 10, 11, 12, 15, 23]. The criterion adopted for the selection of participants was purposive sampling, and subjects were recruited via an online advertisement addressed to all postgraduate students in Architecture. There were no criteria used for the exclusion of volunteers. All tests were performed during the month of May, a period of mixed weather, varying from overcast to clear skies and bright sunshine.

3. Results

3.1 Temporal Variation of Glare Response for Each Visual Task

Among other visual tasks, Figure 3 plots on the y-axis the GRV(IES-GI) calculated on the basis of the votes of glare sensation reported by subjects when presented with Task 3. On the x-axis, the figure provides the times of day when glare assessments were given.

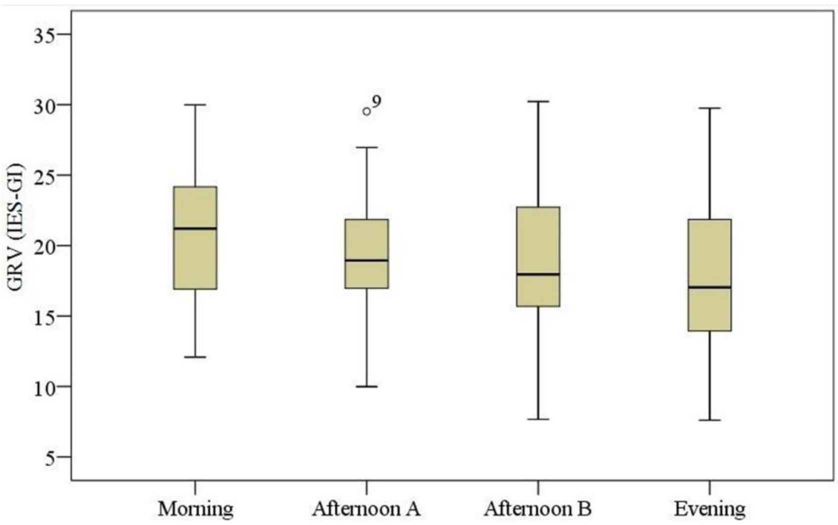


Figure 3. Boxplots of GRV(IES-GI) for *Task 3*

https://doi.org/10.1016/j.buildenv.2015.09.021
https://escholarship.org/uc/item/5g20q4dg

Inspection of the boxplots suggested, for all the 12 visual tasks used in this study, a tendency for statistical values to correspond to lower levels of GRV(IES-GI) as the day progresses. That is, when subjects were presented with the same visual task at later times of day, the constant source luminance was reported as being less visually discomfortable.

Graphical (Q-Q plot) and statistical (Shapiro-Wilk and Kolmogorov-Smirnov tests) analysis of the data revealed normal distributions around the mean and that homogeneity of variance of the dependent variables (glare responses at all test sessions for each visual task) across the independent variables (times of day) was not statistically significant (p>0.05). In addition, tests of homogeneity of regression slopes between the dependent variables and the covariate (fatigue) showed that also this postulation was satisfied [37]. Since basic assumptions were not violated, parametric tests were adopted for this analysis [38].

A Multivariate Analysis of Covariance (MANCOVA) was initially run to compare the GRV(IES-GI) scores for each visual task at all test sessions, while controlling for the effect of self-assessed fatigue. A review of literature and previous findings [19], in fact, provided reasons to suspect that glare sensation could be influenced primarily by this temporal variable. Therefore, the MANCOVA was selected to guard against this confounding factor not experimentally controlled over the independent variables. The results from the MANCOVA showed no statistically significant differences: $F(36, 189.82) = 0.94$, $p = 0.58$ n.s., Wilk's $\Lambda = 0.62$, $p\eta^2 = 0.15$. However, when controlling for the effect of fatigue, a statistically significant difference was detected: $F(12, 64) = 2.00$, $p = 0.04$*, Wilk's $\Lambda = 0.74$, $p\eta^2 = 0.27$. An Analysis of Covariance (ANCOVA) was then performed to search for significant differences between the independent variables for each visual task. Table 2 presents the inferential data from the ANCOVA, whereby the 'Fatigue' column provides information on the effect of the covariate on the dependent variable GRV(IES-GI), and the 'Time of Day' column reports the effect of experimental interest once adjusted for the covariate. For each column, Table 2 shows the test statistic (F) and the degrees of freedom (df), the p-value, and the effect size, which was calculated by making use of equivalence between the standardised measure of the detected difference between independent groups and the partial-eta square ($p\eta^2$) coefficient [39]. The interpretation of the outcome was derived from Ferguson [40], where benchmarks are given for small, moderate, and large effect sizes ($p\eta^2 \geq 0.04$, 0.25, and 0.64 respectively). Values of $p\eta^2$ below 0.04 denote not substantive (i.e., not practically relevant) influences.

Table 2. ANCOVA for each visual task

	Fatigue			Time of Day		
Task	***F* (df)**	***p*-value**	**Effect Size ($p\eta^2$)**	***F* (df)**	***p*-value**	**Effect Size ($p\eta^2$)**
1	0.53 (3)	0.66 n.s.	0.02	0.32 (1)	0.58 n.s.	0.00
2	0.74 (3)	0.53 n.s.	0.03	0.21 (1)	0.65 n.s.	0.00
3	2.62 (3)	0.06 n.s.	0.10	7.57 (1)	0.01**	0.09
4	2.37 (3)	0.08 n.s.	0.09	8.48 (1)	0.01**	0.10
5	5.41 (3)	0.01**	0.19	11.24 (1)	0.00***	0.13
6	2.43 (3)	0.07 n.s.	0.09	5.45 (1)	0.02*	0.07
7	1.90 (3)	0.14 n.s.	0.07	13.69 (1)	0.00***	0.15
8	1.59 (3)	0.20 n.s.	0.06	8.95 (1)	0.01**	0.11
9	1.42 (3)	0.24 n.s.	0.05	6.04 (1)	0.02*	0.08
10	2.85 (3)	0.04*	0.10	6.61 (1)	0.01**	0.08
11	4.62 (3)	0.01**	0.17	7.41 (1)	0.01**	0.09
12	3.10 (3)	0.03*	0.11	4.39 (1)	0.04*	0.06

*weakly significant; **significant; ***highly significant; n.s. not significant
$p\eta^2<0.04$= negligible; $0.04\leq p\eta^2<0.25$= small; $0.25\leq p\eta^2<0.64$= moderate; $p\eta^2\geq 0.64$= large

The 'Fatigue' column of Table 2 shows two statistically significant (Tasks 5 and 11) and two weakly significant differences (Tasks 10 and 12), all with a small yet substantive effect size. The $p\eta^2$ is practically relevant in 10 out of 12 cases ($p\eta^2<0.04$ only for Tasks 1 and 2). After controlling for the effect of fatigue, the 'Time of Day' column presents two highly significant (Tasks 5 and 7), five significant (Tasks 3, 4, 8, 10 and 11), and three weakly significant differences (Tasks 6, 9, and 12), all corresponding to practically relevant effects. To isolate the influences detected, contrasts were made using pairwise comparisons [41].

The ANCOVA showed no statistical or practical significance for Tasks 1 and 2; therefore, these tasks were excluded from the subsequent analysis.

In consideration of the experiment-wise error rate caused by the significance level inflating across multiple tests carried out on the same data – which was calculated as $1-(0.95)^n= 0.95$ (thus, risking a 95% probability of making at least one Type I error), where n= 60, i.e. the number of pairwise comparisons performed – the Least Significant Difference (LSD) method was applied. This test calculates the smallest significant difference (LSD) between two means and declares statistically significant any detected difference that is larger than the LSD. This test has more statistical power compared to other post-hoc correction methods since the alpha-level (*p*-value) is not adjusted [42]. Conversely, the use of Bonferroni correction would have considerably inflated the risk of occurrence of a Type II error [43]. The effect size of each contrast was measured by the Cohen's d, and was considered small, moderate, and large for, respectively, d≥0.41, 1.15, and 2.70 [40].

Based on graphical inspection, the tests adopted the alternative hypothesis that the constant source luminance is reported by subjects as being less visually discomfortable as the day progresses. Table 3 shows the results of the pairwise comparisons for each visual task, providing the adjusted mean (M, controlled for the effect of fatigue) and the standard deviation (SD) for the GRV(IES-GI) score at each test session, the difference between the means (ΔM) and the outcome of its statistical significance (Null Hypothesis Significance Testing (NHST), *p*-value calculated with a one-tailed test), the lower (CI_L) and upper (CI_U) 95% confidence intervals for the ΔM, and the effect size (d).

Table 3. Pairwise comparisons between test sessions for each visual task

Task	Test Sessions	M (SD)	M (SD)	ΔM^{NHST}	$[CI_L,CI_U]$		Effect Size (d)
	Morn. vs. Aft. A	21.20 (2.18)	19.87 (2.16)	1.33 n.s.	-1.79	4.42	0.61
3	Morn. vs. Aft. B	21.20 (2.18)	18.22 (2.13)	2.98*	-0.22	6.17	1.38
	Morn. vs. Even.	21.20 (2.18)	16.73 (2.22)	4.47**	1.14	7.78	2.01

https://doi.org/10.1016/j.buildenv.2015.09.021
https://escholarship.org/uc/item/5g20q4dg

