Faculty of Engineering

Department of Architecture &
Built Environment

Environment, People and Design
(ePad) Research Group

Temporal Effects in Glare Response

Michael George Kent

Thesis submitted to the University of Nottingham for the degree of Doctor of Philosophy

July 2016

Abstract

Discomfort glare is considered to be an annoyance or distraction caused by sources of non-uniform or high luminance within the field of view of an observer. There are still significant gaps in our understanding of the conditions that characterise the magnitude and occurrence of discomfort glare, this being especially evident in the presence of large sources of luminance such as windows. The large degree of scatter that is observed when subjective evaluations of glare sensation are compared against calculated glare indices suggests that discomfort glare may be dependent on other variables beyond the physical and photometric parameters that are commonly embedded in glare formulae (e.g., source luminance, source size, background luminance, and position index). There are strong reasons to believe that some of these variables might be linked to the time of day when the observer is exposed to the glare source. In response, this thesis investigated the research hypothesis that subjective glare sensation is associated with temporal variability. This hypothesis was tested in two stages. The first stage was conducted within a laboratory setting, and sought to examine temporal effects under controlled artificial lighting conditions. The collection of temporal variables and personal factors – thereby examining the scatter in glare responses across the independent variable (time of day) and isolating potential confounding variables – enabled to identify factors that could influence the subjective evaluation of glare sensation along the day. Having established the presence of a temporal effect on glare response, the influences detected were further explored within a test room with direct access to daylight, whereby temporal variables and personal factors were measured in conjunction to glare sensation for them to be statistically masked from the analysis. The results confirmed the hypothesis of an increased tolerance to glare as the day progresses. This supported the conclusion that physical and photometric parameters alone are not sufficient for a robust prediction of discomfort glare.

Acknowledgements

This PhD Thesis could not have not been achieved without the supervision and guidance of many academics, colleagues and friends who have supported me with their time and expertise throughout this study. I would like to express my most sincere gratitude to the following:

- Dr. Sergio Altomonte (*The University of Nottingham*) for first accepting me as a PhD student and for all the immense time he has spent guiding me tirelessly for the last three years;
- Dr. Robin Wilson (*The University of Nottingham*) who has continuously supervised me throughout my time in Nottingham;
- Professor Peter Tregenza (*The University of Sheffield*) for inspiring me to formulate my research question and guiding throughout the entirety of this investigation;
- Dr. Jan Wienold (*École Polytechnique Fédérale de Lausanne*) for taking the time to review through all of the work and for all the meetings and discussions we had;
- Dr. Peter Rutherford (*The University of Nottingham*) who has taken the time to review all of the work conducted over the course of this thesis;
- Professor Steve Fotios (*The University of Sheffield*) who agreed to rigorously examine this Thesis, providing challenging comments and valuable feedback.

I would also like to extend my gratitude to: Diana for all her love and support; all the participants that have kindly volunteered to take part in this investigation, without their dedication this work would not have been possible; the technicians within the *Department of Architecture & Built Environment*; Dr. Martine Knoop (*Technische Universität Berlin*) and Professor Werner Osterhaus (*Aarhus University*) for their critical feedback; to my parents and my younger sister Sophie for supporting me; the members of the *Interdisciplinary Laboratory of Performance-Integrated Design* (EPFL) for all their comments.

Table of Contents

List of Figures

List of Tables

Chapter 1

Introduction

1. Introduction

For many years, building standards and daylighting codes – either implicitly or explicitly – have defined windows as essentially serving two main purposes: the first is to transmit natural illumination to indoor occupied spaces; the second is to provide a view to the outside (Tregenza & Wilson, 2011). However, this definition only partially captures the multiple functions offered by windows (Collins, 1975; 1976). Daylight from windows, in fact, plays a much more comprehensive role, impacting at once on the *physical* senses (i.e., allowing us to see, providing us with colour perception, etc.), while also being a source of *physiological* (i.e., entraining the human circadian system, contributing to stimulate alertness, regulating the sleep-wake homeostasis, etc.) and *psychological* (i.e., providing restorative connections to the dynamic cycles of daily and seasonal variations, orientation in space and time, etc.) health and well-being (Boyce *et al.*, 2003; 2006; Boyce 2014).

An example of the important role of regular exposure to daylight is represented by the relatively recent discovery that light, other than enabling vision and spatial/colour perception, offers a potent cue (*zeitgeber*) for entraining several biological processes via the action of a distinct ganglion cell-type in the mammalian retina, the *intrinsically photosensitive retinal ganglion cell* (ipRGC) (Berson *et al.*, 2002; Brainard *et al.*, 2001). These processes have been found to be mediated primarily by melanopsin (Gooley *et al.*, 2001; 2003; Provencio *et al.*, 1998), a photopigment with maximum sensitivity at a wavelength (λ_{max}) of around 480nm (Berson, 2007). These discoveries have paved the way for a whole stream of grounded research related to the so called *non-visual* effects of light, including the triggering of *neurobehavioral* (i.e., secretion and suppression), and *circadian* (i.e., cyclic rhythms) responses (Lockley, 2009). The ipRGCs are responsible for stimulating these responses and, if not properly entrained by light exposure, many aspects of physiology and behaviour might risk becoming desynchronised from each other with detrimental consequences on health and well-being (Andersen *et al.*, 2012; Andersen, 2015). For example, if building occupants are not regularly exposed to sufficient amounts of light of the right spectrum, for an adequate period of time, and at the right time of the day, reduced physiological functions (e.g., melatonin suppression), neurobehavioral performance (e.g., difficulties concentrating) and irregular sleeping patterns may emerge (Figueiro, 2003a; 2003b).

A potential barrier that arises when designing buildings to accommodate both visual and non-visual human needs is represented by the consideration that an unrestricted use of daylight through windows might be associated with the occurrence of glare, a common source of visual discomfort (other than by concerns for unwanted solar gains).

Discomfort glare is a phenomenon that has not yet been completely characterised (Boyce, 2014; Osterhaus, 2005). There are many mechanisms that have been proposed in the literature as potential causes of discomfort glare. For example, work by Fry & King (1975) explored differences in pupil size, although Hopkinson (1956) had reported almost opposing results, demonstrating that under different photometric conditions the presence or absence of glare sensation occurred independently of pupil size. Along with these studies, visual distraction was proposed by Lynes (1977) as one of the potential causes for glare, arguing that discomfort glare may be assimilated to the definition that acousticians give of ‘noise’ as ‘unwanted sound’. Hopkinson (1956) also examined the relationship between glare and muscular tension, exploring the antagonistic actions of the sphincter and dilator muscles. The sensation of discomfort may originate when the pupil is forced to adjust to conflicting requirements within the luminous environment (i.e., when simultaneously exposed to a bright light source and low background illuminance (Murray *et al.*, 2002)). However, some scholars have speculated that there may be no general specific effect that consistently causes the occurrence of discomfort glare (Jay, 1989). Yet, despite there being no proven causes that trigger discomfort glare sensation, it has often been observed that when subjects are exposed to excessive contrasts from high luminance, or excessive quantity of light from large source areas, symptoms related to visual discomfort are reported (Stone, 2009). This discomfort can cause responses that span from mild irritation to extreme distress (Howarth *et al.*, 1993), although actual task performance (Osterhaus & Bailey, 1992; Stone & Groves, 1968) or visibility (Rea, 2000) might not be necessarily affected. In terms of adaptive responses of subjects, when discomfort glare is present within the field of view, an observer would frequently draw the blinds and significantly reduce daylight levels throughout the space (Figueiro, 2008; Osterhaus, 2005), hence greatly impacting on buildings' energy performance (e.g., due to control of artificial lighting). In essence, discomfort glare is a complex phenomenon that requires further research in order to identify the key underlying mechanisms that are associated to the subjective responses to luminous stimulation under different conditions.

1.1. Problem Identification

Current glare indices are not reliable predictors of glare sensation from large sources of luminance such as windows. If a large number of subjects were asked to report their glare sensation from different light sources characterised by similar glare indices, in fact, the results would show a large scatter (Tregenza & Wilson, 2011).

Figure 1-1 to Figure 1-5 below present the results of studies from the literature illustrating a consistent scatter in photometric and calculated glare index values corresponding to criteria of reported glare sensation. In Figure 1-1 and Figure 1-3, the findings originated from an artificial lighting stimulus while, in Figure 1-2, Figure 1-4, and Figure 1-5, the stimulus utilised was daylight from windows.

In terms of subjective assessment methods, these studies mostly used discrete categorical scales, with descriptors anchored at the polar ends of the measurement generally ranging from ‘imperceptible’ to ‘intolerable’, although with slight variations of the descriptors.

A study using daylight from windows (Velds, 2002) introduced an additional category corresponding to ‘Not perceptible’ (Figure 1-2), therefore acknowledging the occurrence of low levels of brightness that may not be perceived as a cause for visual discomfort. Conversely, Tuaycharoen & Tregenza (2007) used a continuous linear scale featuring descriptors ranging from ‘Just perceptible’ to ‘Just intolerable’ (Figure 1-4). It is important to highlight that, while this graphical display appears to show a high correlation between calculated levels of Daylight Glare Index (x-axis) and Glare Response Votes (y-axis), the predictive power associated with the glare index had been improved by controlling for the effects of luminance uniformity (Tuaycharoen, 2006).

The selection of inferential statistical analysis methods varied from study to study, including: linear regression (Figure 1-1); multiple regression (Figure 1-2); matched-pairs t-test (Figure 1-3); Analysis of Covariance (ANCOVA) (Figure 1-4); and, Analysis of Variance (ANOVA) (Figure 1-5).

This demonstrates that the statistical analysis method is heavily dependent on the measurement criteria related to the experimental procedure adopted for each study (Field & Hole, 2013). None of these studies, however, provided a consistent measure (i.e., an indicator or an index) that could account for the amount of spread (scatter) associated with sources of individual differences related to the reported level of glare sensation.

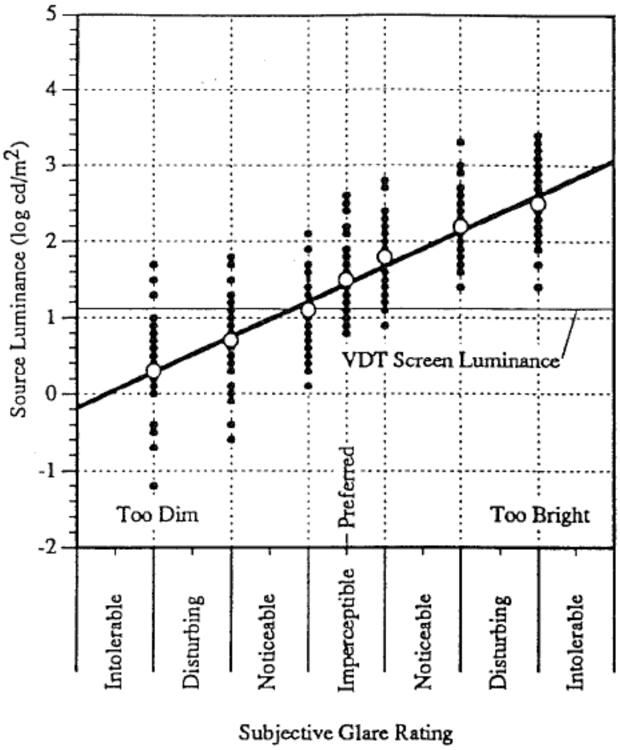


Figure 1-1 Subjective glare criteria (x-axis) against source luminance (y-axis) of a large artificial light source surrounding a Visual Display Terminal (VDT) screen. The mean (indicated by white points) and individual ratings (indicated by black points) correspond to the luminance level at which subjects (N= 26) reported the various criteria of glare sensation (Osterhaus & Bailey, 1992)

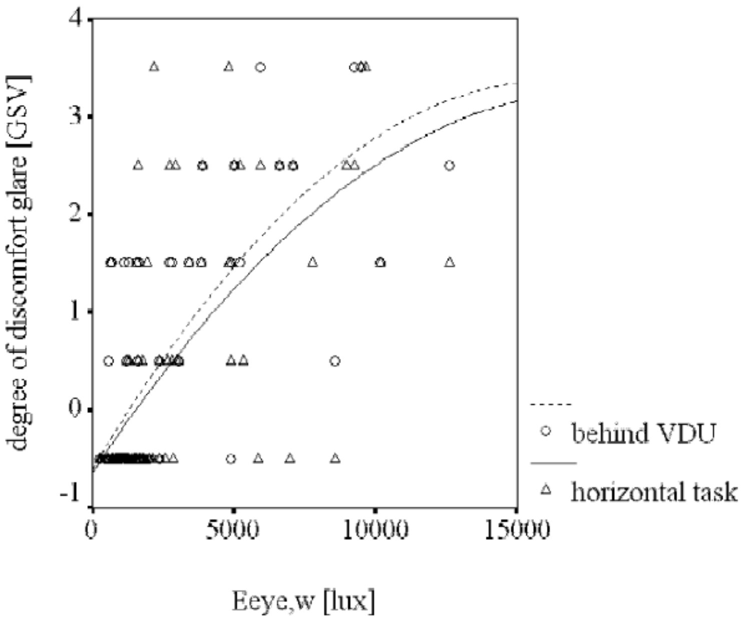


Figure 1-2 Vertical illuminance at the eye (x-axis) against reported criteria of glare sensation (y-axis). Circles correspond to assessments given when subjects (N= 23) were working in front of a VDT, and triangles indicate reported glare sensation when subjects focused on a horizontal (paper-based) task (Velds, 2002)