	Aft. A vs. Aft B	19.87 (2.16)	18.22 (2.13)	1.65 n.s.	-1.50	4.84	0.77
	Aft. A vs. Even.	19.87 (2.16)	16.73 (2.22)	3.14*	-0.14	6.44	1.43
	Aft. B vs. Even.	18.22 (2.13)	16.73 (2.22)	1.49 n.s.	-1.65	4.62	0.68
4	Morn. vs. Aft. A	24.13 (2.54)	24.43 (2.51)	-0.30 n.s.	-3.92	3.33	-0.12
	Morn. vs. Aft. B	24.13 (2.54)	20.62 (2.49)	3.51*	-0.22	7.24	1.40
	Morn. vs. Even.	24.13 (2.54)	20.50 (2.59)	3.63*	-0.26	7.50	1.42
	Aft. A vs. Aft B	24.43 (2.51)	20.62 (2.49)	3.81*	0.11	7.52	1.52
	Aft. A vs. Even.	24.43 (2.51)	20.50 (2.59)	3.93*	3.34	7.77	1.57
	Aft. B vs. Even.	20.62 (2.49)	20.50 (2.59)	0.12 n.s.	-1.52	1.62	0.05
5	Morn. vs. Aft. A	20.55 (1.61)	19.78 (1.60)	0.77 n.s.	-1.54	3.05	0.48
	Morn. vs. Aft. B	20.55 (1.61)	16.41 (1.58)	4.14***	2.79	6.49	2.47
	Morn. vs. Even.	20.55 (1.61)	16.97 (1.64)	3.56**	1.11	6.03	2.20
	Aft. A vs. Aft B	19.78 (1.60)	16.41 (1.58)	3.38**	1.04	5.73	2.12
	Aft. A vs. Even.	19.78 (1.60)	16.97 (1.58)	2.82**	0.39	5.25	1.77
	Aft. B vs. Even.	16.41 (1.58)	16.97 (1.64)	-0.56 n.s.	-2.87	1.76	-0.35
6	Morn. vs. Aft. A	22.10 (1.98)	21.59 (1.96)	0.51 n.s.	-2.31	3.34	0.26
	Morn. vs. Aft. B	22.10 (1.98)	18.73 (1.94)	3.37**	0.45	6.26	1.72
	Morn. vs. Even.	22.10 (1.98)	19.18 (2.02)	2.92*	-0.11	5.94	1.46
	Aft. A vs. Aft B	21.59 (1.96)	18.73 (1.94)	2.86*	-0.04	5.73	1.47
	Aft. A vs. Even.	21.59 (1.96)	19.18 (2.02)	2.41*	-0.51	5.40	1.21
	Aft. B vs. Even.	18.73 (1.94)	19.18 (2.02)	-0.45 n.s.	-3.28	2.42	-0.23
7	Morn. vs. Aft. A	22.41 (1.72)	22.20 (1.70)	0.21 n.s.	-2.28	2.62	0.12
	Morn. vs. Aft. B	22.41 (1.72)	20.20 (1.69)	2.21*	-0.37	4.67	1.30
	Morn. vs. Even.	22.41 (1.72)	19.85 (1.75)	2.56*	-0.10	5.16	1.48
	Aft. A vs. Aft B	22.20 (1.70)	20.20 (1.69)	2.00 n.s.	-0.51	4.50	1.18
	Aft. A vs. Even.	22.20 (1.70)	19.85 (1.75)	2.35*	-0.23	4.97	1.36
	Aft. B vs. Even.	20.20 (1.69)	19.85 (1.75)	0.35 n.s.	-2.08	2.87	0.20
8	Morn. vs. Aft. A	22.73 (1.97)	22.73 (1.95)	0.00 n.s.	-2.80	2.79	0.00
	Morn. vs. Aft. B	22.73 (1.97)	20.62 (1.93)	2.11 n.s.	-0.78	4.98	1.08
	Morn. vs. Even.	22.73 (1.97)	20.13 (2.01)	2.60*	-0.41	5.58	1.31
	Aft. A vs. Aft B	22.73 (1.95)	20.62 (1.93)	2.11 n.s.	-0.74	4.98	1.09
	Aft. A vs. Even.	22.73 (1.95)	20.13 (2.01)	2.60*	-0.36	5.57	1.31
	Aft. B vs. Even.	20.62 (1.93)	20.13 (2.01)	0.49 n.s.	02.32	3.32	0.25
9	Morn. vs. Aft. A	17.78 (1.84)	18.59 (1.82)	-0.81 n.s.	3.44	1.80	-0.44
	Morn. vs. Aft. B	17.78 (1.84)	16.41 (1.80)	1.37 n.s.	-0.13	4.06	0.75
	Morn. vs. Even.	17.78 (1.84)	16.01 (1.88)	1.77 n.s.	-1.03	4.54	0.95
	Aft. A vs. Aft B	18.59 (1.82)	16.41 (1.80)	2.18*	-0.47	4.86	1.20
	Aft. A vs. Even.	18.59 (1.82)	16.01 (1.88)	2.58*	-0.19	5.34	1.39
	Aft. B vs. Even.	16.41 (1.80)	16.01 (1.88)	0.40 n.s.	-2.23	3.02	0.33
10	Morn. vs. Aft. A	21.04 (2.18)	20.31 (2.16)	0.73 n.s.	-2.39	3.82	0.34
	Morn. vs. Aft. B	21.04 (2.18)	18.01 (2.14)	3.03*	-0.18	6.22	1.40
	Morn. vs. Even.	21.04 (2.18)	16.62 (2.22)	4.42**	1.08	7.74	2.01
	Aft. A vs. Aft B	20.31 (2.16)	18.01 (2.14)	2.30 n.s.	-0.86	5.49	1.07
	Aft. A vs. Even.	20.31 (2.16)	16.62 (2.22)	3.69**	0.41	7.00	1.68
	Aft. B vs. Even.	18.01 (2.14)	16.62 (2.22)	1.39 n.s.	-1.73	4.53	0.62
11	Morn. vs. Aft. A	21.55 (2.13)	20.73 (2.11)	0.82 n.s.	-2.18	3.85	0.39
	Morn. vs. Aft. B	21.55 (2.13)	17.55 (2.08)	4.00**	0.90	7.10	1.90

	Morn. vs. Even.	21.55 (2.13)	16.22 (2.17)	5.33***	2.09	8.54	2.48
	Aft. A vs. Aft B	20.73 (2.11)	17.55 (2.08)	3.18*	0.12	6.28	1.52
	Aft. A vs. Even.	20.73 (2.11)	16.22 (2.17)	4.51**	1.32	7.71	2.11
	Aft. B vs. Even.	17.55 (2.08)	16.22 (2.17)	1.33 n.s.	-1.69	4.39	0.61
12	Morn. vs. Aft. A	23.29 (2.55)	23.03 (2.53)	0.26 n.s.	-3.39	3.90	0.10
	Morn. vs. Aft. B	23.29 (2.55)	19.38 (2.50)	3.91*	0.14	7.65	1.55
	Morn. vs. Even.	23.29 (2.55)	18.43 (2.60)	4.86**	0.96	8.78	1.89
	Aft. A vs. Aft B	23.03 (2.53)	19.38 (2.50)	3.65*	-0.08	7.38	1.45
	Aft. A vs. Even.	23.03 (2.53)	18.43 (2.60)	4.60**	0.75	8.49	1.79
	Aft. B vs. Even.	19.38 (2.50)	18.43 (2.60)	0.95 n.s.	-2.70	4.66	0.37

*weakly significant; **significant; ***highly significant; n.s. not significant
$d<0.41$= negligible; $0.41\leq d<1.15$= small; $1.15\leq d<2.70$= moderate; $d\geq 2.70$= large

Coherent with previous findings [19], the descriptive statistics of Table 3 (i.e., means and their differences, ΔM) show for all visual tasks a tendency for the GRV(IES-GI) to decrease at later times of day, signalling a greater tolerance to source luminance as the day progresses. In fact, the ΔM is positive in all but five pairwise comparisons; these five comparisons are, however, consistently not statistically significant. The differences in GRV(IES-GI) calculated between each test session are highly significant in two cases, significant in 11 cases, weakly significant in 20 cases, and not significant in 27 cases, out of a total of 60 comparisons. The effect sizes are moderate ($1.15\leq d<2.70$) in 34 cases and small ($0.41\leq d<1.15$) in 11 cases. Negligible influences ($d<0.41$) are detected in 15 cases. All the non-substantive effect sizes correspond to not statistically significant pairwise comparisons.

For six visual tasks (Tasks 3, 7, 8, 10, 11 and 12), the largest statistically significant differences measured by the effect sizes (and ΔMs) occurred when comparing glare responses at the Morning and Evening sessions, corresponding to a temporal gap of 9 hours between tests. Therefore, descriptive and inferential analysis seems to suggest that, for these visual tasks, the length of the time interval between test sessions may have a direct influence on the increased tolerance to source luminance reported by test subjects as the day progresses.

Contrary to previous tasks, the largest differences in effect sizes (and ΔMs) for Tasks 5 and 6 were detected between the Morning and Afternoon B sessions, corresponding to a time gap of

6 hours between the tests. For these tasks, Figures 4 and 5 plot, on the y-axis, the individual GRV(IES-GI) scores calculated from the votes of glare sensation provided by each subject in the Afternoon A, Afternoon B and Evening sessions and, on the x-axis, the Glare Response Votes provided by participants in the Morning. In both figures, the null hypothesis line is plotted along the diagonal, representing no differences between the GRV(IES-GI) given by subjects at each session. Coherent with the literature [23, 33, 44], the figures also test linear interpolations of the data as a way to visualise the deviation from the null hypothesis and estimate the scatter in the results (as confirmed by the low coefficients of determination).

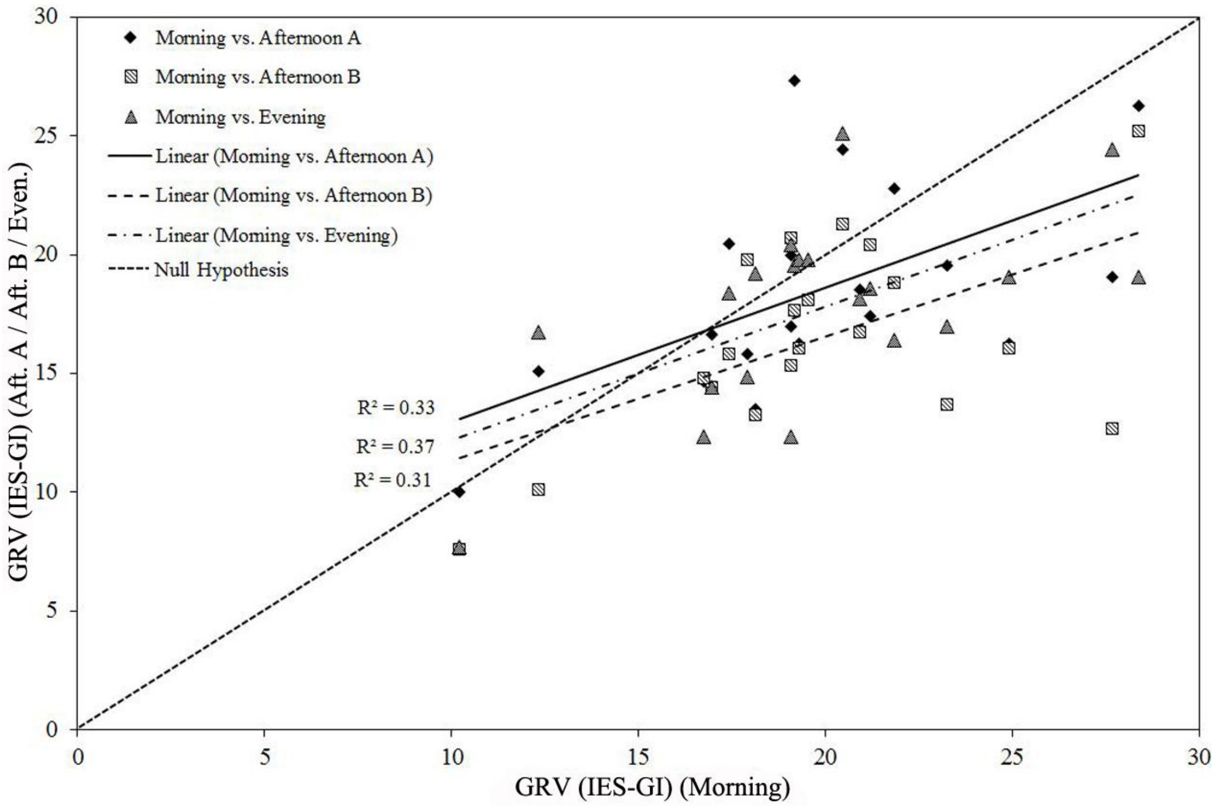


Figure 4. Comparison of GRV(IES-GI) for *Task 5*

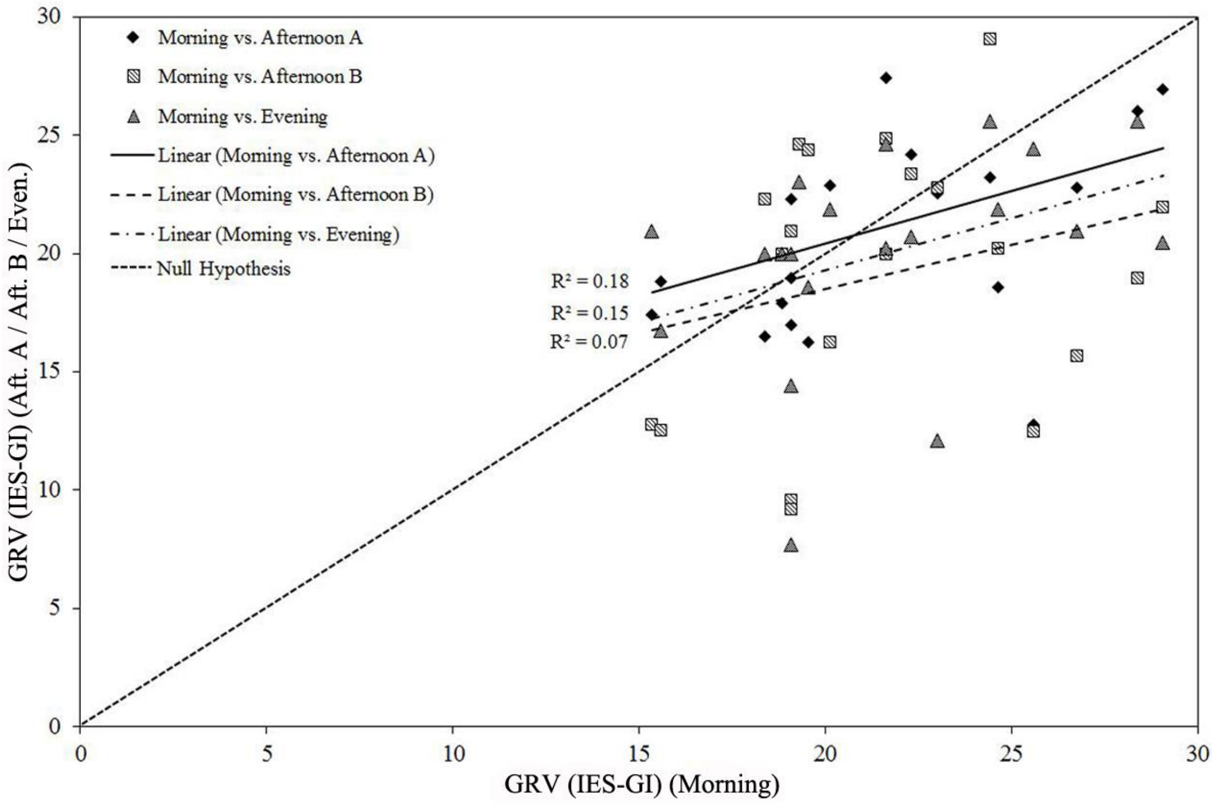


Figure 5. Comparison of GRV(IES-GI) for *Task 6*

Although some GRV(IES-GI) scores are above the null hypothesis line, signalling that some subjects provided higher votes of glare sensation in the Morning, the interpolated lines for each comparison are largely below it and with lower slopes. This confirms that the time interval between sessions may have a direct influence on the increased tolerance to source luminance. Also, at lower levels of Glare Response Vote, both figures show the linear fits above the null hypothesis, leading to speculate that the effect of time of day may be dependent on the level of visual discomfort perceived; that is, as the glare sensation increases, so does the effect of the temporal gap between sessions on tolerance to source luminance. Interestingly, the spread of the data inflates from Task 5 to Task 6, resulting in considerably lower R^2 for the latter. This may possibly result from increased difficulty – the tasks were independently assessed with a difficulty of 1.00 (0.00) for Task 5 and 2.75 (0.79) for Task 6 – thereby leading to a larger scatter of GRV(IES-GI). However, there is not sufficient evidence to infer whether the increased dispersion of Glare Response Votes for Task 6 was caused by task difficulty or by the effect of other confounding factors not experimentally controlled.

For the two remaining tasks, Tasks 4 and 9, the largest significant differences in effect sizes (and ΔMs) were detected between the Afternoon A and Evening, i.e. an interval of 6 hours between sessions. For Task 9, however, only two comparisons were statistically significant. For Task 4, all significant contrasts resulted in a rather narrow range of moderate effect sizes ($1.40 \leq d \leq 1.57$). A similar finding is also apparent for the other tasks that had been independently rated as more difficult in each group (Task 4 was assessed with a difficulty of 4.00 (0.00)). In fact, in the case of Task 8 – whose difficulty was rated as 3.45 (0.83) – the effect size is d= 1.31 for both statistically significant differences. For Task 12 – rated with a difficulty of 4.00 (0.00) – the effect size is between 1.45 and 1.89 across all the statistically significant comparisons. The adjusted means for Tasks 4, 8, and 12 correspond almost invariably to the highest GRV(IES-GI) at all sessions, a result coherent with the literature [3]. This may again suggest that, for more visually demanding tasks, the effect of time of day on temporal variation of glare response may be influenced by other factors such as task difficulty.

3.2 Temporal Variation of Glare Response across Visual Tasks

A further analysis was performed to compare temporal variations of glare response across visual tasks at each time of day, based on the manipulation of their inclusive features and their respective independently-assessed task difficulty. Figure 6 plots for each test session, on the y-axis, the GRV(IES-GI) scores, and, on the x-axis, the three groups of tasks with variation of *font size* (Tasks 1-4), *internal contrast* (5-8), and *character contrast* (9-12).

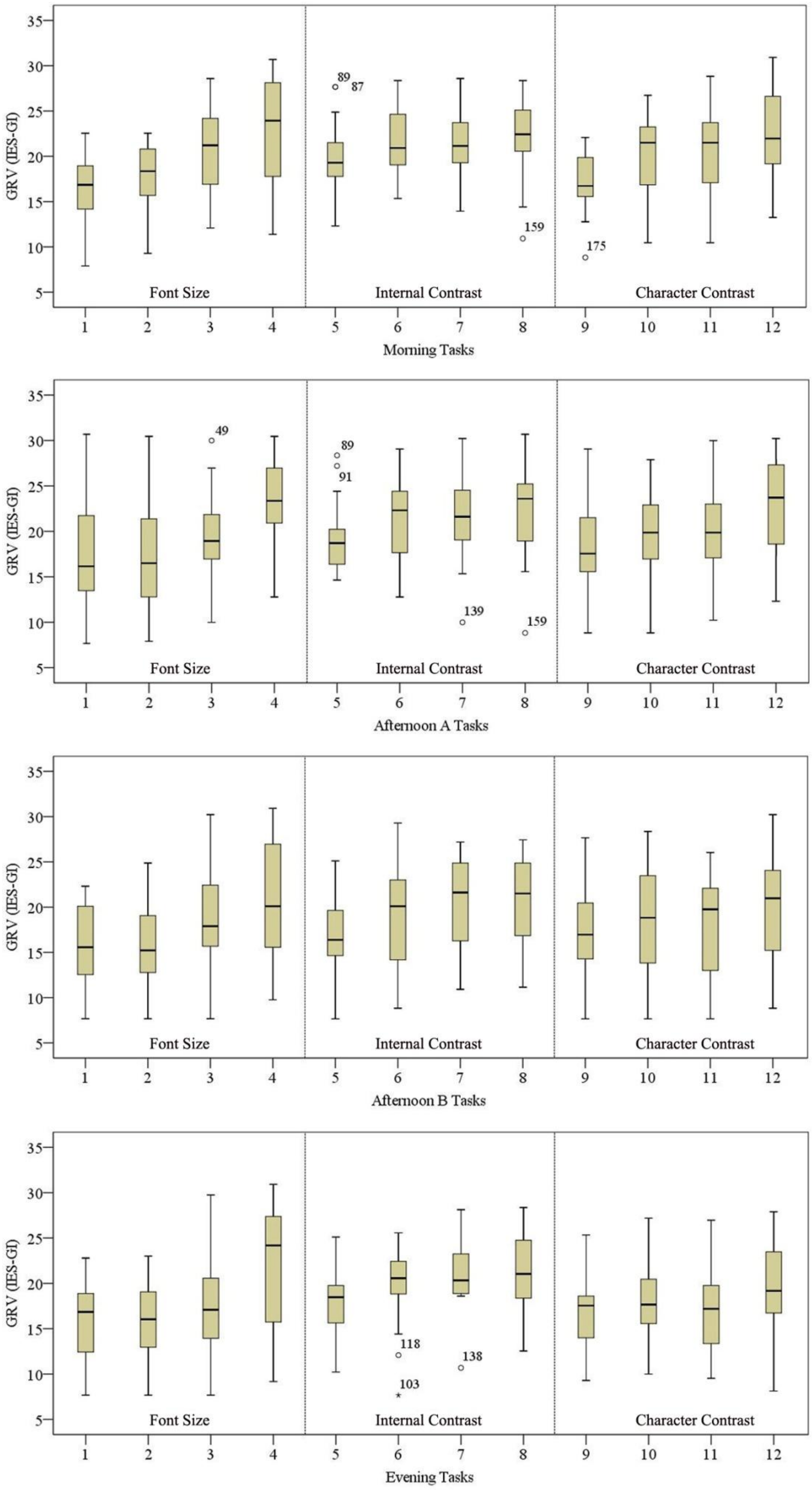


Figure 6. Boxplots of GRV(IES-GI) for groups of visual tasks at each test session

For all sessions, a tendency for increasing statistical values of GRV(IES-GI) can be observed in the first group (Tasks 1-4) when visual tasks presented to subjects feature a progressively smaller *font size* (Task 1 had a font size of 10, which was reduced by 2-point decrements for other tasks). For the *internal contrast* group (Tasks 5-8), the boxplots suggest that when the background saturation of the visual stimulus increases – from 20% of Task 5 to 40% of Task 6 – so does the glare sensation. This trend of intensifying glare response further strengthens when the task information is distorted with vertical (Task 7) and horizontal lines (Task 8). Finally, in terms of *character contrast* (Tasks 9-12), a tendency is detectable where, as the contrast of the characters decreases (from 50% of Task 9 to 15% of Task 12), the glare sensation increases. Almost invariably for all groups, the identified tendencies present temporal consistency; that is, similar trends across tasks can be recognised at all test sessions.

A Multivariate Analysis of Variance (MANOVA) was performed whereby all the dependent variables (GRV(IES-GI) scores from the four test sessions) were compared with the grouping variables (visual tasks). As the 12 visual tasks were presented to participants in a randomised order – and subjects only performed one test session per day – it was assumed for this analysis across visual tasks that the potential influence of temporal variables (e.g., fatigue) and other confounding factors (e.g., learning) on glare sensation were masked from the experiment. Therefore, covariates were considered as controlled [37].

The MANOVA for the visual tasks detected at least one statistically significant difference: F (44, 862.75)= 1.76, p= 0.00**, Wilk's Λ= 0.72, $p\eta^2$= 0.08. Post hoc ANOVA was then performed, breaking down the individual dependent variables to analyse the effect of time of day on glare response. The results for each session are as follows: Morning: F (11)= 4.94, p= 0.00**, $p\eta^2$= 0.19; Afternoon A: F (11)= 3.73, p= 0.00***, $p\eta^2$= 0.15; Afternoon B: F (11)= 2.50, p= 0.01**, $p\eta^2$= 0.11; Evening: F (11)= 3.54, p= 0.00***, $p\eta^2$= 0.15. These findings suggested that not only may glare response be dependent on task difficulty, but that this relationship appears to vary over time, as per the shifting effect sizes detected at each session. To determine differences in temporal variation of glare response across visual tasks, it was decided not to perform multiple pairwise comparisons since the use of standard or sequential Bonferroni correction, required to avoid inflating the significance level, could have made the tests susceptible to Type II errors [43]. Instead, the tasks were divided into 3 grouping variables – *font size* (Size)= Tasks 1-4; *internal contrast* (I. Contrast)= Tasks 5-8; *character contrast* (C. Contrast)= Tasks 9-12 – that were compared in an additional ANOVA (Table 4).