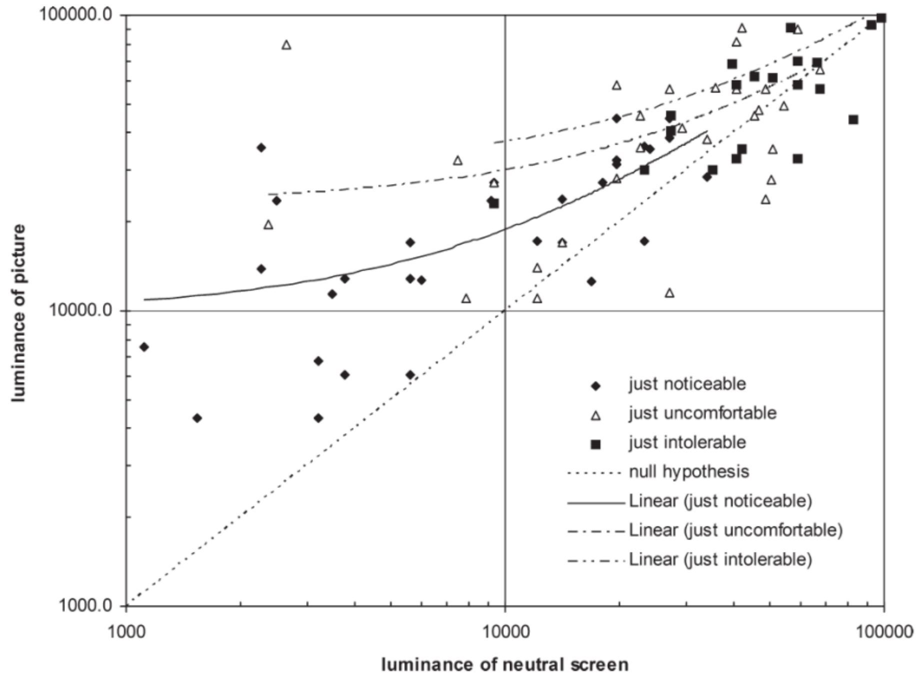


Figure 1-3 Multiple regression comparing the luminance of a ‘neutral’ plain image (x-axis) against the luminance of a ‘picture’ (y-axis), when subjects (N= 8) reported various criteria of glare sensation (Tuaycharoen & Tregenza, 2005)

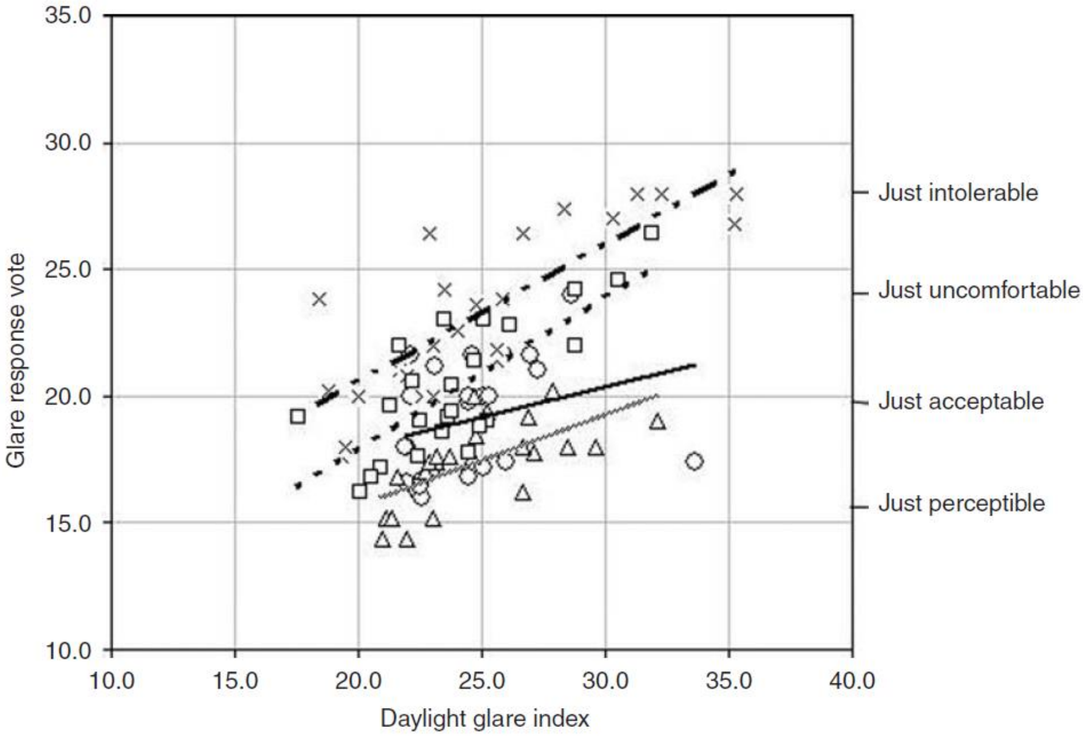


Figure 1-4 Multiple regression comparing calculated levels of Daylight Glare Index (x-axis) against subjective Glare Response Votes from subjects (N= 72) (y-axis). Circles correspond to no-view conditions (diffusive screen), triangles correspond to a low-rated view (concrete wall), and crosses correspond to a high-rated view (ground, trees and sky) (Tuaycharoen & Tregenza, 2007)

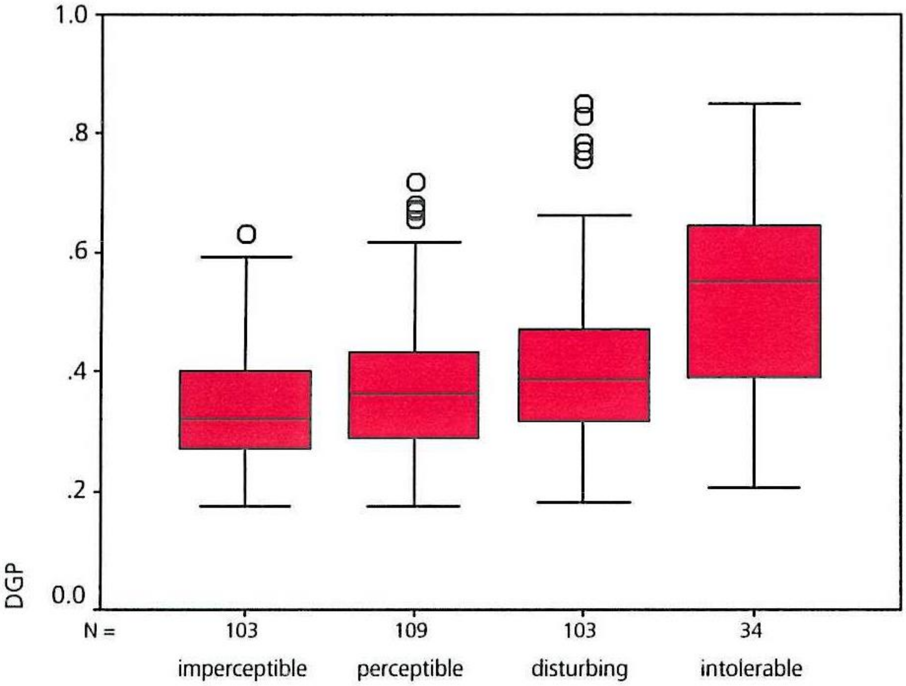


Figure 1-5 Boxplots and corresponding statistical parameters (minimum, interquartile ranges, median, maximum and outliners (circles)) used to compare the criteria of glare sensation reported by subjects (N= 76) (x-axis) against calculated values of Daylight Glare Probability (y-axis) (Wienold, 2009)

In this context, effect sizes – although these are rarely reported within research (Fritz *et al.*, 2012) – could offer a systematic method for identifying and comparing the amount of scatter within different studies regardless of their sample size, differences in rating scales, and the statistical significance (*p*-value) level used to measure the independent variable(s) (Field, 2011). The effect size is an inferential statistical parameter that can be derived from various types of statistical analysis, providing a standardised (i.e., unit- and scale-free) measure of the magnitude of the difference or strength of association across independent variables (Cohen *et al.*, 2003). Furthermore, effect sizes from similar studies can also be combined, leading to more robust calculations associated with statistical inferences (i.e., meta-analysis), since larger studies provide better estimates of population parameters (Field, 2016).

The scatter detected in the results of the studies from the literature presented in Figures 1-1 to 1-5 suggests that there must be other variables than those included in glare formulae causing variability in the level of glare sensation reported by subjects. However, despite the poor association generally detected between subjective glare responses and calculated values of discomfort glare (Stone & Harker, 1973; Hopkinson, 1971; 1972, Wienold & Christoffersen, 2006, Velds, 1999; 2002), all current methods for predicting perceived glare sensation from windows, and from artificial lighting sources, rely solely on a combination of physical and photometric parameters.

To more accurately quantify glare sensation, there is therefore a need to move beyond the four parameters commonly found in glare indices: luminance of the source, the size of the source (expressed in terms of the visual solid angle as well as shape), the general background luminance, and the angle between the direction of view and the direction of the glare source (Einhorn, 1961; Hopkinson & Collins, 1970).

1.2. Research Hypothesis and Aim

As mentioned in the Introduction, recent research has revealed that light, other than enabling visual performance, triggers several *non-visual* processes at different times of the day. Studies by Cajochen *et al.* (2000) and Zeitzer *et al.* (2000; 2005) have demonstrated that lower illuminance at the eye is required during biological night-time (00:00-05:00) rather than day-time (12:00-17:00) to stimulate non-visual responses (90-180 lux instead of 1,000 lux) (Phipps-Nelson *et al.*, 2000; Smolders *et al.*, 2012), although Cajochen (2007) argued that a complete phase response curve – including information related to both daytime and night-time effects – is needed before concluding that the alerting action of light is dependent on time of day. Comparison of bright light exposure between biological daytime (12:00-16:00) and night-time (00:00-04:00) have also been associated with time-dependency effects on heart rate, secretion of cortisol (stress hormone), core body temperature, fatigue and sleepiness (Daurat *et al.*, 1996; 2000; Ruger *et al.*, 2006).

In essence, a review of the literature suggests that the system responsible for biological stimulation presents significant changes in response to light over the time of day. Although there is currently insufficient evidence to demonstrate a link between the visual and the non-visual photoreceptive systems, it can be hypothesised that the mechanisms governing visual discomfort may exhibit similar temporal variation to those triggering non-visual responses.

The hypothesis that this investigation seeks to address is, therefore, that omitted consideration of temporal effects in the evaluation of glare sensation may (in part) explain the poor correlation that currently exists between reported and predicted levels of discomfort glare.

In response to this hypothesis, this study has been designed with the following **aim**:

Investigate whether temporal effects could be detected on the level of glare sensation reported by test subjects in the presence of daylight from a window.

However, designing an experiment to test temporal variation in glare response from a window with direct access with daylight might be problematic (Boyce, 2015), since it requires consideration of a number of uncontrollable and unknown variables (i.e., daylight fluctuates continuously within a short period of time).

Therefore, to answer the research hypothesis proposed in this investigation, a systematic experimental design approach was embraced, based on two stages of analysis.

The first stage was conducted within a laboratory setting, and sought to examine the effect of time of day on subjective glare response under highly controlled artificial lighting conditions. It was expected, in fact, that under tightly controlled conditions the effect of experimental interest would be much more sensitive and, hence, easier to isolate. Given the anticipated degree of scatter in the results, the four variables found in most glare indices needed to be controlled. Then, all other parameters that – from the literature – have been suggested to influence perceived levels of discomfort glare, but are not included within conventional glare formulae, needed to be isolated. Along with this, various extraneous variables needed to be 'masked' from the experiment (i.e., temperature, humidity, and sound).

The results from the preliminary laboratory studies permitted the identification of potential confounding variables – which also vary with the time of the day – that informed the design of a follow-on experiment where these could be controlled and quantified. The second stage of analysis took place in a test room with direct access to daylight, hence affording less controlled conditions over the variables used in estimating glare sensation. In linking these two experiments, it was assumed that similar temporal relationships could be detected for a small artificial lighting source and for daylight from a window.

1.3. Thesis Outline

The outline below provides an illustration of the various steps that have been undertaken to design and conduct the experiments, to analyse the data collected, to inform the inferences, made, and to lead to the conclusions of this work. Consistent with the study conducted by Tuaycharoen (2006), which systematically investigated the effects of different types of view on glare sensation, an extensive methodological section has been included in the description of each experiment described in Chapter 3 ‘Laboratory Studies’ and Chapter 4 ‘Test Room Studies’. This editorial choice was aimed at purposes of clarity and repeatability of the studies described, so that all the various tests conducted might be – if required by other researchers – extrapolated from the context of this work.

This Thesis follows the structure outlined below.

Chapter 1 Introduction: The introduction to this Thesis presents a brief synthetic overview of issues in lighting design surrounding the visual and non-visual systems, and illustrates the research hypothesis.

Chapter 2 Review of the Literature: Presents a comprehensive review of the literature, including aspects of discomfort glare and methods for evaluating glare sensation from different combinations of physical and photometric conditions, and from small and large sources. A detailed review of studies on the non-visual (non-image forming) effects of light is provided, documenting the research that was used to formulate the initial research hypothesis.

Chapter 3 Laboratory Studies: The experiments conducted under controlled laboratory conditions are reported in this Chapter, initially examining the research question in response to a visual stimulus that was provided by a small diffusive screen lit by artificial lighting. Personal factors and temporal variables were explored to determine whether they influence perceived visual discomfort as the day progresses. A follow-on study was designed to establish whether the detected temporal variation in reported glare response was protected from confounding variables (e.g., learning and experience). Finally, the effect of visual task difficulty (i.e., the ability of an observer to extract information from a visual stimulus) and the influences of temporal variables on perceived glare sensation were explored.