Table 4. ANOVA for groups of visual tasks at each test session

Session	Source	Sum of Squares	df	Mean Square	F	*p*-value	Effect Size ($p\eta^2$)
Morning	Size	103.57	3	34.52	8.62	0.00***	0.25
	I. Contrast	9.90	3	3.30	1.16	0.33 n.s.	0.05
	C. Contrast	56.09	3	18.70	4.84	0.00**	0.16
Afternoon A	Size	107.37	3	35.79	6.80	0.00***	0.21
	I. Contrast	17.70	3	5.90	1.56	0.21 n.s.	0.06
	C. Contrast	35.77	3	11.92	2.76	0.05*	0.10
Afternoon B	Size	75.42	3	25.14	4.50	0.01**	0.15
	I. Contrast	41.56	3	13.85	3.15	0.03*	0.11
	C. Contrast	17.50	3	5.83	0.85	0.47 n.s.	0.03
Evening	Size	80.80	3	26.93	4.91	0.00**	0.16
	I. Contrast	24.43	3	8.14	2.69	0.05*	0.10
	C. Contrast	13.88	3	4.63	1.08	0.36 n.s.	0.04

*weakly significant; **significant; ***highly significant; n.s. not significant
$p\eta^2<0.04$= negligible; $0.04\leq p\eta^2<0.25$= small; $0.25\leq p\eta^2<0.64$= moderate; $p\eta^2\geq0.64$= large

Table 4 indicates that the differences between the GRV(IES-GI) scores calculated for each group of visual tasks at all test sessions are highly significant in two cases, significant in three cases, weakly significant in three cases, and not significant in four cases. The detected differences have a generally substantive effect size, ranging between moderate ($p\eta^2$= 0.25 for the Size group in the Morning session) in one case and small ($0.04\leq p\eta^2<0.25$) in 10 cases.

These findings provide evidence to postulate that inclusive features of the task (font size, internal contrast, and character contrast) may affect the temporal variation of glare sensation.

For the *font size* of the visual task (Size), the data present the largest statistically significant and practically relevant difference in the Morning (p= 0.00***, pη^2= 0.25), followed by the Afternoon A session (p= 0.00***, pη^2= 0.21). Likewise, the effect of varying the *character contrast* of the task (C. Contrast) shows a similar trend, although with a weaker magnitude as indicated by lower effect sizes at all test sessions. Also in this case, in fact, the largest statistically and practically significant difference was detected in the Morning (p= 0.00**, pη^2= 0.16), followed by the Afternoon A (p= 0.05*, pη^2= 0.10). This leads to hypothesise that *direct* manipulation of the information of interest within the visual task (i.e., the size and the contrast of the characters) may have its largest influence on variation of glare response at earlier times of the day and that its effect decreases as the day progresses.

Conversely, modification of the *internal contrast* of the task (I. Contrast) shows an opposite tendency. In fact, the influence on temporal variation of glare sensation given by increased saturation and distortion of the background appears to strengthen with time of day, as shown by increasing levels of effect size (pη^2 increasing from 0.05-0.06 to 0.10-0.11 along the day). This seems to indicate that *indirect* manipulation of the information within the visual task (i.e., its background) may have a larger effect on glare sensation at later times of the day.

However, these results have to be treated with caution since the ANOVA did not detect statistically significant differences in the Afternoon B and Evening sessions for the C. Contrast group, and in the Morning and Afternoon A sessions for the I. Contrast tasks.

3.3 Influence of Temporal Variables on Glare Response

The GRV(IES-GI) score was used to evaluate the influence of fatigue, food intake, caffeine ingestion, mood, prior daylight exposure and sky condition on temporal variation of glare response while subjects performed visual tasks of various difficulties. To allow permutation testing, each temporal variable has been organised in ordered categories dividing the size of the visual analogue scales (10 cm) in four equal segments [5]: 0 to 2.5 cm; 2.6 to 5 cm; 5.1 to 7.5 cm; and, 7.6 to 10 cm. To facilitate the interpretation of the outcomes, each ordered category of the continuous temporal variables has been linked to a descriptor. However, only the descriptors at the ends of each visual analogue scale were presented to subjects when they provided their self-assessments of temporal variables (e.g., *not fatigued* and *very fatigued*).

The influence of task difficulty was not directly included in this analysis due to the large number of permutations that would have resulted. Rather, the 12 visual tasks were grouped into the respective times of day when the glare assessments were provided. Therefore, for each session, 240 votes of glare sensation were considered (20 subjects per 12 visual tasks).

For all temporal variables, graphical and statistical inspection revealed that some of the data were not normally distributed around the mean, thus violating one of the assumptions for a parametric test. In addition, the Levene's test of homogeneity of variance returned in most cases high statistical significance with substantive effect sizes [45]; hence, non-parametric tests were adopted. Since independent groups had different variances associated with them, the interpretation of the alternative hypothesis was supported by consideration of statistical significance as well as the mean ranks of the groups [46]. Directionality of the hypothesis was confirmed by inspection of central tendencies and graphical displays. However, if no convincing trend could be determined, a two-tailed hypothesis was applied [47].

For all variables, the Jonckheere-Terpstra test was performed to analyse the statistical significance of priori ordering effects and evaluate the magnitude and directionality of trends within the data. The Kruskal-Wallis one-way analysis of variance was used to compare the independent samples and detect statistically significant differences. Post-hoc Mann-Whitney U tests were then performed to isolate the effects detected, comparing against each other all

https://doi.org/10.1016/j.buildenv.2015.09.021
https://escholarship.org/uc/item/5g20q4dg

permutations between ordered categories of variables at each session. To counterbalance the experiment-wise error rate caused by the significance level inflating across multiple pairwise comparisons, Bonferroni corrections were applied [43]. The emphasis of the inferential analysis was placed on the effect size; this was calculated by the Pearson's coefficient r, and was considered small, moderate, or large, respectively for $r \geq 0.20$, 0.50, and 0.80 [40].

For each temporal variable, the sections below provide the graphical displays of distributions, the outcomes of the Jonckheere-Terpstra and Kruskal-Wallis tests, and the results of the Mann-Whitney U pairwise comparisons. For each contrast, Tables 5-10 illustrate the sample size of independent groups (N), the median (M_{dn}) and inter-quartile range (IQR) of the GRV(IES-GI) scores, the median difference (ΔM_{dn}) and its statistical significance (NHST, with Bonferroni-corrected *p*-value calculated with a two-tailed test), the mean ranks of the groups, the U-value test statistic, and the effect size (r). In interpreting the outcomes of Mann-Whitney U tests, effect sizes have to be considered using their absolute value (this is due the test statistic always being based upon the sample with the lowest rank sum) [48].

Fatigue. Figure 7 presents the boxplots of the data related to the four ordered categories in which participants were distributed based on their reported levels of *fatigue*. As with all other temporal variables, the figure plots the GRV(IES-GI) calculated on the individual glare assessments provided by subjects at the four test sessions.

Inspection of the graphical displays at each test session led to hypothesise a prevailing direct relationship between glare response and fatigue, as suggested by higher GRV(IES-GI) often corresponding to increasing levels of this temporal variable.

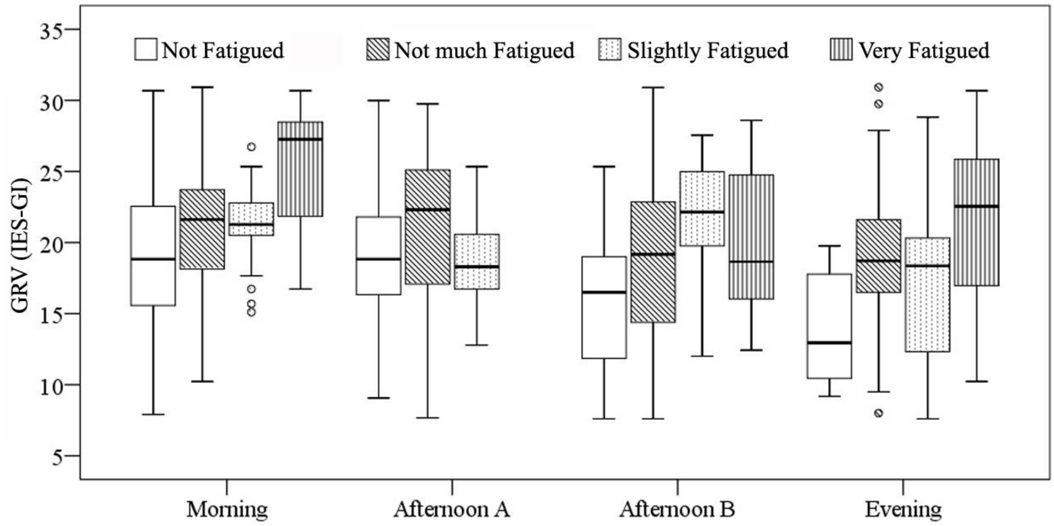


Figure 7. Boxplots of GRV(IES-GI) for *Fatigue*

The one-tailed Jonckheere-Terpstra tests showed evidence of two highly significant and practically relevant differences at the Morning (J-T= 6639.50, p= 0.00***, r= 0.27) and Afternoon B sessions (J-T= 6268.00, p= 0.00***, r= 0.30). The Kruskal-Wallis tests resulted in highly significant differences for all test sessions: Morning: $H(3)$= 20.10, p= 0.00***; Afternoon A: $H(2)$= 12.62, p= 0.00***; Afternoon B: $H(3)$= 24.82, p= 0.00***; Evening: $H(3)$= 29.21, p= 0.00***. Multiple post-hoc Mann-Whitney U tests were performed to isolate the effects detected. The results support the postulated tendency of a direct relationship between fatigue and glare response, with higher votes of glare sensation being reported by subjects as their perceived level of fatigue increases. This is corroborated by both the analysis of descriptive statistics (e.g., mostly positive median differences, ΔM_{dn}) and effect sizes (Table 5). The differences of GRV(IES-GI) detected at all sessions and levels of fatigue are statistically significant for all but five comparisons (only contrasts that are both statistically significant and practically relevant are reported in Tables 5-10). The magnitude of the effects is generally substantive, with a practically relevant influence being detected in 13 out of 21 pairwise comparisons.

https://doi.org/10.1016/j.buildenv.2015.09.021
https://escholarship.org/uc/item/5g20q4dg

Table 5. Pairwise comparisons for *Fatigue*

Session	Categories	N (x_1,x_2)	M_{dn} (IQR)	M_{dn} (IQR)	ΔM_{dn}^{NHST}	$MRank_{x1}$	$MRank_{x2}$	U	Effect Size (r)
Morn.	3 vs. 1	24, 108	21.27 (2.30)	18.83 (6.98)	2.44*	83.06	62.82	898.50	-0.20
	4 vs. 1	12, 108	27.26 (8.08)	18.83 (6.98)	8.43***	92.83	56.91	260.00	-0.31
	4 vs. 2	12, 96	27.26 (8.08)	21.62 (5.58)	5.64**	79.33	51.4	278.00	-0.28
	4 vs. 3	12, 24	27.26 (8.08)	21.27 (2.30)	5.99*	25.17	15.17	64.00	-0.45
Aft. A	2 vs. 1	132, 84	22.31 (8.20)	18.83 (5.58)	3.48**	119.17	91.73	4135.50	-0.21
Aft. B	2 vs. 1	132, 60	19.18 (8.54)	16.50 (7.52)	2.68***	106.16	75.24	2684.50	-0.26
	3 vs. 1	24, 60	22.14 (5.46)	16.50 (7.52)	5.64***	61.98	34.71	252.50	-0.51
	4 vs. 1	24, 60	18.65 (8.77)	16.50 (7.52)	2.15**	54.96	37.52	421.00	-0.32
Even.	2 vs. 1	132, 12	18.71 (5.23)	12.95 (7.73)	5.76***	76.22	31.58	301.00	-0.30
	4 vs. 1	36, 12	22.55 (9.04)	12.95 (7.73)	9.60***	28.78	11.67	62.00	-0.53
	3 vs. 2	60, 132	18.36 (8.16)	18.71 (5.23)	-0.35***	78.92	104.49	2905.00	-0.21
	4 vs. 2	36, 132	22.55 (9.04)	18.71 (5.23)	3.84*	103.28	79.38	1700.00	-0.20
	4 vs. 3	36, 60	22.55 (9.04)	18.36 (8.16)	4.19***	63.29	39.63	547.50	-0.41

Ordered Categories: Not Fatigued= 1; Not much Fatigued= 2; Slightly Fatigued= 3; Very Fatigued= 4
*weakly significant; **significant; ***highly significant; n.s. not significant
$r<0.20$= negligible; $0.20\leq r<0.50$= small; $0.50\leq r<0.80$= moderate; $r\geq 0.80$= large

Food intake. Figure 8 presents the boxplots related to the temporal variable *food intake*.