Chapter 4 Test Room Studies: After the initial controlled laboratory experiments, further studies in a test room with direct access to daylight were conducted. The test room studies aimed at isolating an effect of time of day on reported levels of glare sensation, while controlling for the effects of key temporal variables identified under controlled settings. Finally, chronotype (a personal factor) and the importance of the view to the outside (self-reported by subjects) were analysed to investigate the potential causes of scatter in the data that could not be experimentally controlled.

Chapter 5 Conclusions: The analysis of the data collected from Chapter 3 and Chapter 4 is discussed in this final chapter, conclusions are drawn, and limitations of the study are identified. The Chapter closes by proposing areas of future research that flow from this study.

Chapter 2

Review of the Literature

2. Review of the Literature

2.1. Introduction

There are strong reasons to believe that there could be other variables that influence glare sensation beyond the four that are typically embedded in the calculation of glare indices. These variables might drive personal differences in the perceived level of visual discomfort experienced within illuminated environments (Tregenza & Wilson, 2011).

The occurrence of glare has been associated to extreme non-uniformity, or very high luminance, within the field of view (Boyce, 2015). Vos (1999) suggested that there may be up to eight different forms of glare, the most common being *disability* and *discomfort* glare (Boyce, 2015). Osterhaus (2005) argued that the difference between these two types of glare can be better explained using the German-derived terms of *physiological* (referring to disability glare) and *psychological glare* (referring to discomfort glare). Disability glare has been linked to the scattering of light that fills the eye, forming a luminous veil over the retinal image of adjacent parts of the visual scene (Vos, 2002). When disability glare is present, the observer will immediately notice a reduction in their ability to see and perform visual tasks (Osterhaus, 2005). Within the literature, it has been acknowledged that disability glare is well understood. Identification of plausible underlying psychophysical mechanisms has led to successful methods of quantifying disability glare in terms of the reduction of contrast associated with the veiling phenomenon (Boyce, 2015; Chauvel *et al.*, 1982). Conversely, discomfort glare has not yet been comprehensively characterised, and there is a widely held view that current glare index measures – published in the scientific literature and in international standards – do not provide reliable predictors of the perceived discomfort reported by subjects (Tregenza & Wilson, 2011; Wienold & Christoffersen, 2006).

2.2. Evaluating Discomfort Glare from Small Sources

Discomfort glare experienced in indoor environments is frequently caused by small lighting sources of very high luminance – natural and artificial, including incandescent lamps (Paul, 1999) – and has been the focus of research for almost a century (Boyce, 2015).

A review of the literature reveals that Luckiesh & Holladay (1925) were among the first to apply psychological appraisal methods to study the sensation of discomfort glare. From a laboratory experiment, they proposed a scale of comfort/discomfort (or degrees of sensation) ranging from *scarcely noticeable* to *painful sensation* generated by the glare source. Their work investigated the effects of tungsten lamps of various sizes on perceived levels of discomfort glare at various positions in the field of view. Starting at an angular displacement of 5^o above the line of sight (i.e., normal to the position of the observers' eyes), the lamps were raised at 5^o intervals until an angular displacement of 30^o was achieved. At each interval, the subject was instructed to provide an assessment of the visual discomfort experienced. The findings revealed that the sensation of glare experienced changes according to the position of the source in the visual field. Luckiesh & Guth (1949) later used data collected from these early studies to develop the *Guth's position index*. This was described as the change in discomfort glare experienced relative to the azimuth and elevation of the source from the position of the observer (Wienold & Christoffersen, 2006). In following studies, this was introduced in the formulation of glare indices proposed by other researchers.

2.2.1. British Glare Index

While working at the Building Research Station (BRS), Peterbridge & Hopkinson (1950) developed the *British Glare Index* (BGI). Their study was conducted under controlled experimental conditions and was based on a number of assumptions around the factors that cause glare sensation (i.e., changes in source luminance, source area, angular displacement between glare sources, direction of viewing, and additive nature of glare sources). In their experiments, a continuous source of glare was presented to test subjects, whose sensation was rated from 'Just Imperceptible' to 'Just Acceptable', 'Just Uncomfortable' and 'Just Intolerable'. As also reported in Hopkinson (1950; 1955) and Hopkinson & Collins (1970), the study argued against the appraisal methods – i.e., the comfort/discomfort scale (Luckiesh & Holladay, 1925) – that in earlier studies had attempted to measure glare sensation. Instead, the authors (Peterbridge & Hopkinson, 1950) stated that test subjects should not be requested to provide judgements of the glare stimuli in terms of numerical or categorical scales, and that ratings of perceived magnitude of glare sensation should not correspond to a limited number of related criteria. As an alternative, a multiple-criterion technique was proposed, whereby judgements of glare sensation corresponded to criteria linked to word descriptions related to the observer's own experience.

Two formulae derived from the study by Peterbridge & Hopkinson (1950), based on a single glare source (Equation 2-1) and on multiple glare sources:

$$BGI = \frac{L_s^{1.6} \cdot \omega^{0.8}}{L_b \cdot P^{1.6}} \qquad \text{Equation 2-1}$$

Where:

- L_s = Luminance of the glare source (cd/m^2);
- ω = Solid angle of the source (sr);
- L_b = Luminance of the background (cd/m^2);
- P = Guth's position index of the source, which relates to its displacement from the line of sight (-).

The ‘position factor’ P accounts for the position of the individual glare sources relative to the viewing position and the direction of viewing. This was later included in the formula described in Equation 2-2 (Hopkinson *et al.*, 1966):

$IES - GI = 10Log_{10} \cdot 0.478 \sum BGI$ Equation 2-2

This then became known as the *Illuminating Engineering Society Glare Index* (IES-GI), which was proposed from the revision of the 1955 edition of the IES Code by the Technical Committee panel (Robinson *et al.*, 1962).

Referring specifically to street lighting, Peterbridge & Hopkinson (1950) (as cited in Hopkinson, 1940) also found that the perceived level of glare was additive when considering small sources. This was demonstrated by the fact that, when various sources were present in the field of view, the surrounding brightness corresponding to a particular degree of glare was found to be equivalent to the sum of the surrounding brightnesses necessary to produce the same degree of glare when each source was presented separately. Therefore, simple arithmetic addition of the glare constants – given from a number of independent sources – gave a glare value that corresponded to the glare sensation from an array of multiple sources. This implied that the glare given from an array of sources was equal to that from a single source of equal apparent area at the centroid of the array. Hopkinson (1957) later showed that the additive nature of glare sources in the field of view was a more complex phenomenon in respect to the saturation characteristics of the adaptation mechanism. In fact, it was found that the additivity function was not only dependent on the luminance of the sources, but also on their position within the field of view relative to the fixation point.

2.2.2. Visual Comfort Probability (VCP)

The work by Luckiesh & Holladay (1925) and Holladay (1926) also paved the way for the development of the *Visual Comfort Probability* (VCP), which was proposed by Luckiesh & Guth (1949). Within a controlled laboratory study using uniform background luminance, test subjects were briefly exposed to a view under a maintained adaption luminance and were requested to describe the visual sensation experienced. The results of this study led to the subjective evaluation criterion otherwise known as the *Borderline between Comfort and Discomfort* (BCD). This is based on the premise that all brightnesses within the visual field, regardless of their magnitude (i.e. luminance), area, and position, affect visibility and comfort independently of whether the observer is aware or not of their effect (Luckiesh & Guth, 1949). However, this subjective criterion has been criticised since the BCD is effectively unable to account for large individual differences associated with glare sensation. In fact, its discrete nature forces observations into dichotomous categories. Conversely, multi-criterion scales have been much preferred in the literature as they are able to more comprehensively characterise the glare sources detected by the observer and the associated perceived magnitude of glare sensation (MacGowan & Clear, 2013).

To be noted here that, unlike the work conducted by Peterbridge & Hopkinson (1950), Luckiesh & Holladay (1925) had exposed test subjects for short periods of time to the glare source, arguing that when an observer is distracted by glare they briefly observe the source and then focus on the visual task again.

For a single glare source, the relationship between subjective glare rating and experimental work established by Luckiesh & Guth (1949) led to the formula for glare sensation (M) (Equation 2-3):

$$M = \frac{0.5 \cdot L_S \cdot Q}{F \cdot P^{0.44}} \qquad \text{Equation 2-3}$$

Where:

- L_s = Luminance of the glare source (cd/m^2);
- $Q = 20.4\omega + 1.52\omega^{0.2} - 0.075$;
- ω = Solid angle of the source (sr);
- F = Average luminance of the entire field of view including the glare source (cd/m^2);
- P = Position index of the source (-).

The effect of multiple glare sources can be derived by first calculating the *Discomfort Glare Rating* (DGR) from the individual glare sensation, and then converting this to VCP. The DGR can be found from (Equation 2-4):

$$DGR = \left(\sum_{i=1}^{n} M\right)^{a} \quad \text{Equation 2-4}$$

Where:

- M = Glare sensation given from a single glare source (Equation 2-3);
- $a = n^{-0.0914}$;
- n = the number of glare sources in the field of view.

A recommended procedure for computing VCP from DGR for interior lighting was published in the IES Lighting Handbook 1984 Reference volume (Kaufman, 1984) and makes use of (Equation 2-5):

$$VCP = \frac{100}{\sqrt{2\pi}} \int_{-\infty}^{6.374-1.3227\,ln(DGR)} e^{\frac{-1^2}{2}}\, dt \quad \text{Equation 2-5}$$

In this equation, the VCP represents the percentage of people who, under a combination of physical and photometric conditions, would be expected to find the conditions represented by the DGR to be acceptable (Boyce, 2015).

2.2.3. CIE Glare Index (CGI)

At the 1979 Kyoto CIE Session, Einhorn (1969; 1979) argued that existing methods described in international standards for calculating discomfort glare displayed a wide range of discrepancy in predictions. In response, the CIE (Commission Internationale de l'Eclairage) engaged in work that would derive a unified glare formula (Equation 2-6):

$$CGI = 8Log_{10}\, 2 \cdot \frac{\left(1+\frac{E_d}{500}\right)}{E_d + E_i} \cdot \sum_{i=1}^{n} \frac{L_s^2 \cdot \omega_s}{P^2} \qquad \text{Equation 2-6}$$

Where:

- E_d = Direct vertical illuminance at the eye from all the glare source (lux);
- E_i = Indirect illuminance at the eye from all the glare sources (lux);
- L_s = Luminance of the glare source (cd/m^2);
- ω = Solid angle of the source (sr);
- P = Guth position index of the source which relates to its displacement from the line of sight (-).

Interestingly, the adaptation term – conventionally taken as the background luminance – was replaced in this formula by E_i, representing the background illumination of the room surfaces, against the glare source under observation (E_d). The quantity $(1+E_d/500)/E_d+E_i)$ allows for co-variance (i.e., it implies that there is a direct relationship between glare and the direct illuminance at the eye) in the numerator and adaptation in the denominator.

2.2.4. Unified Glare Rating (UGR)

At the 1987 CIE Session in Venice, Sorensen (1987) proposed a *Unified Glare Rating* (UGR) index, which incorporates the Guth's position index and combines aspects of both the CGI (Einhorn) and BGI (Hopkinson) to evaluate glare sensation in an environment illuminated by artificial lighting. This was in response to increasing criticism and recognition of weaknesses and limitations in existing glare prediction methods. The CIE, in fact, sought to develop a glare index that would address known weaknesses in previous methods, while retaining their advantages (Osterhaus, 2005). Most notably, discussions were held about retaining the adaptation term defined through the background luminance (L_b) (i.e., rather than the indirect illuminance (E_i)) and omitting the direct illuminance (E_d) received at the eye, since calculation of this parameter is difficult (Einhorn, 1998). The UGR was finally approved and published in the Technical Committee Report (CIE, 1993), based on the following formula (Equation 2-7):

$$UGR = 8log_{10} \frac{0.25}{L_b} \cdot \sum_{i=1}^{n} \frac{L_s^2 \cdot \omega_2}{P^2} \qquad \text{Equation 2-7}$$

Where:

- L_b = Luminance of the background (cd/m^2);
- L_s = Luminance of the source (cd/m^2);
- ω = Solid angle of the source (sr);
- P = Guth position index of the source (-).

The correlation between subjective glare rating and UGR has been tested by various studies (e.g., Akashi *et al.*, 1996). However, according to the literature, the UGR is strictly limited to glare sources subtending a solid angle between 0.0003-0.1 steradians at the eye, and is therefore unsuitable for evaluating glare sensation from indirect lighting or windows (Boyce, 2015). Moreover, for very small sources (with a projected area below 0.005 m^2), the formula requires adjustment to avoid ratings of intolerable glare being suggested from light sources that are known to provide widely acceptable luminances (Einhorn, 1998).

2.3. Evaluating Discomfort Glare from Large Sources

Conventional glare indices for artificial lighting are designed to be used with small and uniform sources below 0.01sr (Chauvel *et al.*, 1982; Velds, 1999; 2002; Wienold, 2009a). However, it has been argued that, as the glare source grows in size (>0.01sr), the perceived glare sensation does not increase to the same extent as predicted by glare indices (Chauvel *et al.*, 1982). In fact, when occupying large parts of the visual field, the glare source raises the adaption level of the eye and reduces the contrast between source and background, hence lowering the perceived level of glare sensation. Following a symposium held at Cornell University in 1956, and work at the BRS led by Hopkinson (1963), a daylight glare formula was proposed that can be used to evaluate glare sensation from large sources (Boyce, 2015).

2.3.1. Daylight Glare Index (DGI)

The work undertaken at Cornell University and the BRS led to the development of the Cornell formula. This is a modification of the original BGI (Equation 2-1), based upon results from experiments with large glare sources. This approach was proposed as the most practical way to account for the problems of adaptation, and of position in the field of view, when dealing with large area diffused glare. This was done so that the prediction of glare given from large sources could be 'aligned' with glare from small sources (Chauvel *et al.*, 1982).

In a laboratory experiment conducted under artificial lighting conditions, a bank of closely packed fluorescent lamps, whose light was diffused by an opal plastic screen, was set as a large surface of uniform luminance. Test subjects were requested to adjust the luminance of the large area source – ranging between 3.5 and 15,000 cd/m^2 – until a certain degree of discomfort glare was reached. Judgements of glare sensation were provided by subjects using Hopkinson's multiple criterion technique after adapting to the scene, based on the following:

- Just perceptible glare (Criterion D);
- Just noticeable glare (Criterion C);
- Just uncomfortable glare (Criterion B);
- Just intolerable glare (Criterion A).

The application of the BGI (the BRS formula) was considered to be invalid to evaluate glare from large sources. The reason was related to consideration of the position of the glare source relative to the fixation point. In fact, when the glare source becomes large, the point of fixation can no longer be taken as a point-source defined by a single position in space; rather, the area of the source remote from the direction of view is perceived at a lower rating of glare than the part in direct line of sight (fixation point). In response to these concerns, a correction was applied to account for different positions in the field of view (Hopkinson & Bradley, 1960). Based on the findings, the BRS formula was modified, and the degree of glare from a large source was expressed through a *Daylight Glare Index* (DGI) (Equation 2-8):

$$DGI = 10log_{10}0.478\sum_{i=1}^{n}\left(\frac{L_s^{1.6}\cdot\Omega^{0.8}}{L_b+(0.07\omega^{0.5}L_s}\right) \quad \text{Equation 2-8}$$

Where:

- L_s = Luminance of the glare source (cd/m^2);

- Ω = Solid angular subtense of the glare source, modified for the effect of its position in the field of view by means of position index P (sr);
- L_b = Luminance of the background without the luminance of glare source (cd/m^2);
- ω = Solid angular subtense of the glare source (sr).

The Cornell formula started to be used in the 1960s for predicting glare sensation from windows. However, field studies to verify the proposed formula were not reported until the 1970s (Hopkinson, 1971; 1972). These studies were conducted in classrooms and hospital wards, and were used to validate the Cornell formula in two stages.

In the first stage, groups of observers examined a wide range of daylit situations and reported their glare sensation by the multiple criterion technique and the perceived level of acceptability in consideration of the purpose of the room. Testing was conducted under stable sky conditions and at points in which the visible sky luminance did not vary over the area of the window (i.e., the window represented a uniform source of glare). This study was subject to some criticism since calculated values of glare index such as 'Just Intolerable' were rarely, or never, reported. In the second stage, subjects were requested to make judgements as to whether they agreed with recommended values of glare index in the context of their testing environment. For example, subjects had to evaluate whether a situation was 'Just Uncomfortable' when it was represented by a calculated glare index of 22.

The validation studies provided evidence that the correlation between perceived levels of glare sensation from windows and the calculated glare value was not as strong as with artificial lighting. Table 2-1 (Hopkinson, 1971) presents a comparison of glare indices based on the mean values generated from the responses of subjects tested under various conditions, displaying calculated values of IES-GI and DGI with their corresponding criteria of perceived glare sensation.

Table 2-1 Comparison between criteria of glare sensation and glare indices (IES-GI and DGI)

Glare Criteria	IES Glare Index (GI)	Daylight Glare Index (DGI)
Just Imperceptible	10	16
	13	18
Just Acceptable	16	20
	19	22
Just Uncomfortable	22	24
	25	26
Just Intolerable	28	28

The results suggested that test subjects were more tolerant to mild levels of discomfort glare from the sky seen through windows than to glare from artificial lighting sources of comparable size. This tolerance did not, however, linearly extend to high levels of glare sensation ('Just Intolerable'). To account for this discrepancy between glare sources (i.e. natural vs. artificial), the following formula was derived to adjust the DGI values (Equation 2-9):

$$DGI = \frac{2}{3} \cdot (IES - GI + 14) \qquad \text{Equation 2-9}$$

Further work to explore the validity of the DGI has revealed several limitations, particularly in relation to the low degree of correlation that exists between predicted DGI values and subjective glare sensation votes given by observers under different luminous conditions (Hopkinson, 1971; 1972). A study by Boubekri & Boyer (1992) found that subjective evaluations of daylight glare were consistently lower than predicted calculations using the DGI formula, and speculated that observers may have been influenced by the pleasant view they were exposed to in the experiments. This relationship was later confirmed by the studies of Tuaycharoen & Tregenza (2005; 2007), who discovered that when test subjects were exposed to a view that contained information judged to be 'interesting', an increased tolerance to glare was reported. Conversely, Waters *et al.* (1995) argued that a non-uniform glare source – i.e., windows with views containing surfaces with different luminance or a sky with partial cloud cover – can cause a higher degree of glare sensation than a uniform source. Considering that the DGI was based on experiments conducted in artificial conditions with uniform sources of glare, these subsequent findings suggest that it might not be applied to non-uniform light sources.

2.3.2. Chauvel's Daylight Glare Index

A study by Chauvel *et al.* (1982) – conducted in a test room with direct access to daylight from an unobstructed sky, in the absence of direct or reflected sunlight – detected a difference between the reported level of glare sensation from a real window and the glare experienced from a large source of artificial light of the same subtended area. The authors explained this difference as being in part due to the interest in the view of the sky coming from the window, which influenced the perceived experience of visual discomfort. Additionally, Chauvel *et al.* (1982) stated that discomfort glare from a window is practically independent of the window size and its distance from the observer, but is highly dependent on the luminance of the visible portion of the sky seen through the window. A new formula was developed in consideration of these findings, as follows (Equation 2-10):

$$DGI = 10log_{10}0.478 \sum_{i=1}^{n}\left(\frac{L_s^{1.6}\cdot\Omega^{0.8}}{L_b+(0.07\omega^{0.5}L_w}\right) \qquad \text{Equation 2-10}$$

Where:

- L_s = Luminance of the visible patch of sky seen from the position of the observer (cd/m^2);
- Ω = Solid angular subtense of the glare source, modified by the effect of its position in the field of view by means of its position index P (sr);

- L_b = Average luminance background of the interior surfaces of the room (cd/m^2);
- ω = Solid angle of the glare source (sr);
- L_w = Average luminance of the window (cd/m^2).

2.3.3. Daylight Glare Probability (DGP)

The basic photometric data from which the DGI formulae were developed were obtained using simulated windows of uniform luminance. In reality (i.e., in the case of real windows associated with non-uniform luminance distributions), the calculation of discomfort glare requires recording luminance values that contain a wide range of point-to-point variations (Boyce, 2015). Wienold & Christoffersen (2005; 2006) carried out experiments in test rooms with direct access to daylight at two separate locations: the Danish Building Research Institute (SBi, Denmark), and the Fraunhofer Institute for Solar Energy Systems (ISE, Germany). In each location, two identical test rooms were utilised with independent functions. One room (reference room) was used to collect subjective evaluations of glare (i.e., test subjects were requested to provide evaluations of luminous conditions). A secondary room (equipment room) was used to make measurements of the luminous environment experienced by the subjects. Therefore, in total four test rooms were used in the collection of subjective and photometric measurements.

During each test session, subjects were required to perform a series of visual tasks (reading an article, letter searching, and typing), which were displayed within a Visual Display Unit (VDU). At each location, these visual tasks were repeated under three different window sizes (subtending angles (ω_s) of, respectively: SBi test rooms: ‘small’= 1.12sr, ‘medium’= 2.00sr, and ‘large’= 3.89sr; ISE test rooms: ‘small’= 0.96sr, ‘medium’= 2.06sr, and ‘large’= 4.21sr), with three different blind options (white venetian blinds, specular blind, and vertical foil (ISE only)), and under two different viewing positions (one at a 45° angle from the window, and one at a 90° angle to the window) (Wienold, 2009a). After the completion of each visual task, subjects were requested to provide evaluations of the luminous conditions using a four-point scale corresponding to the magnitude of glare sensation experienced: ‘Imperceptible’, ‘Noticeable’, ‘Disturbing’, and ‘Intolerable’. Within the equipment room located adjacent to the reference room, a Charged Couple Device (CCD) camera was positioned at the exact position of the subject’s head position. The camera continuously took High Dynamic Range Images (HDRI) every 20 seconds, which were then combined with the corresponding evaluations of glare sensation reported by tests subjects. At the conclusion of the test sessions, a total of 349 glare assessments were collected. Initial inspection of the results revealed a large spread within the data due to individual differences. Therefore, the authors argued that examining the probability that an observer would be *disturbed* by the glare sources would help reduce individual differences in glare perception. To this aim, the scale was reduced to two categories, whereby reported levels of glare sensation of ‘Imperceptible’ and ‘Noticeable’ were combined into a category of “Not Disturbed”, and glare sensations of ‘Disturbing’ and ‘Intolerable’ were merged in a category of “Disturbed”. Using this dichotomous method to characterise glare sensation from a daylight source, a new glare formula was proposed. The resulting index evaluated the probability that an observer would be *disturbed* by the physical and photometric values in a given field of view, based on the following (Equation 2-11):

$$DGP = 5.87 \cdot E_v + 9.18 \cdot 10^{-2} \cdot log \sum_i 1 + \left(\frac{L_{s,i}^2 \cdot \omega_{s,i}}{E_v^{1.87} \cdot P_i^2}\right) + 0.16 \qquad \text{Equation 2-11}$$

Where:

- E_v = Vertical illuminance at the eye (lux);
- L_s = Source luminance (cd/m^2);

- ω_s = Solid angle of the source (sr);
- P = Position index (-)

Within the DGP, the position index of the glare source P takes into consideration differences in discomfort glare relative to the line of sight. To remind here that, in the development of the Guth Position index by Luckiesh & Guth (1949), subjective responses of glare sensation were collected from small glare sources located *above* the line of sight (i.e., only the effect of artificial light located above the line of vision was assessed). More recent studies have revealed that observers are more sensitive to glare when the source is located *below* rather than above the line of sight (Iwata & Tokura, 1998). Interestingly, in (Equation 2-11), all sources located above the fixation point are calculated using the Guth position index. Conversely, any glare source located below the line of sight is calculated using the work by Iwata & Tokura (1997) and the expression proposed by Einhorn (1997).

The DGP formula uses the vertical illuminance at the eye (E_v) as the adaptation term. As stated by Wienold (2009a), when a glare source becomes large, the background luminance is no longer independent from the source luminance and, therefore, the adaptation level can be expressed as the average luminance (i.e., considering both, the source and its surrounding background luminance). When evaluating discomfort glare from complex fenestration systems (i.e., featuring venetian blinds), in order to reduce difficulty in estimating average luminance, the vertical illuminance received from the field of view can be taken as the adaptation term since the two parameters are heavily correlated.

Table 2-2 (Wienold, 2009a; 2009b) presents the number (N) of glare assessments collected (cases), the mean DGP values, the root mean square deviation (RMSD), the standard error, and the lower (CI_L) and upper (CI_U) 95% confidence intervals for the mean DGP values corresponding to reported levels of glare sensation given by test subjects from descriptive one-way Analysis of Variance (ANOVA) tests.

Table 2-2 Comparison between criteria of glare sensation and DGP statistical parameters

Glare Criteria	Cases (N)	Mean DGP	RMSD	Standard Error	CI_L	CI_U
Imperceptible	103	0.33	0.10	0.01	0.31	0.35
Perceptible	109	0.38	0.11	0.01	0.36	0.40
Disturbing	103	0.42	0.15	0.02	0.39	0.45
Intolerable	34	0.53	0.18	0.03	0.46	0.59

At the conclusion of their study, the authors rated the new DGP as the most reliable tool to assess glare sensation in many office situations, as it overcomes many of the limitations of its predecessors (i.e., direct and reflected sunlight, non-uniformity of the source, etc.). However, as stated by the authors, further validation is required from additional studies (Christoffersen & Wienold, 2008).

2.3.4. Summary

Section 2.2. and Section 2.3 have provided a synthetic overview of some of the existing methods to evaluate discomfort glare from small (BGI, IES-GI, VCP, CGI and UGR) and large sources (DGI, Chauvel's DGI and DGP). These indices have all been published in the scientific literature and in international standards, and are all implemented within the glare evaluation software *Evalglare* (Wienold & Christoffersen, 2006; Wienold, 2009) – this tool will be discussed in detailed and utilised in Chapter 4, 'Daylight Studies'. However, existing methods of evaluation are not limited to the above, but also include:

- The *German Glare Limiting System*, which specifies limiting luminance distribution required for a luminance of average size to achieve a specified degree of glare (Bellchambers *et al.*, 1975);
- The *new Daylight Glare Index* (DGI_N), which is a variation of the Chauvel's DGI (Nazzal, 2001; 2005; Nazzal & Chutarat, 2001) that is valid under direct sunlight. In this, it was argued that a glare source from a window covers too large of an area on the retina for it to be clearly distinguished from the background (Nazzal, 2005). Therefore, the luminance background cannot be used to accurately express the adaption term and was replaced with the immediate surrounding luminance;
- The *Predicted Glare Sensation Vote* (PGSV), which was developed by a team of Japanese researchers from a series of experiments conducted in a laboratory setting and using simulated (Iwata *et al.*, 1992a) and real windows (Iwata *et al.*, 1992b). In their studies, a modification was made to the Hopkinson multiple criterion technique (Hopkinson, 1950). In fact, the authors used a continuous scale called the Glare Sensation Vote (GSV) with discrete categories of perceived glare corresponding to adaptations of, respectively, 'Just Imperceptible', 'Just Acceptable', 'Just Uncomfortable' and 'Just Intolerable' (Iwata *et al.*, 1992a; 1992b, Tokura *et al.*, 1996). Coherent with the observations made by Hopkinson (1971; 1972) and Boubekri & Boyer (1992), Iwata *et al.* (1992b) found that the DGI consistently overestimated the glare sensation that was reported by test subjects when expressing their personal experience in the presence of a real window.