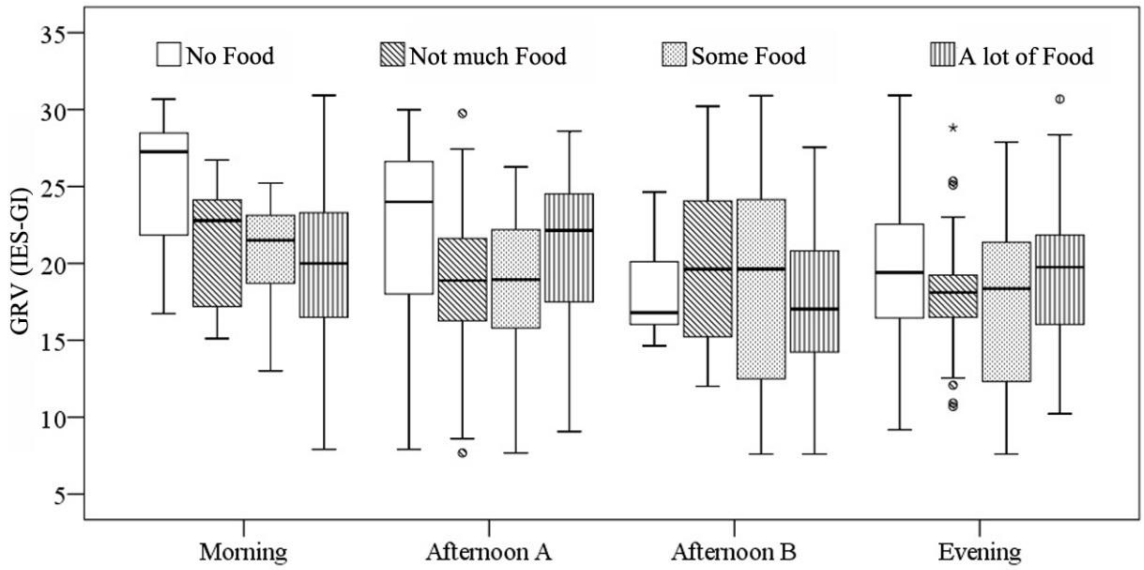


Figure 8. Boxplots of GRV(IES-GI) for *Food intake*

A non-directional alternative hypothesis was adopted for this temporal variable following analysis of graphical displays. The Jonckheere-Terpstra tests detected consistently negative test statistics and effect sizes at all sessions, with a highly significant difference in the Morning (J-T= 3676.00, p= 0.00***, r= -0.19) and a weakly significant influence at the Afternoon B session (J-T= 8609.00, p= 0.03*, r= -0.14). This leads to hypothesise that subjects who had lower food intake reported higher glare response. However, the magnitude of the effects was, for both statistically significant influences, below the threshold of practical relevance. The Kruskal-Wallis tests returned significant differences at all sessions but the Afternoon B: Morning: $H(3)$= 12.49, p= 0.01**; Afternoon A: $H(3)$= 22.54, p= 0.00***; Afternoon B: $H(3)$= 6.10, p= 0.11 n.s.; Evening: $H(3)$= 9.96, p= 0.01**. The Mann-Whitney U contrasts confirmed the postulated inverse relationship between variables in the Morning session (Table 6), as indicated by consistently negative values of median differences (ΔM_{dn}). At this time of day, the detected statistically significant and practically relevant pairwise comparisons support the hypothesis that lower levels of food intake result in higher votes of glare sensation. The tendency of an inverse relation between glare response and food intake is also apparent in two statistically significant and substantive comparisons isolated in the Afternoon A. However, this trend is not verified when analysing the other two significant permutations between ordered categories of this temporal variable at this test session.

Table 6. Pairwise comparisons for *Food intake*

Session	Categories	N (x_1,x_2)	M_{dn} (IQR)	M_{dn} (IQR)	ΔM_{dn} NHST	$MRank_{x1}$	$MRank_{x2}$	U	Effect Size (r)
Morn.	2 vs. 1	12, 12	22.87 (7.31)	27.26 (8.08)	-4.39*	8.83	16.17	28.00	-0.52
	3 vs. 1	24, 12	21.50 (4.76)	27.26 (8.08)	-5.76*	15.00	25.50	60.00	-0.47
	4 vs. 1	192, 12	19.99 (6.89)	27.26 (8.08)	-7.27**	99.18	155.67	514.00	-0.23
Aft. A	2 vs. 1	60, 60	18.88 (5.35)	24.00 (9.07)	-5.12***	49.12	71.88	2947.00	-0.33
	3 vs. 1	72, 60	18.94 (6.45)	24.00 (9.07)	-5.06***	55.30	79.94	1353.50	-0.32
	4 vs. 2	48, 60	22.41 (7.35)	18.88 (5.35)	3.53*	64.17	46.77	2806.00	-0.28
	4 vs. 3	48, 72	22.41 (7.35)	18.94 (6.45)	3.47*	71.92	52.89	1180.00	-0.27

Ordered Categories: No Food= 1; Not much Food= 2; Some Food= 3; A lot of Food= 4
*weakly significant; **significant; ***highly significant; n.s. not significant
$r<0.20$= negligible; $0.20\leq r<0.50$= small; $0.50\leq r<0.80$= moderate; $r\geq 0.80$= large

Caffeine ingestion. Figure 9 presents the boxplots in relation to *caffeine ingestion*.

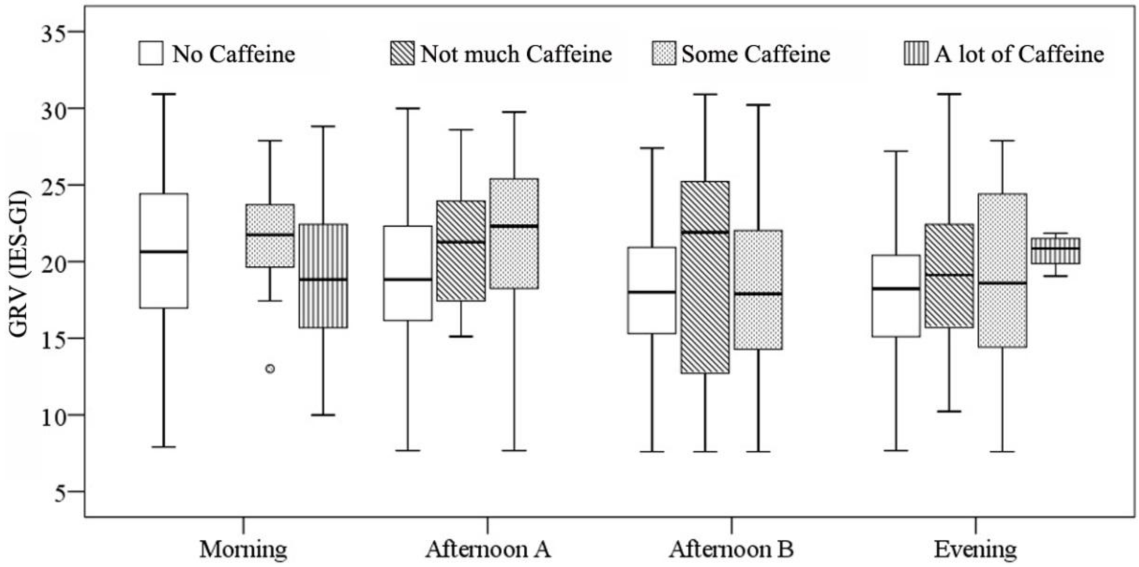


Figure 9. Boxplots of GRV(IES-GI) for *Caffeine ingestion*

Graphical inspection did not lead to identify a convincing prevailing directionality of central tendencies also due to the lack of scores in some ordered categories; hence, a two-tailed alternative hypothesis was adopted. The Jonckheere-Terpstra tests showed evidence of two statistically significant direct trends at the Afternoon A (J-T= 11101.50, p= 0.001***, r= 0.22) and Evening sessions (J-T= 10607.00, p= 0.02*, r= 0.15), suggesting that higher caffeine ingestion leads to an increase in glare response. For the later session, however, the magnitude of the influence detected was of negligible size. The results from the Kruskal-Wallis tests indicated significant differences between groups at all sessions but the Afternoon B: Morning: $H(2)$= 9.14, p= 0.01**; Afternoon A: $H(2)$= 11.68, p= 0.01**; Afternoon B: $H(2)$= 3.63, p= 0.16 n.s.; Evening: $H(3)$= 8.87, p= 0.03*. The Mann-Whitney U tests detected statistically significant and practically relevant differences in the Afternoon A and in the Evening (Table 7) that support the hypothesis of a direct relationship between the level of caffeine ingestion and the reported glare sensation. This trend is, however, not substantiated in a significant and practically relevant comparison isolated at the Morning session.

https://doi.org/10.1016/j.buildenv.2015.09.021
https://escholarship.org/uc/item/5g20q4dg

Table 7. Pairwise comparisons for *Caffeine ingestion*

Session	Categories	N (x_1,x_2)	M_{dn} (IQR)	M_{dn} (IQR)	ΔM_{dn}^{NHST}	$MRank_{x1}$	$MRank_{x2}$	U	Effect Size (r)
Morn.	4 vs. 3	48, 36	18.83 (6.86)	21.74 (4.13)	-2.91***	35.17	52.28	512.00	-0.35
Aft. A	2 vs. 1	48, 108	21.27 (6.51)	18.83 (6.21)	2.44*	92.04	72.48	1942.00	-0.20
	3 vs. 1	84, 108	22.31 (7.30)	18.83 (6.21)	3.48**	110.38	85.70	3370.00	-0.22
Even.	4 vs. 1	12, 120	20.85 (1.98)	18.23 (5.25)	2.62**	96.83	63.47	356.00	-0.25

Ordered Categories: No Caffeine= 1; Not much Caffeine= 2; Some Caffeine= 3; A lot of Caffeine= 4
*weakly significant; **significant; ***highly significant; n.s. not significant
r<0.20= negligible; 0.20≤r<0.50= small; 0.50≤r<0.80= moderate; r≥0.80= large

Mood. Figure 10 presents the boxplots of the data related to the temporal variable *mood*.