With respect to the PSGV, Iwata & Tokura (1998) then expressed several limitations that need to be taken into consideration:

- The PGSV is not independent of source size; when the size increases and coincides with the whole field of view, the PGSV becomes independent from the background;
- The PGSV does not include a position index term and only aims to evaluate glare sensation from windows located within the line of vision;
- The PGSV was developed in experiments with uniform glare source (Tokura, 1996) and, therefore, should not be used to estimate the glare sensation experienced from non-uniform sources.

In summary, most methods used to evaluate the discomfort glare experienced by an observer have been derived under artificial settings (Wienold & Christoffersen, 2006). The exceptions are Chauvel's DGI – which was derived from the BGI and DGI formula in absence of direct sunlight or reflections (Chauvel *et al.*, 1982) – and the DGP – which was developed under office-like conditions (Wienold & Christoffersen, 2006). In this context, Nazzal (2005) stated that the difference between an artificial and a real window in terms of visual content is obvious. Therefore, it is difficult to apply most existing glare evaluation methods directly to the estimation of discomfort glare from real windows. Moreover, glare prediction methods all rely solely on physical and photometric measures and, as a consequence, fail to account for the likely effects brought by psychological and personal factors (Nazzal, 2001).

From a review of the literature, it has therefore been acknowledged that glare rating methods from small sources are not designed to deal with glare sources larger than 0.01 steradians or non-uniform areas of luminance (Boubekri & Boyer, 1992; Velds, 2000; 2002; Wienold, 2009). When the glare source occupies a large part of the field of view, this raises the adaption level of the eye, reducing the contrast between source and background and lowering the glare sensation (Hopkinson & Bradley, 1960).

2.4. Characteristics of Discomfort Glare

According to the literature, discomfort glare is characterised by large individual differences (Stone & Harker, 1973). In fact, in some of the earliest studies on discomfort glare, test subjects were selectively sampled for purposes of reliability of their responses rather than for their representative characteristics (i.e., age, culture, etc.) (Hopkinson, 1963). Boyce (2004; 2015) argued that discomfort glare is also dependent on the context in which the source of visual discomfort is presented to an observer. Indeed, lighting conditions that may be considered as uncomfortable in one situation may be comfortable in another. For example, when watching a sunset, direct glare is desirable, but the same glare sensation while working on a visual display terminal (e.g., on a computer screen) may impair the visual task. In addition, the *interpretation* of glare sensation may also change according to time-based socio-cultural values (MacGowan *et al.*, 1993). In fact, as reported by MacGowan (2010), Hopkinson (1957) had observed that the Borderline between Comfort and Discomfort (BCD) (Luckiesh & Guth, 1949) fell between the intervals of the 'Just Acceptable'= 15 and 'Just Uncomfortable'= 21 ratings for small glare sources. As a consequence, a glare index of 18 was initially assigned to the BCD (Hopkinson, 1957). However, in the 1980s, it was proposed that the glare index corresponding to the point in which test subjects perceived the BCD should rather be considered as equal to 25 (using the DGI formula for large luminance sources, Equation 2-12) (MacGowan, 2010).

From the review of the literature presented in Sections 2.2 and 2.3, some of the potential variables discussed so far that may influence individual responses to glare, but that are not consistently considered within conventional glare formulae, are:

- Position of the source relative to the line of sight (Iwata & Tokura, 1997, embedded within DGP);

- The connection with the outside view (Iwata *et al.*, 1992b; Hopkinson, 1971; 1972; Boubekri & Boyer, 1992);
- Uniformity of the source (Waters *et al.*, 1995).

The following sections provide a comprehensive overview of other potential variables that, based on research studies, may influence the personal evaluation of glare sensation.

2.4.1. Age and Glare Sensation

Based on the literature, aging causes a 'yellowing' effect that determines a decreased optical transmittance of the eye (Pokorny *et al.*, 1987). In addition, as the eye grows older, the pupil gets smaller and the lens becomes thicker, hence absorbing more light and causing a significant increase of light scattering. The scattering effect produces a "luminous veil" in the retina, which reduces the contrast and sharpness of the images under observation (Fry & Alpern, 1955). Over time, this can cause increased sensitivity to disability and discomfort glare in older observers (Torrington & Tregenza, 2007). A strong direct relationship has been detected between the age of the observer and the preferred illuminance levels of visual tasks (Tregenza *et al.*, 1974). A initial study by Kuhn *et al.*, (2013), conducted in a test room, requested 49 test subjects to perform a series of visual tasks (landolt ring, typing, and concentration) under two different shading settings (white venetian blinds and fabric roller blinds). Responses were distributed over independent categories, corresponding to the participants' demographic age: '20-30', '30-40', '50-60', and '60-70' years old. A small effect of age was detected, suggesting that glare was more frequently reported by older observers. A follow-up field study – conducted with 462 subjects in 9 buildings and 16 different offices in Germany – provided more evidence that older observers reported glare more often than younger subjects.

2.4.2. Culture and Glare Sensation

Pulpitlova & Detkova (1993) detected a higher tolerance to discomfort glare in Japanese rather than in European subjects. Similar observations were made by Akashi *et al.* (1996) and Cai & Chung (2013), following studies conducted under artificial lighting conditions, suggesting a strong correlation between the Unified Glare Rating (UGR) and reported individual glare sensation, although the UGR consistently overestimated the subjective responses of Asian (Japanese and Chinese) observers. The authors concluded that a possible cause for subjective ratings being lower than the calculated UGR glare indices may be due to how multiple luminaires are combined in the summation term ($\sum$) within the glare formula. Another possibility suggested by Rowlands (1996) is that the sensitivity to discomfort glare may not be universal across different countries and cultures.

2.4.3. View Interest and Glare Sensation

Studies by Hopkinson (1970; 1972) provided reasons to believe that observers could tolerate higher levels of glare sensation from a daylight source in comparison to artificial lighting of equivalent size and brightness. From these findings, Hopkinson hypothesised that the view to the outside may play a role of mediating or enhancing factor. In this context, Tuaycharoen & Tregenza (2005) analysed, in a laboratory setting, the relationship between the content of the view and the level of visual discomfort reported by test subjects, suggesting that the information contained within an image judged as ‘interesting’ can enhance the tolerance to discomfort glare. The authors also showed that specific ‘features’ of the view can be related to decreased sensitivity to glare sensation. These ‘features’ include horizontal stratification (Markus, 1967) – i.e., views that contain the three layers of foreground (ground), landscape (buildings) and background (sky) – and the presence of natural elements (Ulrich, 1981; 1984, Tennessen & Chimprich, 1995). The influence of view interest on reported glare sensation was also detected in studies that used real windows (Tuaycharoen & Tregenza, 2007). This, however, required adjusting the effect of relative maximum luminance, i.e. the maximum luminance of the glare source against its instantaneous mean luminance. Further research by Tuaycharoen (2011) found that additional factors, extrapolated from the Kaplan & Kaplan (1995) landscape preference model (i.e., 'complexity', 'mystery', coherence' and 'legibility'), could be associated to an increased tolerance to glare sensation in both an artificial setting and in the presence of a real window.

2.4.4. Photosensitivity and Glare Sensation

Rodriquez & Pattini (2014) and Rodriquez *et al.* (2015) found that subjects who self-reported to be less sensitive to light could tolerate higher levels of calculated discomfort glare. Stone (2009) characterised differences in light sensitivity using the term 'photophobia' (i.e., a fear or dislike to light), this being an effect causing abnormal sensitivity to artificial or natural light and producing a range of symptoms from mild annoyance to pain. Although an observer may not actually experience symptoms of pain on exposure to the light source, this effect has been used to describe a condition of ocular discomfort (Evans & Digre, 2003). Authors have suggested that photophobia is a common symptom that is associated with almost all forms of migraines, even if little is yet known about its pathological and neurophysiological basis (Chronicle & Mulleners, 2002; Evans & Digre, 2003).

2.4.5. Task Difficulty and Glare Sensation

Visual discomfort can result from either a combination of physical or photometric conditions present within the field of view or from the visual task itself (Boyce, 2015). In the literature, the measureable dependencies of visual task difficulty have been often expressed by the size and contrast of the characters contained within the stimulus (Boyce, 1974; Tregenza & Loe, 2014; Tregenza & Wilson, 2011). In this context, Sivak *et al.* (1989) detected an effect of task difficulty on discomfort glare by varying the size of a gap that test subjects had to detect in a stimulus that was presented simultaneously with a luminous 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 study (Flannagan *et al.*, 1990), the gap size was varied in factorial combination with whether the gap location task was required to subjects, 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 the luminance of the stimulus influenced discomfort ratings even when the stimulus was not relevant to task performance.

2.4.6. Time of Year (Circannual)

In a post-occupancy evaluation study conducted in an office building during early morning in winter, Altomonte (2009) found that daylight exposure and contrast ratios from an East-facing window signalled the potential occurrence of visual discomfort, although the occupant reported no glare sensation and had no desire to close the venetian blinds. However, when the study was repeated at the same time in summer, regardless of the presence or absence of obvious glare sources in the field of view of the observer, the blinds were drawn by the user. An interview with the occupant revealed that, in winter, the abundant morning daylight ingress in the room was preferred to feel awake at the start of the day. Conversely, during summer, by having walked to their workplace, sufficient luminous stimulation had already occurred before entering the office. This suggests that individual glare perception and luminous preferences could greatly differ depending on the time of the year (i.e., based on a circannual rhythm) when an observer is exposed to the glare source. A similar finding was reported by Christoffersen *et al.* (2000), showing that observers had a higher acceptance of the presence of sunlight in winter rather than in summer.

2.4.7. Further Variables

Among other variables that could potentially influence individual glare responses, some studies have suggested that the spectral composition of light may also affect the perceived levels of discomfort glare. For example, research has demonstrated that, when observers were exposed to a monochromatic stimuli at ~480nm (i.e., in the blue band), increased glare sensation was reported, hence signalling a lower tolerance to source luminance. Conversely, more tolerance to discomfort glare was detected when observers were exposed to a mid-wavelength luminous source at ~577nm (i.e. in the yellow band) (Flannagan *et al.*, 1989; 1994). A study by Boyce *et al.* (2006) also postulated interactions between visual sensation, duration of exposure, individual control of lighting, surface reflectance, and task characteristics, which all might have a potential influence on visual discomfort. Finally, a potential connection between visual sensation and *non-visual* effects of light (e.g., neuro-endocrine, neuro-behavioural and neuro-physiological responses) has also been suggested, hypothesising a link between visual acceptance and mood, alertness, sleepiness, prior daylight exposure, and time of day (Borisuit *et al.*, 2015).

2.4.8. Summary

From a comprehensive review of the literature, various variables that could potentially be linked to individual differences in subjective glare response have been discussed. These variables may be in part associated to the large scatter in levels of glare sensation provided by test subjects, which has often been detected in experimental research.

With respect to the aims of this study, evidence from the literature suggests a need for controlling as many variables as possible within the proposed experiments under controlled artificial lighting conditions (stage one of this investigation), so that their influence on the measureable outcome (i.e., glare response) can be masked from the analysis. In terms of variables that cannot be controlled – or for which control is not practically feasible (e.g., sleepiness, fatigue, prior daylight exposure, etc.) – but that are postulated to have a potentially significant impact on perceived levels of glare sensation, there is a need to measure these variables in conjunction with reported levels of visual discomfort, so that their influence can be analysed independently. This will allow isolating the effects of specific variables (counterbalancing the impact of uncontrolled confounding factors), in order for their influences to be controlled when evaluating the effect of experimental of interest within a daylit setting (stage two of this investigation).

2.5. Non-Image Forming System

Other than stimulating the *visual* photoreceptive system (i.e., what we see), light reaching the retina has several *non-visual* effects on human physiology and behaviour (Boyce, 2015). This places a wide array of biological processes under direct or indirect retinal control (Lucas *et al.*, 2014). Circadian (*circa-dian*, about a day) biological rhythms have a periodicity of around 24 hours and regulate many processes such as core body temperature, hormone levels, sleep, cognitive performance, etc. (Berson, 2003). These rhythms are primarily entrained by the light received at the eye, which synchronises (*zeitgeber*) bodily functions with the light/dark cycle governed by the Earth's rotation, although non-photic cues also play a role (Berson, 2003; Hannibal, 2002).

2.5.1. Physiology

The relatively recent discovery of a distinct class of ganglion cells, *intrinsically photosensitive Retinal Ganglion Cells* (ipRGCs), in the mammalian retina has led, over the last 15-20 years, the first steps of an entirely new stream of photobiological research (Berson *et al.*, 2002). The newly detected photoreceptor, ipRGC, is not found in the same area of the retina as the rods and cones that are used by the visual photoreceptive system (Berson 2003). The visual system, in fact, presents a larger concentration of photoreceptors (cones) within the fovea (Provencio *et al.*, 2002), which is the central and most sensitive area of the inner eye. Conversely, the ipRGCs, which constitutes a tiny percentage (<1~2%) of all the retinal ganglion cells (RGCs) present in the retina, appear to be mostly located away from the fovea, in the synaptic lamina (Provencio *et al.*, 1998; 2000). This has been associated with a higher probability of photon capture (i.e., detection of light) in comparison to the visual photoreceptive system, giving it a faster reaction time (Boyce, 2015). However, the response of ipRGCs to luminous stimulation is much more sluggish than the visual system (Berson, 2003), and appears to require much higher levels of illumination to produce a sustained response (Berson, 2007).

The ipRGCs reach phototransduction by expression of an opsin-like protein called *melanopsin*, which is characterised by different properties with respect to the rod-opsin and cone-opsins found in rods and cones (Berson, 2003; Hatter *et al.*, 2002; Provencio *et al.*, 1998; 2000 Gooley *et al.*, 2001). Provencio *et al.* (1998; 2002) initially hypothesised that melanopsin was responsible for mediating the non-visual photoreceptive system, having isolated this photopigment in the outermost synaptic lamina of amphibians. Notably, this opsin appeared not to be contained within conventional retinal photoreceptors, but was restricted within specific ganglion cells that were not responsible for image formation (Gooley *et al.*, 2001; Provencio *et al.*, 2002). This is the reason why the non-visual effects mediated by ipRGCs are also often referred to as *non-image forming* (al Enezi *et al.*, 2011).

iPRGs have a dedicated neural connection to the *suprachiasmatic nuclei* (SCNs) in the hypothalamus, and respond to light by depolarising the stimuli. In fact, even under high illumination, relatively few photons are absorbed, thereby amplifying the optic signal in order to encode physiologically relevant events to light intensities (Lucas *et al.*, 2014). After light stimulates both visual and non-visual photoreceptors, signals are fed either to the primary optic tract (POT) from rods and cones, or to the SCN from the ipRGCs via the retinohypothalamic tract (RHT) (Gooley *et al.*, 2001). However, neural signals from the rods and cones have also been found to contribute to non-visual responses (Hanifin & Brainard, 2006; Lucas, 2013; Lucas *et al.*, 2014; Rea & Figueiro, 2014). The SCN functions as a *pacemaker* (or master clock) for the body, and is located above the crossing of the optic nerves (Lockley & Gooley, 2006; Roenneberg *et al.*, 2007; Berson, 2003). The SCN is responsible for synchronising the timing of several physiological events, being connected to parts of the brain (Hannibal, 2002) that, among other processes, modulate the pineal glands and adrenal cortex responsible for hormonal (e.g., melatonin and cortisol) secretion and suppression (Berson, 2007).

2.5.2. Circadian System

Central to the non-visual photoreceptive system is the circadian *biological clock*, which regulates many daily patterns throughout the body such as sleep/wake cycles and changes in behaviour (Rea *et al.*, 2002). The ipRGCs convert luminous radiation into neural signals that are carried to the SCN, which is responsible for synchronizing physiological functions to light/dark cycles (Berson, 2007; Lockley, 2009; Rea & Figueiro, 2014). The SCN presents an endogenous periodicity (*internal time*) that is slightly different from 24 hours (this varies between individuals). Therefore, in order for the circadian pacemaker to timely regulate physiological and behavioural proxies, the internal clock is highly dependent on exogenous cues (*external time*) (Duffy & Czeizler, 2009; Duffy *et al.*, 1996; Lockley, 2009). Circadian disruption can have serious consequences on health and well-being (Andersen *et al.*, 2012; Andersen, 2015). A disruption to daily patterns is characterised by *phase shifts*, which determine changes in the timing of circadian rhythms (Duffy *et al.*, 1996). One of the primary factors for the circadian pacemaker to remain synchronised to the external environment is represented by the timing of exposure to light (Lockley, 2009; Minor *et al.*, 1991). In fact, bright light exposure early in the biological night may lead to a *phase delay*, while bright light presented late in the night may lead to a *phase advance* (Boyce, 2015).

2.5.3. Melatonin Regulation

The most common biological marker to quantify any disruption to the circadian pacemaker is represented by the secretion and suppression of the pineal hormone melatonin (Rea & Figueiro, 2014). In fact, in all mammalian species, the pineal gland – receiving neural signals from the SCN (Vanecek *et al*., 1987; Figueiro & Rea, 2010) through the anatomical pathway of the brain (Arendt, 1995; Stone, 1999) – synthesizes and secretes high levels of melatonin during the night, while production ceases during the day (Brainard *et al*., 1988; Lockley, 2009; Rea *et al.*, 2002). Melatonin provides an internal representation of the environmental photoperiod (Lockley, 2009). Early studies have detected nocturnal suppression of melatonin under high levels of artificial illumination, which were not found when light levels were decreased (Lewy *et al.*, 1980). Further studies have shown that melatonin suppression depends on light intensity. In a night-time experiment (between 00:00 and 03:00), McIntyre *et al.* (1989a) used light intensities of 10 lux (controlled group), 200 lux, 400 lux, and 600 lux at the eye with independent sample groups, confirming a direct relationship between increasing light intensity and melatonin suppression. At an illuminance of 1,000 lux at the eye, nocturnal melatonin levels could be suppressed down to daytime levels (McIntyre *et al.*, 1989b). Work by Figueiro *et al.* (2011) displayed similar findings in relation to intensity and melatonin suppression from computer monitors. A study by Rea *et al.* (2001) used lighting conditions with different spectral power distributions (SPD) – one warm (yellowish) and one cold (bluish) in appearance – each delivering approximately the same photopic (cone-based) illuminance, but with different scotopic (rod-based) weighting. The findings indicated that at higher scotopic illuminance there was greater melatonin suppression, providing an early indication that the non-visual photoreceptive system may have a response that is effectively blue-shifted in terms of lighting spectral composition.

2.5.4. Neuroendocrine, Neurophysiological and Neurobehavioral Effects of Light

In addition to resetting and regulating the circadian pacemaker, light also triggers a variety of other *neuroendocrine, neurophysiological* and *neurobehavioral* responses (Hanifin & Brainard, 2006). In addition to contributing to melatonin suppression (Lockley, 2009), light exposure directly stimulates cortisol production in the morning. However, unlike melatonin, cortisol rhythms seem to be closely tied to transitions from periods of dark to light (Figueiro & Rea, 2010). Research has also investigated neurophysiological effects of light exposure (Cajochen *et al.*, 2005) – such as changes in heart rate (Scheer *et al.*, 2004) and body temperature (Scheer *et al.*, 2005) – and neurobehavioral responses – such as increase in alertness, psychomotor vigilance, lapses in attention, and changes in mood (Cajochen, 2007; Dumont & Beaulieu, 2007; Lockley *et al.*, 2006). From a review of the literature, it appears that the crucial factors to influence these non-visual effects are represented by the *timing*, *intensity*, *spectral composition*, *duration*, and *prior history* of the luminous stimulus (Lockley, 2009). These factors will be briefly discussed in the following sections.

2.5.5. Timing

According to the literature, circadian regulation is highly dependent on a range of subset factors that characterise the luminous stimulus (Berson, 2003; Hannibal, 2002; Roenneberg *et al.*, 2007), the most important being the timing of exposure (Minors *et al.*, 1991; Lockley, 2009). Precise information can be provided by a *Phase Response Curve* (PRC) that illustrates the overall relationship between phase shift magnitude and the circadian phase of the master pacemaker (Figueiro *et al.*, 2014; Khalsa *et al.*, 2003). Figure 2-1 displays data from a study using 21 test subjects, with pre- and post-stimulus constant routines, under dim (~2-7 lux) and bright light exposures, with a fixed gaze (~10,000 lux) for 6 minutes and a free gaze (~5,000-9,000 lux) for other 6 minutes, and for a total of 6.5 hours of light exposure under all conditions. The graph plots on the x-axis the circadian phases 6-18 (i.e., 06:00 to 18:00) related to the time of the day, and on the y-axis the phase shift. The filled circles (black) represent the selective data from plasma melatonin collected from test subjects, and the open circle (white) present data collected from salivary melatonin. The curve describes the relationship between a bright light exposure and a response over time, in this instance a shift in the circadian rhythm as illustrated by melatonin concentration. A shift corresponding to a positive value indicates a phase-advance, and a negative value indicates a phase-delay. In this context, it is important to highlight that exposure to light at specific times of the day may either be beneficial or detrimental, depending on the intended responses that are affected by the stimulation (e.g., bright light exposure early in the night may be beneficial for night-shift workers, but harmful to sleep) (Andersen *et al.*, 2012; St. Hilaire *et al.*, 2012).

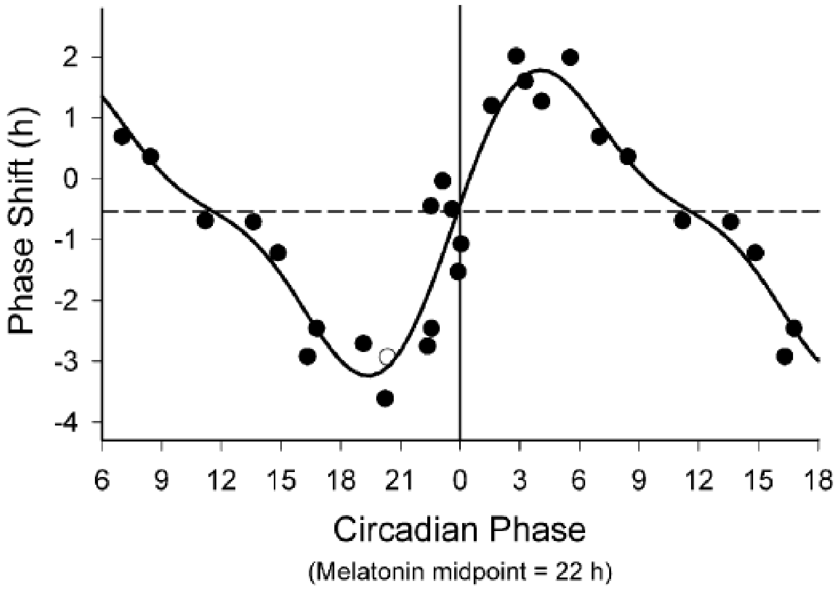


Figure 2-1 Melatonin as the circadian phase biomarker to a bright light stimulus. The time of day (circadian phase) at which plasma melatonin levels (x-axis) were collected from test subjects (N= 21) is plotted against estimates of circadian phase shifts (y-axis) resulting from 6.7h of bright light exposure (Khalsa *et al.*, 2003)

2.5.6. Intensity

The intensity of the light received at the retina plays an important role in resetting the circadian pacemaker (Lockley, 2009). Studies have shown a non-linear relationship between light intensity and its phase resetting effects (expressed by melatonin suppression), such that exposure to relatively dim conditions (~100lux (Zeitzer *et al.*, 2000); ~50-160lux (Zeitzer *et al.*, 2005)) for 6.5 hours at night could stimulate 50% (or half-saturation) of the maximum effect detected with an illuminance of equal duration up to 100 times greater (Figure 2-2).

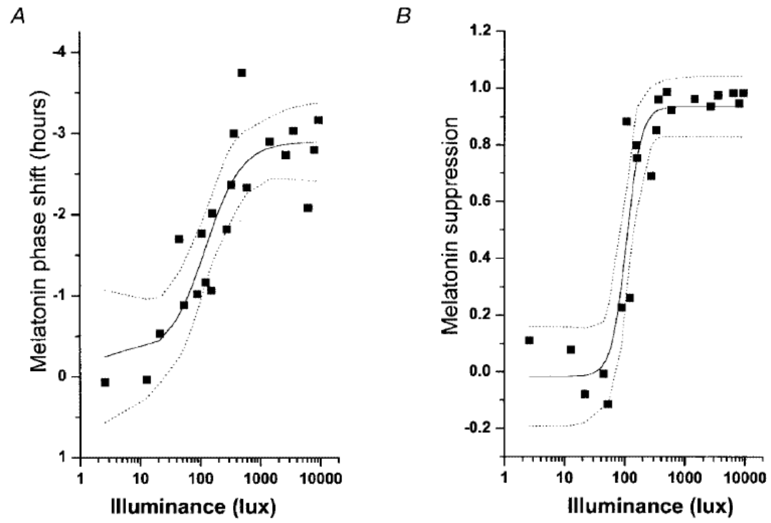


Figure 2-2 Dose-response relationship between levels of illuminance (x-axis), melatonin phase shift response (left, y-axis), and melatonin suppression (right, y-axis), measured from melatonin plasma concentration levels following 6.5h of bright light exposure on (N= 21) test subjects during early biological night (Zeitzer *et al.*, 2000)

Further studies by Cajochen *et al.* (2000), conducted during the biological night, also revealed that the intensity of the light source displayed a non-linear response relationship with heightened subjective alertness, slow eye movements (SEMs), and reduced measure electroencephalogram (EEG) theta-activity (Figure 2-3).