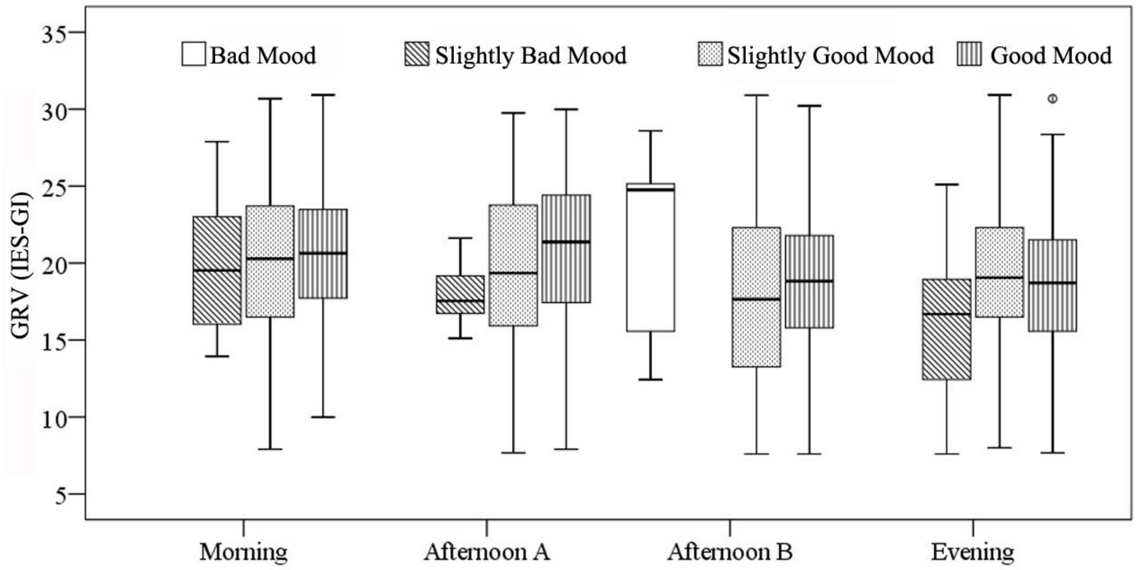


Figure 10. Boxplots of GRV(IES-GI) for *Mood*

Examination of graphical displays did not allow any consistent interpretation of a prevailing relationship between variables also due to the absence of scores in some ordered groups, thus suggesting the adoption of a two-tailed alternative hypothesis. The Jonckheere-Terpstra tests detected one statistically significant difference in the Afternoon A session (J-T= 9224.00, p= 0.01**, r= 0.16), although the magnitude of this effect was below the threshold of practical relevance (r<0.20). The Kruskal-Wallis tests resulted in significant differences only in the Afternoon A and Evening sessions: Morning: $H(2)$= 1.33, p= 0.51 n.s.; Afternoon A: $H(2)$=

https://doi.org/10.1016/j.buildenv.2015.09.021
https://escholarship.org/uc/item/5g20q4dg

6.80, p= 0.03*; Afternoon B: $H(2)$= 3.89, p= 0.14 n.s.; Evening: $H(2)$= 11.47, p= 0.00**. Post hoc Mann-Whitney U tests resulted in three statistically significant and practically relevant pairwise comparisons (Table 8), whereby the detected tendency would lead to infer that better mood corresponds to higher votes of glare sensation reported by subjects.

Table 8. Pairwise comparisons for *Mood*

Session	Categories	N (x_1,x_2)	M_{dn} (IQR)	M_{dn} (IQR)	ΔM_{dn}^{NHST}	$MRank_{x1}$	$MRank_{x2}$	U	Effect Size (r)
Aft. A	4 vs. 2	108, 12	21.38 (6.98)	17.54 (2.73)	3.84*	62.81	39.71	398.50	-0.20
Even.	3 vs. 2	84, 36	19.06 (5.69)	16.68 (6.69)	2.38**	67.22	44.82	947.50	-0.30
	4 vs. 2	120, 36	18.71 (6.22)	16.68 (6.69)	2.03**	84.23	59.40	1472.50	-0.23

Ordered Categories: Bad Mood= 1; Slightly Bad Mood= 2; Slightly Good Mood= 3; Good Mood= 4
*weakly significant; **significant; ***highly significant; n.s. not significant
r<0.20= negligible; 0.20≤r<0.50= small; 0.50≤r<0.80= moderate; r≥0.80= large

Prior Daylight Exposure. Figure 11 presents the boxplots related to consideration of the *prior daylight exposure* that subjects declared to have experienced before each test session.

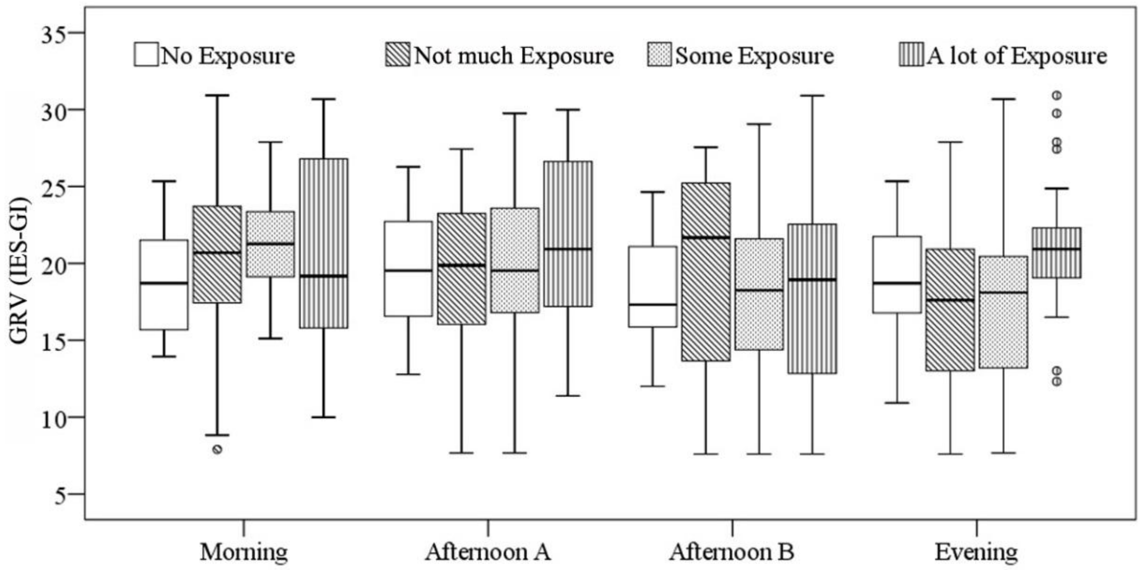


Figure 11. Boxplots of GRV(IES-GI) for *Prior daylight exposure*

The boxplots did not lead to identify any univocal trend of distributions of statistical parameters, this resulting in an alternative hypothesis supporting either a direct or inverse relationship between variables. The Jonckheere-Terpstra tests revealed one statistically significant effect in the Evening (J-T= 11902.50, p= 0.01**, r= 0.16), whose magnitude was, however, non-substantive. Also the Kruskal-Wallis tests showed significant differences between groups only at the Evening session: Morning: $H(3)$= 5.50, p= 0.14 n.s.; Afternoon A: $H(3)$= 4.55, p= 0.21 n.s.; Afternoon B: $H(3)$= 1.38, p= 0.71 n.s.; Evening: $H(3)$= 22.11, p= 0.00***. Post-hoc Mann-Whitney U tests in the Evening signalled three significant and practically relevant pairwise comparisons (Table 9). For all these contrasts, a direct trend was detected between GRV(IES-GI) scores and levels of prior exposure to daylight.

Table 9. Pairwise comparisons for *Prior daylight exposure*

Session	Categories	N (x_1,x_2)	M_{dn} (IQR)	M_{dn} (IQR)	ΔM_{dn} NHST	$MRank_{x1}$	$MRank_{x2}$	U	Effect Size (r)
	4 vs. 1	48, 36	20.92 (3.25)	18.71 (5.06)	2.21**	48.47	34.54	577.50	-0.28
Even.	4 vs. 2	48, 72	20.92 (3.25)	17.60 (8.03)	3.32***	75.61	50.42	1002.50	-0.35
	4 vs. 3	48, 84	20.92 (3.25)	18.10 (7.17)	2.82***	85.43	55.68	1107.50	-0.37

Ordered Categories: No Exposure= 1; Not much Exposure= 2; Some Exposure= 3; A lot of Exposure= 4
*weakly significant; **significant; ***highly significant; n.s. not significant
r<0.20= negligible; 0.20≤r<0.50= small; 0.50≤r<0.80= moderate; r≥0.80= large

Sky Condition. Figure 12 presents the boxplots related to consideration of the prevailing *sky condition* that participants reported before the start of each test session.

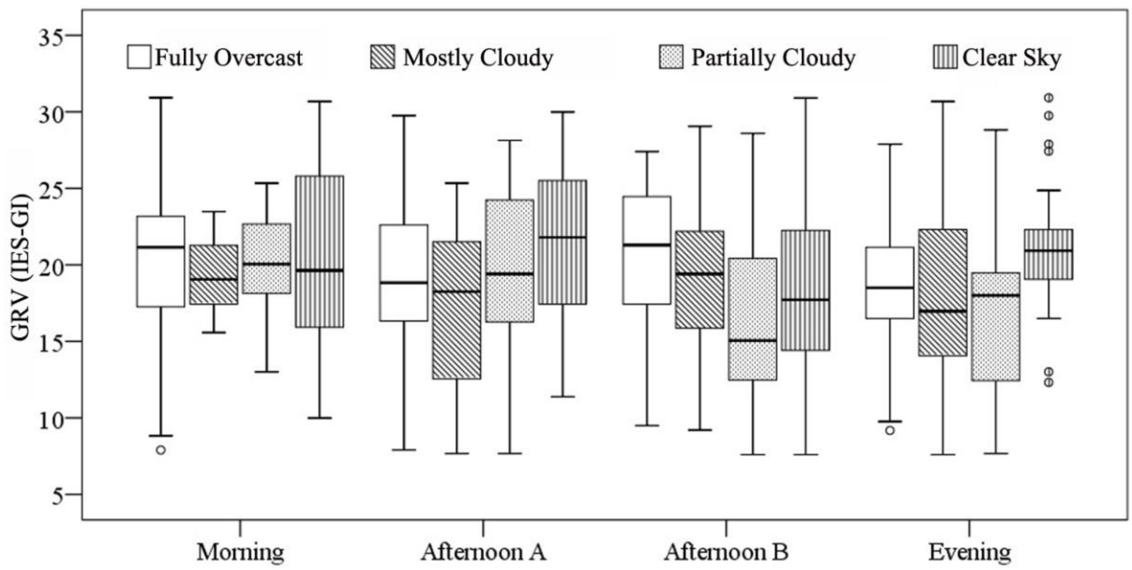


Figure 12. Boxplots of GRV(IES-GI) for *Sky condition*

The displays of distributions did not reveal any prevailing tendency between variables, hence supporting the adoption of two-tailed inferential analysis. The Jonckheere-Terpstra tests showed one statistically significant difference between categories at the Afternoon A session (J-T= 12046.00, p= 0.01**, r= 0.19), although with an effect at the borderline of practical relevance. The Kruskal-Wallis tests detected significant differences at three test sessions: Morning: $H(3)$= 0.84, p= 0.83 n.s.; Afternoon A: $H(3)$= 12.78, p= 0.01**; Afternoon B: $H(3)$= 9.92, p= 0.02*; Evening: $H(3)$= 24.15, p = 0.00***. Post-hoc Mann-Whitney U tests isolated three statistically and practically relevant differences in the Evening, leading to hypothesise a higher glare sensation being reported by subjects who experienced a clearer sky before the test sessions (Table 10). A similar tendency was found for one significant pairwise comparison in the Afternoon A. However, two statistically significant and practically relevant inverse trends were detected in the Afternoon B session, hence hindering the possibility of inferring a consistent relationship between subjective glare response and sky condition.

Table 10. Pairwise comparisons for *Sky condition*

Session	Categories	N (x_1,x_2)	M_{dn} (IQR)	M_{dn} (IQR)	ΔM_{dn}^{NHST}	$MRank_{x1}$	$MRank_{x2}$	U	Effect Size (r)
Aft. A	4 vs. 2	84, 24	21.79 (8.23)	18.24 (9.59)	3.55**	59.46	37.15	591.50	-0.30
Aft. B	3 vs. 1	48, 24	15.05 (7.99)	21.45 (7.75)	-6.40**	31.81	45.88	351.00	-0.32
	3 vs. 2	48, 48	15.05 (7.99)	18.93 (6.65)	-3.88*	42.26	54.74	852.50	-0.22
Even.	4 vs. 1	48, 72	20.92 (3.25)	18.49 (4.76)	2.43***	74.56	51.13	1053.00	-0.33
	4 vs. 2	48, 72	20.92 (3.25)	16.97 (8.52)	3.95***	72.72	52.35	1141.50	-0.29
	4 vs. 3	48, 48	20.92 (3.25)	18.00 (7.24)	2.92***	62.23	34.77	1669.00	-0.49

Ordered Categories: Fully Overcast= 1; Mostly Cloudy= 2; Partially Cloudy= 3; Clear Sky= 4
*weakly significant; **significant; ***highly significant; n.s. not significant
$r<0.20$= negligible; $0.20\leq r<0.50$= small; $0.50\leq r<0.80$= moderate; $r\geq 0.80$= large

4. Discussion

Previous laboratory experiments by the authors provided reasons to postulate an influence of time of day on glare response. This effect was found to be more substantive when considering a larger time gap between test sessions and unlikely to be affected by learning, although potentially confounded by fatigue [19]. Further research detected influences of temporal and personal factors on the variation of tolerance to source luminance as the day progresses [20]. The findings from the study presented here support the previous conclusion that the length of the time gap between test sessions shows a direct influence on the increased tolerance to source luminance along the day. Moreover, the results from the inferential statistics suggest that the effect of time of day on glare sensation may be affected by the level of visual discomfort experienced and may be masked by other factors such as the difficulty of the task. To examine more comprehensively the relations between task difficulty, temporal variation of glare response, and time of day, Figure 13 plots – for the tasks reported in Table 3 (Tasks 3 to 12) – the effect sizes (d) of the pairwise comparisons between the Morning vs. Afternoon A, Morning vs. Afternoon B, and Morning vs. Evening test sessions.

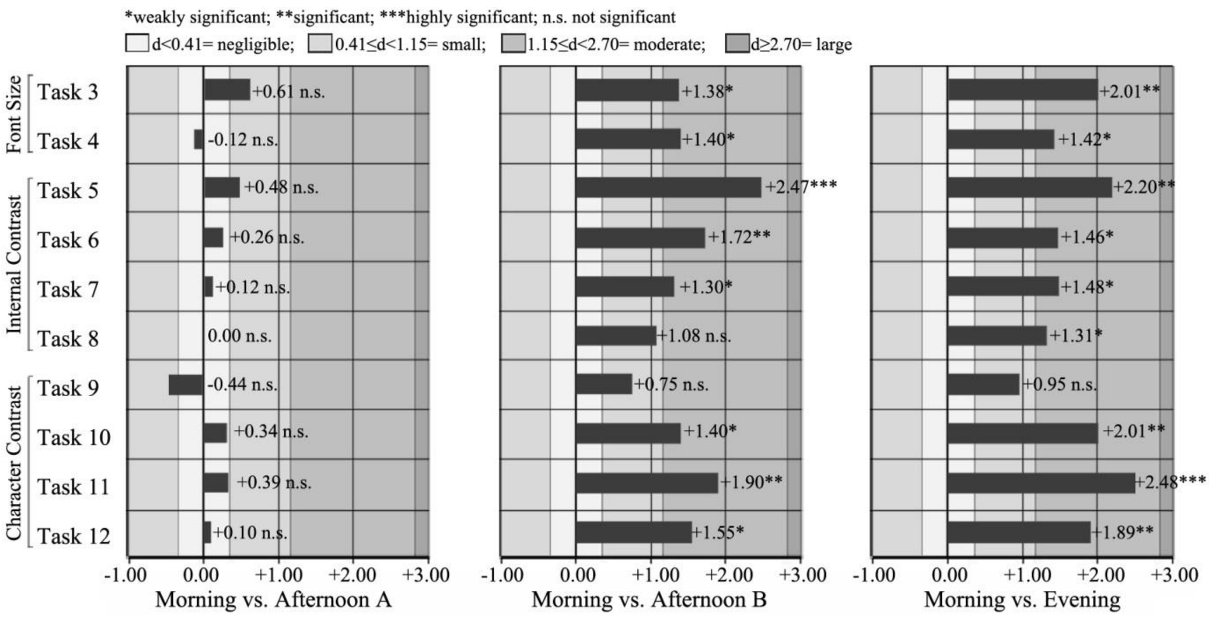


Figure 13. Comparison of effect sizes (d) for visual tasks

Although all comparisons between the Morning and Afternoon A sessions (corresponding to an interval of 3 hours between tests) were not statistically significant, the graphical displays reveal a tendency for the magnitude of the effect sizes to strengthen with the increase in the

https://doi.org/10.1016/j.buildenv.2015.09.021
https://escholarship.org/uc/item/5g20q4dg

temporal gap between sessions. The rate of variation of the effect sizes is, however, not consistent between tasks and test sessions, hence supporting the hypothesis that visual task difficulty may have an influence on glare response along the day. The only task showing a regular gradient of variation across the independent variable (time of day) is Task 3, for which a coefficient of determination for a linear fit of R^2= 0.9967 was detected. To note that the range of variation of effect sizes between tests sessions is broadly in line with what had been measured in previous experiments for the 'Just Uncomfortable' GSV criterion [19], which corresponds to the constant luminance of the glare source set for this study.

With respect to individual task manipulation, the tasks that varied in *internal contrast* (Tasks 5 to 8) show a trend, at each comparison between sessions, for the effect size to decrease when increasing the difficulty of the tasks. That is, tasks that were independently rated as more difficult correspond to lower effect sizes as the day progresses. Conversely, the tasks that varied in *character contrast* (Tasks 9 to 12) show an opposite tendency, with the effect size increasing as the tasks become more visually demanding. In this group, an exception is represented by Task 12, which was rated with the highest task difficulty of 4.00 (0.00). No clear trend could be identified for the two tasks featured in the group varying *font size* (Tasks 3 and 4), both rated with a high level of difficulty, respectively 3.00 (0.00) and 4.00 (0.00). Therefore, the graphical displays seem to support the hypothesis that direct or indirect manipulation of the information of interest within the task may lead to different effects on the variation of glare response along the day. Also, the data suggest that, for highly demanding tasks, the effect of time of day may be masked by task difficulty or other temporal influences.

With respect to the latter, Figure 14 compares the temporal variation of the effect sizes (r) extracted from the Jonckheere-Terpstra tests. These tests were performed to detect ordered differences between categories of all temporal variables across the dependent variable (glare response) and evaluate the magnitude and the sign of the trends as the day progresses [37].

For all tests, the alternative hypothesis supported either a direct or an inverse relationship (i.e., as the temporal variable under consideration intensifies, the glare response respectively increases or decreases), with the exception of fatigue for which initial graphical inspection had suggested a prevailing directional trend, hence leading to a one-way test.

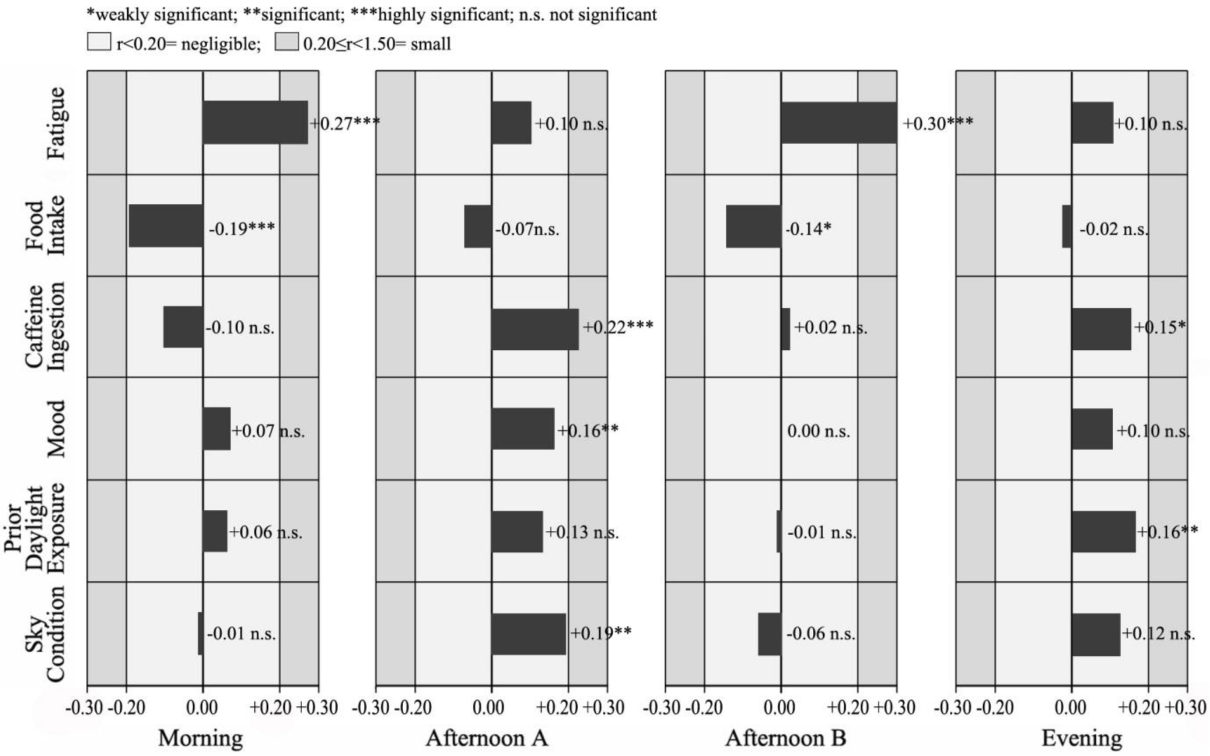


Figure 14. Comparison of effect sizes (r) for temporal variables

Consideration of self-assessed *fatigue* shows the strongest influences on variation of glare response, as indicated by the largest detected effect sizes. This is particularly evident in the Morning (r= +0.27***) and Afternoon B (r= +0.30***) sessions, with a consistent direct relationship between reported levels of fatigue and glare sensation along the day. This finding is consistent with the results of the Kruskal-Wallis and the Mann-Whitney U tests, and with postulations from earlier studies [19, 20]. However, the absence of a visual task in previous experiments had not led to identify any univocal significant and practically relevant effect.

Analysis of *food intake* suggests a consistently inverse influence of this variable on glare response as the day progresses, this being statistically significant in the Morning (r=-0.19***) and Afternoon B sessions (r= -0.14*), albeit with a marginally non-substantive effect size. Signals of this inverse relationship were also apparent in the contrasts performed on the data.

Conversely, with the exclusion of the Morning, a significant and substantive direct influence of *caffeine ingestion* on glare sensation is evident in the Afternoon A (r= +0.22***) and, to a lower extent, in the Evening (r= +0.15*) sessions. This finding is in line with other inferential testing and with earlier results that detected higher tolerance to source luminance for subjects not having ingested caffeine, and with an effect tending to decrease in the mid-afternoon [20].

A direct relationship is noticeable between ordered categories of *mood* and reported levels of glare sensation, this tendency being statistically significant in the Afternoon A session (r= +0.16**), although at a magnitude that is below the threshold of practical relevance. The pairwise comparisons detected evidence of this trend also in the Evening session.