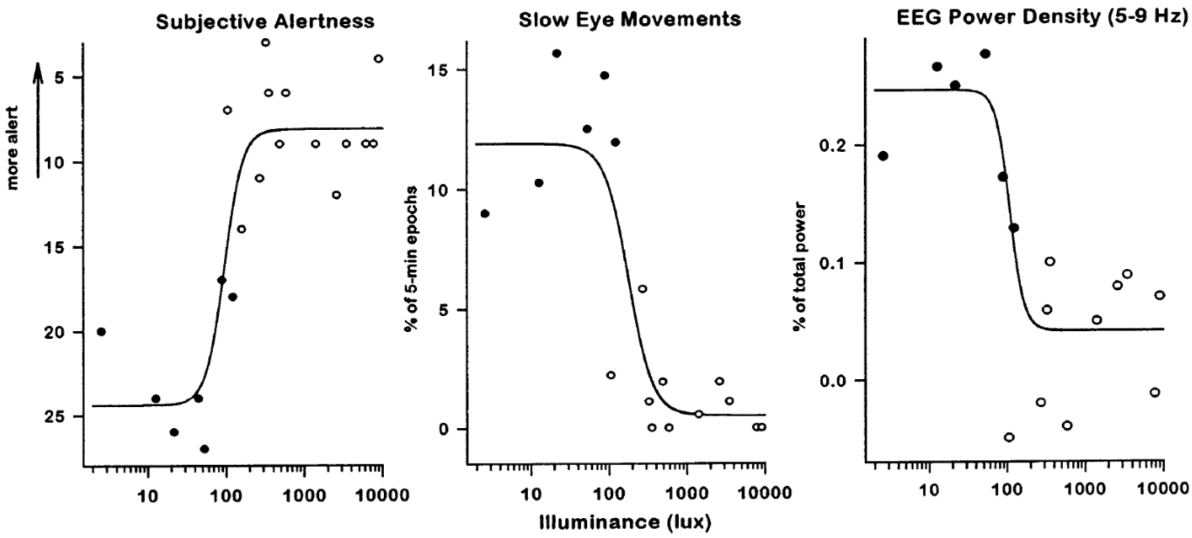


Figure 2-3 Dose-response relationship between levels of illuminance (x-axis), subjective alertness (left, y-axis), incidences of measured SEMs (middle, y-axis), and EEG theta-alpha activity (5-9Hz) (right, y-axis), following 6.5h of bright light exposure on (N= 22) test subjects during early biological night (Cajochen *et al.*, 2000)

Bright light exposure during the daytime, using a mean illuminance of 1,056 lux at the eye, has been associated with subjective sleepiness, minimised incidences of SEMs, and improved reaction times (Phipps-Nelson *et al.*, 2003). These alerting effects have been detected without any suppressive changes to melatonin levels in the body. In addition, fMRI (functional magnetic resonance imagining) studies have shown that daytime exposure to bright light can increase subjective alertness and enhance activity due to effects in the posterior thalamus of the brain (Vandewalle *et al.*, 2006). Similar findings were reported by Smolders *et al.* (2012), showing that, during daytime in a simulated office environment, illuminance levels of 1,000 lux at the eye were sufficient to reduce sleepiness, increase alertness, improve reaction times on a psychomotor vigilance task, and increase physiological arousal. These findings were also directly dependent on time of the day, since bright light exposure in the morning (09:00 or 11:00) for one hour compared to the afternoon (13:00 or 15:00) yielded more substantive effects from the luminous stimulus. Similar results of bright light exposure on subjective sleepiness, vitality, and mood were also detected by Smolders & de Kort (2014).

2.5.7. Spectral Sensitivity

Studies have demonstrated that light presented over a spectral region between 450 nm and 555 nm provides the strongest stimulation to circadian responses in mammals (Brainard & Hanifin, 2005). Brainard *et al.* (2001) and Thapan *et al.* (2001) independently developed a Spectral Response Curve (SRC) for melatonin suppression C(λ) (Figure 2-4). Brainard *et al.* (2001) formed the SRC for the non-visual proxy melatonin suppression by comparing eight different wavelengths resulting in predicted peak sensitivity at 464 nm. Thapan *et al.* (2001) used monochromatic light that differed in wavelength (424 nm, 456 nm, 472 nm, 496 nm, 520 nm and 548 nm) and irradiance (0.70-65 μW/cm^2), administered for 30 minutes via a sphere (45 cm in diameter), proposing a SRC peaking at 459 nm. The resulting action spectra for melatonin suppression C(λ) show short-wavelength maximum sensitivity, hence being very different from the known scotopic S(λ) and photopic V(λ) spectral sensitivities.

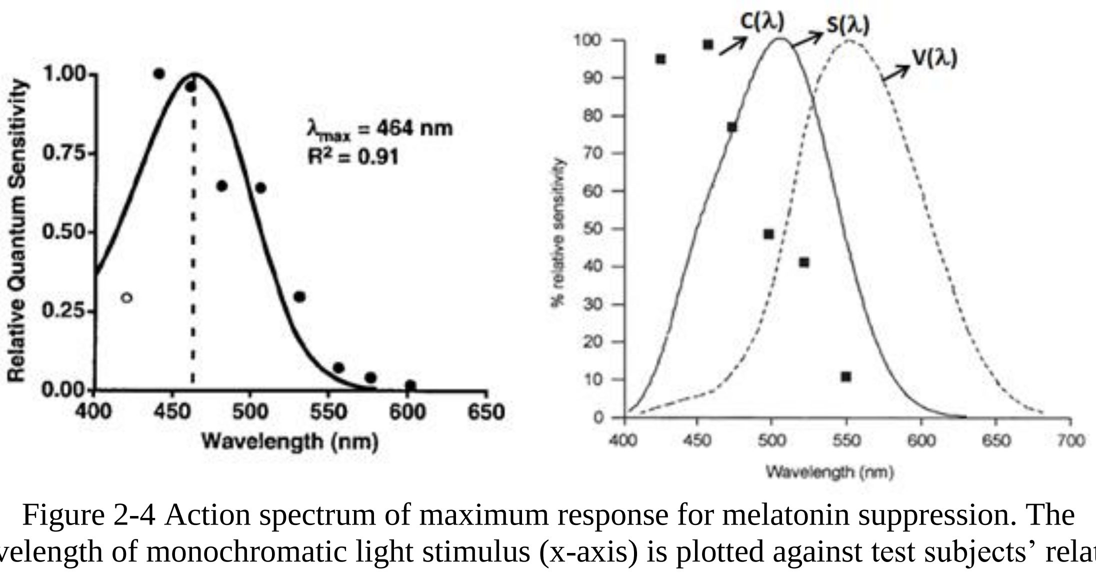


Figure 2-4 Action spectrum of maximum response for melatonin suppression. The wavelength of monochromatic light stimulus (x-axis) is plotted against test subjects' relative sensitivity (y-axis) (Brainard *et al.*, 2001, left; Thapan *et al.*, 2001, right).

The differences in SRCs for each system (C(λ), S(λ), V(λ)) are dependent on the specific photoreceptor involved and their related photopigments. The response of ipRGCs –

associated with C(λ) – is driven by the spectral sensitivity of the photopigment melanopsin (Lucas *et al.*, 2015). This is markedly different from the rods – containing rod-opsin, which has peak sensitivity (λ_{max}) at ~500nm – and cones – S-cones, M-cones and L-cones – associated with the scotopic and photopic systems. S-cones contain the short wavelength sensitive cyanolabe-opsin, with peak sensitivity at ~420nm; M-cones feature the photopigment clorolable-opsin, with peak sensitivity at ~535nm; L-cones contain a red-shifted erythrolable-opsin, with peak sensitivity at ~565nm (Stockman & Sharpe, 2000).

Despite only measuring melatonin suppression responses to varying wavelengths of light, Cajochen *et al.* (2005) and Lockley *et al.* (2006) also found that short wavelength sensitivity to the alerting effects of light were more significant at light exposure of 460 nm compared to 555 nm during the biological night. From a review of the literature, it is therefore evident that the peak spectral sensitivity of the non-visual system is somewhere within the short-wavelength (blue) part of the electromagnetic spectrum. That is, the action spectra related to the non-visual system are significantly divergent from the sensitivity associated with the visual system. This implies that the most widely used photometric unit of measure in lighting research – i.e., the photopic lux – is not appropriate to predict the response of the melanopsin photoreceptor (al Enezi *et al.*, 2011; Lucas *et al.*, 2013). In essence, measuring light exposures using traditional photopic (and scotopic) illuminance (as discussed in Section 2.5.6) gives an incorrect indication of the intensity of illumination experienced by the neural pathway that is responsible for the non-visual effects of light (Veitch, 2005; Cajochen, 2015).

2.5.8. Duration

Studies have shown that the duration of light exposure could be just as significant as timing to trigger non-visual effects. Gronfier *et al.* (2004) found that separate exposures of light consisting of six 15 minute pulses of ~9,500 lux, interspaced by 60 minutes of very dim light (<1 lux), had a greater effect in delaying phase shifts in the human circadian pacemaker than a continuous light source of ~9,500 lux (Figure 2-5). Both conditions were more effective at phase delaying than continuous very dim light conditions.

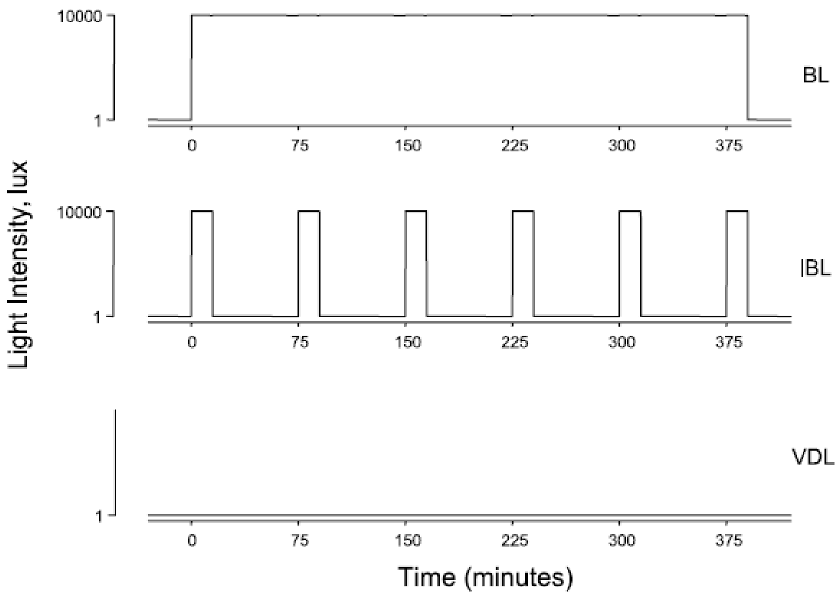


Figure 2-5 Times and duration of exposure (x-axis) against light intensity (y-axis) in experimental bright light exposure conditions for 6.5h (N= 20); BL= Bright Light Condition (continuous) (top); IBL= Intermittent Bright Light Condition (interval pulses) (middle); and, VDL= Very Dim Light (continuous) (bottom) (Gronfier *et al.*, 2004)

Chang *et al.* (2012) used a wide-range of exposure durations, with a target corneal light illuminance of >6,000 lux, using ceiling-mounted 4,100K fluorescent lamps. The study demonstrated that brief durations of bright light were more efficient for phase-shifting, melatonin suppression, and increasing alerting effects, than longer periods of stimulation. St. Hilaire *et al.* (2012) showed that a one hour exposure to bright light resulted in 40% of non-visual response when compared to bright light exposure of 6.7 hours. This study used ceiling-mounted 4,100 K fluorescent lamps producing vertical plane light levels of 8,020 lux.

2.5.9. Prior Light Exposure

Wong *et al.* (2005) found that the ipRGCs adapt to both light and dark conditions, and are strongly influenced by prior lighting history. These authors noted that the observed adaptation appears to reflect a change in the neural pathway between melanopsin to the light active ion channel in the plasma membrane. Their study suggested that the ability of adaptation shown by the ipRCGs helps photic entrainment operate under a wide range of ambient lighting conditions. For example, daylight exposure in winter is less intense than at other times of the year, and the ipRGCs would generate a much weaker signal unless they were able to adjust their sensitivity. However, according to the literature, the effect of prior light history on non-visual processes is not yet fully characterised in comparison to the other known factors that mediate photobiological responses (Chang *et al.*, 2011).

2.5.10. Stimulants Intake: Caffeine Ingestion and Food Intake

According to the literature, circadian rhythms can be entrained by time-dependent schedules of light exposure, caffeine, and food intake (Mistlberger, 2011).

Studies have found that the combination of caffeine ingestion and bright light exposure (≥2,000lux) supressed melatonin secretion and reduced the normal night-time decrease in core body temperature (Wright Jr *et al.*, 1997: 2000). This effect resulted to be larger than the individual influence of these stimuli, suggesting that the effects of caffeine ingestion and bright light exposure were additive. Caffeine ingestion has also been associated with increases in alertness and reduction in fatigue particularly during night-time work, and has been related to behavioural responses (i.e., mood) (Smith, 2002). In fact, abstinence or withdrawal from caffeine has been linked to negative mood effects (Smith *et al.*, 2005).

Food intake can also have profound behavioural and physiological effects depending on the time of day (Stephan, 2002). Studies have suggested that feeding regimes and metabolism could also play a key role in the resetting of the circadian clock, which could lead to changes in life span, aging, well-being, and the occurrence of diseases (Froy & Miskin, 2010).

2.5.11. Summary

For more than a century, lighting scientists and researchers have considered the rods and the cones to be the only photoreceptors in the human retina, with sole responsibility over visual photoreception (van Bommel, 2006). In response, building regulations have been designed to exclusively meet requirements for vision. Nevertheless, some consideration to the non-image forming effects of light has recently started to be given in standards (for example, within the Society of Light and Lighting code (SLL, 2012)), although findings from photo-biological research are not yet fully grounded, also due to the fact that studies to date have mostly focused on exposure to monochromatic or static artificial lighting (Andersen, 2015).

There is a need to fully characterise the impacts of light received by the non-visual system, so that recommendations can be made also for healthy non-visual stimulation (CIE, 2015; Lucas *et al.,* 2013). To transfer the findings of non-visual research in the design of the built environment, it is first of all important to consider which proxies should be triggered (and when), as well as the criteria required for their stimulation. At biological night, in fact, the sensitivity of the non-visual system to light exposure is heightened. This means that artificial lighting of a certain wavelength may be capable of triggering melatonin suppression and phase-delaying responses in the human body. Disruptions of these rhythms can have negative consequences on health and well-being, and therefore have to be carefully considered in the design of buildings. On the other hand, it is reasonable to assume that the high alerting effects of light during the biological day would be beneficial in an office environment. Nevertheless, the lower sensitivity of the non-visual system at daytime means that much higher target illuminances are required to trigger alerting responses. This may present a risk of increased visual discomfort and overheating from solar gains, resulting in costly energy demands.