For *prior daylight exposure* and *sky condition*, the Jonckheere-Terpstra tests did not identify univocal temporal trends between variables. However, some evidence of a statistically significant direct influence on glare response of the amount of daylight that subjects were exposed before the sessions was detected in the Evening (r= +0.16**), this being supported by the results of the Kruskal-Wallis and Mann-Whitney U tests. Conversely, a direct relationship between sky condition and glare sensation is recognisable in the Afternoon A (r= +0.19**) session, as also substantiated by the pairwise comparisons performed on the data.

In interpreting these outcomes, it is reasonable to hypothesise that some temporal variables may interact with each other along the day. In fact, the direct effect of fatigue on glare response in the Morning and Afternoon B sessions seems to complement the inverse relationship between food intake and visual sensation at these times of day. That is, subjects who declared to have ingested lower amount of food may have also given a self-assessment of greater fatigue, this resulting in a higher glare response. Also, consideration of fatigue and food intake did not lead to isolate any significant effect on glare sensation in the Afternoon A and Evening sessions. However, at these times of day, a significant and relatively substantive direct influence of caffeine ingestion could be detected on glare response. In essence, it would be plausible that higher caffeine ingestion or the lack of food intake may act as personal *stressors* that, respectively, mask or enhance the subjective self-assessment of fatigue, this resulting in a lower tolerance to source luminance and a higher evaluation of glare sensation. This is in line with research in the psychophysiological and behavioural sciences that have, among others, detected anxiety trait and long work hours to be associated with heightened central arousal and sensitisation to luminous stimuli [49].

Similarly matching patterns of variation of effect sizes can be identified when considering the temporal variables of mood and, particularly, prior daylight exposure and sky condition.

5. Conclusions, Limitations and Future Work

This study explored the relationships between task difficulty, temporal variables, and glare response as the day progresses. The main conclusions to be drawn from this analysis are:

- When performing visual tasks of various difficulties under a constant artificial lighting, the temporal gap between test sessions shows a direct influence on the increased tolerance to source luminance. This supports previous findings suggesting that the perception of glare sensation tends to decrease as the day progresses.
- The effect of time of day on glare response appears to be dependent on the level of visual discomfort experienced and to be influenced by task difficulty. Also, direct or indirect manipulation of the information of interest contained within the visual task leads to different effects on the variation of glare sensation along the day.

- A significant and substantive direct relationship was found between fatigue and glare response, particularly in the morning and during the post-lunch afternoon sessions. At these times of day, an inverse effect of food intake on visual discomfort was detected. Also, a direct influence of caffeine ingestion was found on the reported level of glare sensation, with significant effects in the early afternoon and in the evening.
- Some significant and relevant evidence was detected of a direct influence on glare response of mood, prior daylight exposure, and sky condition. However, particularly for the latter, the scatter of the data did not allow inferring a convincing relationship between these variables and subjective visual perception.
- Analysis of temporal influences on glare sensation at different times of day leads to the hypothesis that some temporal variables may interact with each other and significantly affect the variation of glare response as the day progresses.

In interpreting the influences detected and projecting them on the design of the built environment, some methodological and experimental limitations should be acknowledged.

In first instance, it should be highlighted that the between-subject tests adopted in this study were characterised by uneven sample sizes (Tables 5-10). These tests are also less likely to detect statistical significance than repeated-measures experiments, due to differences between participants that cannot be controlled [32]. In addition, all data were collected under an artificial lighting laboratory setting, and with a rather small sample size of postgraduate students, which may not be fully representative of a general population [33]. Ongoing investigations, conducted with a larger sample size and in a test room with direct access to daylight, are furthering the exploration of the influences here reported.

Nevertheless, the findings from this study provide supportive evidence that time of day, inclusive features and difficulty of the visual task, and temporal variables, play a significant and substantive role in the process of visual discomfort, and should be accounted for to interpret the scatter normally observed when regressing individual votes of glare sensation.

(Day)lighting is increasingly being considered as a research area situated at the interface between environmental and psycho-physiological human factors [50]. With this in mind, the results of this study suggest that physical and photometric parameters alone – as conventionally embedded in glare formulae and lighting standards – may not be sufficient to provide accurate measurements and predictions of visual discomfort in built spaces.

Acknowledgements

The data for this study were produced as part of the ongoing PhD research of the second author at the University of Nottingham funded by an EPSRC Doctoral Training Award.

References

[1] Boyce PR. Human factors in lighting. CRC Press, 3rd edition; 2014.

[2] Fotios S. Research Note: Uncertainty in subjective evaluation of discomfort glare. Lighting Research & Technology 2015; 47:379-383.

[3] Sivak M, Flannagan M, Ensing M, Simmons CJ. Discomfort glare is task dependent. The University of Michigan Report UMTRI-89-27; 1989.

[4] Osterhaus W. Discomfort glare assessment and prevention for daylight applications in office environments. Solar Energy 2005; 79:140-158.

[5] Wienold J. Daylight glare in offices. PhD thesis, Fraunhofer ISE; 2009.

[6] Tregenza P, Loe D. The design of lighting. Routledge, 2nd edition; 2014.

[7] Hopkinson RG. Glare from daylighting in buildings. Applied Ergonomics 1972; 3(4):206-15.

[8] Ostberg O, Stone PT, Benson RA. Free magnitude estimation of discomfort glare and working task difficulty. Göteborg Psychological Reports 1975; 15:1-18.

[9] Gunnarson E, Ostberg O. The physical and mental working environment in a terminal-based data system. Swedish National Board of Occupational Safety and Health Department of Occupational Medicine Report 35; 1977.

[10] Flannagan M, Wintraub D, Sivak M. Context effects on discomfort glare: task and stimulus factors. The University of Michigan Report UMTRI-90-35; 1990.

[11] Dugas Garcia K, Wierwille W. Effect of glare on performance of a VDT reading-comprehension task. Human Factors 1985; 27(2):163-173.

[12] Osterhaus W, Bailey I. Large area glare sources and their effect on discomfort and visual performance at computer workstations. Proceedings of IEEE Industry Applications Society Annual Meeting, Houston; 1992.

[13] Lynes JA. Discomfort glare and visual distraction. Lighting Research & Technology 1997; 9(1):51-52.

[14] Boyce PR. Current knowledge of visual performance. Lighting Research & Technology 1973; 5(4):204-12.

[15] Rodriguez R, Pattini A. Tolerance of discomfort glare from a large area source for work on a visual display. Lighting Research & Technology 2014; 46(2):157-170.

[16] Rodriguez R, Pattini A. Effects of a large area glare source in cognitive efficiency and effectiveness in visual display terminal work. Leukos 2012; 8(4):283-299.

[17] Ko P, Mohapatra A, Bailey IL, Sheedy J, Rempel DM. Effect of font size and glare on computer tasks in young and older adults. Optometry and Vision Science 2014; 91(6):682-9.

[18] Wolska, A. Human aspects of lighting in working interiors. In Karwowski W. International Encyclopaedia of Ergonomics and Human Factors. Taylor and Francis; 2006.

[19] Kent MG, Altomonte S, Tregenza PR, Wilson R. Discomfort glare and time of day. Lighting Research & Technology 2014; doi: 1477153514547291.

[20] Kent MG, Altomonte S, Tregenza PR, Wilson R. Temporal variables and personal factors in glare sensation. Lighting Research & Technology 2015; doi: 10.1177/1477153515578310.

[21] Roenneberg T, Wirz-Justice A, Merrow M. Life between clocks: daily temporal patterns of human chronotypes. Journal of Biological Rhythms 2003; 18(1):80–90.

[22] Velds M. User acceptance studies to evaluate discomfort glare in daylit rooms. Solar Energy 2002; 73(2):95-103.

[23] Tuaycharoen N, Tregenza PR. Discomfort glare from interesting images. Lighting Research & Technology 2005; 37(4):329-341.

[24] Iwata T, Tokura M. Position index for a glare source located below the line of vision. Lighting Research & Technology 1997; 29(3):172-178.

[25] Hopkinson RG. Glare from windows. Using the glare index in daylighting design. Construction Research and Development 1971; 3:23-28.

[26] Boyce PR. Illuminance, difficulty, complexity and visual performance. Lighting Research & Technology 1974; 6:222-226.

[27] Tregenza P, Wilson M. Daylighting. Architecture and lighting design. Routledge; 2011.

[28] Clear R, Berman S. Speed, accuracy and VL. Illuminating Engineering Society 1989; 19:124-31.

[29] Rea MS. Visual performance with realistic methods of changing contrast. Illuminating Engineering Society 1981; 10:164-77.

[30] Slater AL, Perry MJ, Crisp VHC. The applicability of the CIE visual performance model to lighting design. Proceedings of the CIE 20th session, Amsterdam; 1983.

[31] Markus TA. The function of windows - A reappraisal. Building Science 1967; 2:97-121.

[32] Field A, Hole G. How to design and report experiments. Sage; 2013.

[33] Tuaycharoen N, Tregenza PR. View and discomfort glare from windows. Lighting Research & Technology 2007; 39(2):185-200.

[34] CIBSE. Code for Interior Lighting. CIBSE; 1994.

[35] Chauvel P, Collins JB, Dogniauz R. Glare from windows: current views of the problem. Lighting Research & Technology 1982; 14(1):31-46.

[36] Tokura M, Iwata T, Shukuya M. Experimental study on discomfort glare caused by windows - Part 3: Development method for evaluating discomfort glare from a large light source. Architecture, Planning and Environmental Engineering 1996; 489:17-25.

[37] Field A. Discovering statistics using IBM SPSS statistics. Sage, 4th edition; 2013.

[38] Siegel S. Nonparametric statistics. The American Statistician 1957; 11(3):13-19.

[39] Schiavon S, Altomonte S. Influence of factors unrelated to environmental quality on occupant satisfaction in LEED and non-LEED certified buildings. Building and Environment 2014; 77:148-159.

[40] Ferguson CJ. An effect size primer: a guide for clinicians and researchers. Professional Psychology: Research and Practise 2009; 40(5):532–538.

[41] Bird KD, Hadzi-Pavlovic D. Controlling the maximum familywise Type I error rate in analyses of multivariate analysis. American Psychological Association 2014; 19(2):265-280.

[42] Williams LJ, Abdi H. Fisher's Least Significant Difference (LSD) test. In Salkind Neil, editor. Encyclopaedia of Research Design. Sage; 2010.

[43] Cabin RJ, Mitchell RJ. To Bonferroni or not to Bonferroni: when and how are the questions. Bulletin of the Ecological Society of America 2000; 81(3):246-248.

[44] 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.

[45] Nordstokke DW, Zumbo BD, Cairns SL, Saklofske DH. The operating characteristics of the nonparametric Levene test for equal variances with assessment and evaluation data. Practical Assessment, Research and Evaluation 2011; 16(5):1-8.

[46] Hart A. Mann-Whitney test is not just a test of medians: differences in spread can be important. Education and Debate 2001; 323:391-393.

[47] Hauschke D, Steinijans VW. Directional decision for a two-tailed alternative. Journal of Biopharmaceutical Statistics 1996; 6:211-213.

[48] Sawilowsky SS. Mann-Whitney U Test (Wilcoxon Rank-Sum Test). In Salkind NJ and Rasmussen K. Encyclopaedia of measurement and statistics. Sage; 2007.

[49] Emdad R, Belkic K, Theorell T, Cizinsky S, Savic C, Olsson K. Psychophysiologic sensitization to headlight glare among professional drivers with and without cardiovascular disease. Occupational Health Psychology 1998; 3(2):147-160.

[50] Andersen M. Unweaving the human response in daylighting design. Building and Environment 2015; 91:101-117.