An important output of photo-biological research is represented by findings that have practical application in real-word settings, whereby the benefits of increased alertness, improved health and well-being, can be captured and replicated. To date, the translation of photo-biological research into target lighting goals has looked specifically at whether these objectives can be met in indoor spaces through computational simulation or photometric measurements. However, existing criteria for the prediction of discomfort glare may stand as a potential barrier to an “active” use of daylight in buildings. The evaluation of the relationship between visual discomfort and non-visual stimulation is, therefore, necessary if the potential benefits of photo-biological research are ever to be realistically achieved.

Section 2.4 and Section 2.5 have provided a comprehensive overview of the literature in visual discomfort and in non-visual responses to light. In this context, Boyce (2004) suggested that, within lighting research related to the built environment, visual performance, visual discomfort, and the non-visual system can be considered as the three main areas of interest. Nonetheless, to date, very little attention has been focused on the interactions between these areas. One reason for this is that research related to these domains is often being conducted independently from each other, disregarding the comprehensive interfaces and overlaps that potentially exist between them and resulting in gaps in our understanding. For example, Mardaljevic *et al.* (2012) stressed that the interactions across the proxies (see Section 2.5.5 to Section 2.5.9) related to the non-visual system are not yet well characterised.

The literature shows that the non-visual system presents changes over time, as documented in Section 2.5.5. Although there is currently not sufficient evidence to demonstrate that the processes underlying visual and non-visual responses are connected, it can be hypothesised that the visual discomfort system may exhibit similar temporal variations across the time of day to the non-visual one. In addition, a number of variables have been associated with non-visual responses to light. For example, intense light exposure has been linked to increased alertness and decreased levels of fatigue (Section 2.5.6); prior light history has been found to influence the non-visual system's sensitivity to light (Section 2.5.9), etc.

In considering the potential influence of these variables on the primary effect of interest of this study (i.e., the effect of time of day on glare response), it must be taken into account that these factors cannot be experimentally controlled. The same can be said for other variables (i.e., temporal and personal factors), also varying across the time of day, which may influence the outcomes of this investigation. In this regard, it may be expected that, if these variables are not promptly identified under controlled laboratory conditions (stage one of the study), a large scatter in the reported levels of glare sensation might appear across the time of the day, consistently with the literature. This would ultimately confound the presence of the effect of experimental interest under less controlled daylit conditions (stage two of the study).

Rodriquez & Pattini (2014) stated that although it may not be feasible to completely eliminate the scatter commonly associated with the subjective evaluation of glare sensation, individual differences may be mitigated if the effects of influencing variables (e.g., fatigue, prior daylight exposure, etc.) are systematically identified. In response, a series of experiments under controlled conditions initially needs to be conducted so as to identify the influences of these variables in conjunction to the detection of the effect of time of day on glare response. These studies are necessary in preparation to examining the experimental hypothesis in the presence of daylight from a window.

Chapter 3

Laboratory Studies

3. Laboratory Studies

This chapter describes a series of experiments conducted under controlled laboratory conditions. The first of these experiments aimed to isolate the research hypothesis, i.e. to search for an effect of time of day on the perceived level of glare sensation reported by subjects at four testing periods equally spaced at 3-hour intervals. A follow-on experiment was designed to exclude that the detected trend (i.e., the effect of experimental interest) was statistically related to a possible confounding effect of learning, hence providing supportive evidence of an effect of time of day on glare sensation. During these tests, several temporal variables (e.g., fatigue, food intake, caffeine ingestion, mood, previous daylight exposure, sky condition), and personal factors (e.g., age, gender, ethnicity, chronotype, photosensitivity), were also measured to analyse their influence on the reported levels of visual discomfort along the day. Finally, an experiment was designed to explore the relationships between visual task difficulty, temporal variables, and glare response as the day progresses.

3.1. Discomfort Glare and Time of Day

3.1.1. Introduction

The literature suggests that glare sensation is characterised by large individual differences (Stone & Harker, 1973; Tregenza & Wilson, 2013); yet, there are still significant gaps in our understanding of the phenomenon (Boyce, 2014; Fotios, 2015). The research studies documented in the previous sections provide strong reasons to suspect that individual glare response might depend on variables that differ from the four physical and photometric parameters commonly embedded in glare calculation models. On this basis, this investigation was set to initially explore the effect of time of day on glare sensation under laboratory conditions. In fact, a tightly controlled experiment allowed controlling the four variables that are found in most conventional glare indices (source luminance, source size, background luminance, and solid angle), while isolating several other parameters that – from a review of the literature – are alleged to influence discomfort glare but that are not included in glare formulae. Along with this, all extraneous variables (e.g., daylight, temperature, and humidity) could be 'masked' within this experiment that was fully conducted in the presence of artificial lighting. This was done to reduce the anticipated scatter in responses commonly associated with subjective evaluation of glare sensation (Tregenza & Wilson, 2011), hence increasing the sensitivity of the experimental effect of interest, and enabling the potential detection of the postulated effect of time of day also in case of small influence on the dependent variable (glare response).

3.1.2. Method

3.1.2.1.1. Experimental Setup

To verify whether an effect of time of the day on the reported level of glare sensation could be detected, a laboratory experiment under controlled artificial lighting conditions was considered appropriate. This was based on the postulation that the influence of time of the day on visual response might be general, and hence can be initially isolated by means different than real windows where several other parameters would also be continuously changing over time (Tuaycharoen & Tregenza, 2005). The design of the apparatus (Figure 3-1 and Figure 3-2) was informed by the study from Tuaycharoen & Tregenza (2005), which used a similar setup to investigate the influence of view interest on discomfort glare.

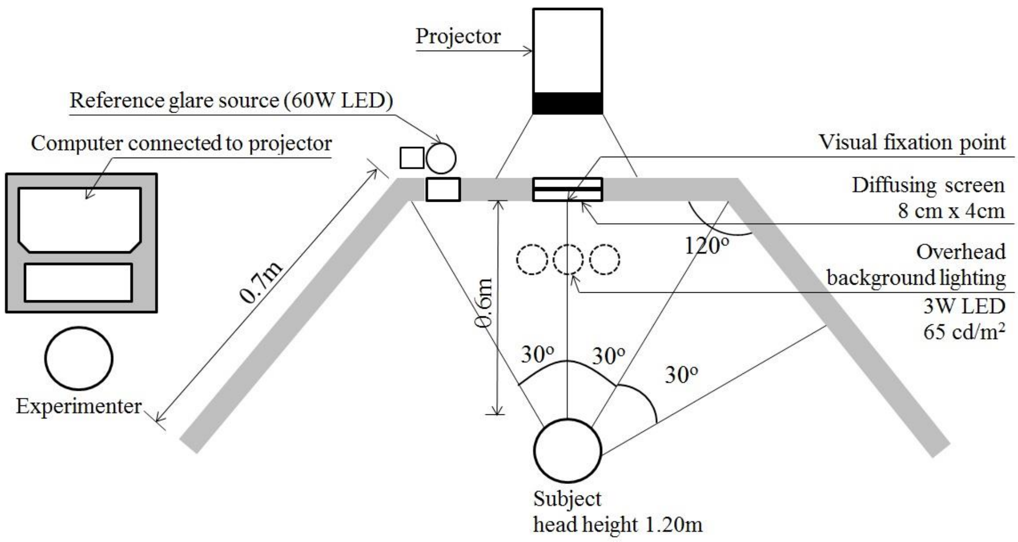


Figure 3-1 Layout of the experimental lighting chamber


Figure 3-2 Internal view of the experimental lighting chamber

The chamber was semi-hexagonal in plan, the interior surfaces (2.7 m in height) were painted matte white, and three 3W LED lamps, mounted from above, were used to produce an average of 65 cd/m^2 lighting, maintaining a constant background luminance distribution throughout the experiment. Photometric measures were checked before and after each experimental test session, to ensure they did not vary (Tuaycharoen, 2006). The shape and size of the cubicle was designed to cover the entire field of view of the observer for binocular vision, spanning from 60° left to 60° right, and from 53° above and 67° below relative to the line of sight (Boff & Lincoln, 1988). The subjects' head position was located at a height of 1.2 m from the floor and at the centre of the apparatus, facing two light sources. The first was positioned on the direct line of sight of the subject, in the middle of the central 'wall', and comprised a small diffusive screen made from two sheets of tracing paper (0.08 x 0.04 m^2), which was mounted directly in front of a projector connected (via VGI) to a computer. The second was presented at an angle of around 30° to the left of the line of sight, and used a 60W LED lamp to provide a small reference glare source (Figure 3-3).

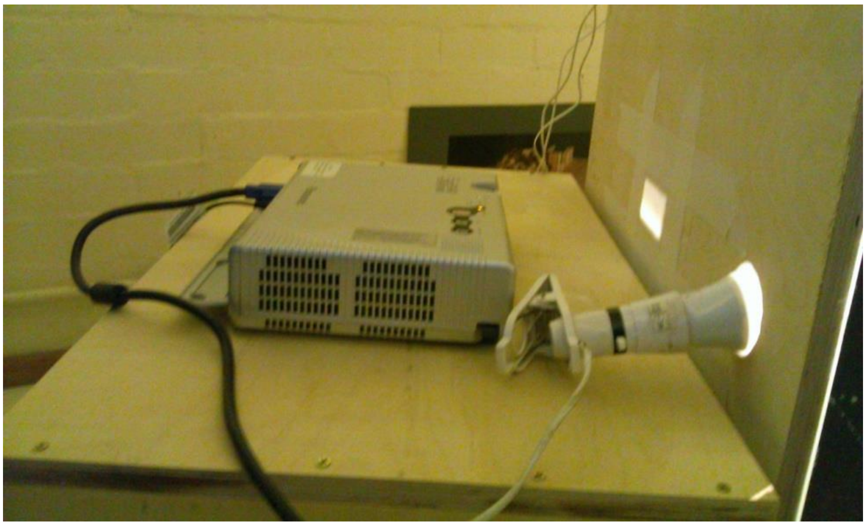


Figure 3-3 Computer projector, diffusive screen, and 60W LED lamp used in the experiment

According to the literature, conventional glare indices used for evaluating glare sensation from artificial lighting are not designed to quantify the effect of small sources above 0.01 steradians or non-uniform sources (Velds, 1999; Wienold & Christoffersen, 2006). For these reasons, the diffusive screen represented a condensed source of glare, subtending an angle at the eye of 0.009 steradians and providing a uniform source of luminance that could be varied in the range between 400 cd/m^2 and 20,000 cd/m^2. The additional reference glare source was connected to a dimmer directly controlled by the test subject. Experimental conditions were kept constant throughout the testing, with the only difference being represented by the variation of the luminance of the small diffusive screen controlled by the computer.

3.1.2.2. Experimental Procedure

The experimental procedure was designed so that parameters known to influence discomfort glare – i.e., source luminance, source size, background luminance, position index, luminance distribution, source uniformity, visual task, view, etc. – could be controlled across the variable of experimental interest (time of day). Since no established methodology could be found in the literature to detect a potential effect of time of the day on the reported level of glare sensation, the framework developed by Mardaljevic *et al.* (2012) to describe how the non-visual system responds to light at different times of the day was adopted for this experiment. Based on this, the experimental procedure requested test subjects to participate, on the same day, to four test sessions evenly distributed at 3-hour intervals:

- Morning: 09:00 or 09:30;
- Midday: 12:00 or 12:30;
- Afternoon: 15:00 or 15:30;
- Evening: 18:00 or 18:30.

A total of 30 test subjects volunteered to take part to this experiment, which was carried out between the months of January and March. Membership to the sample group was tightly controlled. All participants were postgraduate architecture students, 17 male and 13 female, the mean age was 24.10 (standard deviation (SD)= 3.21), and 10 wore corrective lens.

Subjects could elect whether to enter the sequence of sessions at 09:00 or rather at 09:30, while keeping tests always distributed at 3-hour intervals for each participant. Conducting two sequences of experimental sessions per day made it possible to minimising the total testing time, therefore reducing the potential influence of seasonal changes on the data collected. In fact, the literature reveals that during winter months observers may become more acceptant of discomfort glare (Altomonte, 2009; Christoffersen *et al.*, 2000). This consideration becomes relevant when testing a large number of subjects, so steps had to be taken to concentrate the tests on a relatively short period of time so as to avoid any influence of changes in daily lighting patterns on visual response. In addition, the results of a pilot study (N= 10) – which utilised the same procedure of this experiment but was based on a between-subject design – had revealed that no significant differences ($p>0.05$) could be detected if intervals between test sessions were kept below 1.5 hours. Hence, a gap of 30 minutes between the two sequences of test sessions was considered sufficiently guarded from any temporal discrepancy in collecting results from different subjects. Finally, an interval of 3 hours between the test sessions that each participant was required to attend was deemed adequate to monitor the response of subjects on a free-running day and prevent disruption from the participants' daily routine, which could have potentially influenced their glare response.

In the design of the study, it was also considered that maintaining the same sequence of test sessions for all subjects – under systematic variation (Field & Hole, 2013) – could have potentially masked an effect of learning into the design procedure. Indeed, Hopkinson (1950) hypothesised that a less experienced subject could not be able to interpret in a consistent manner the meaning of discomfort glare descriptors, hence possibly biasing the responses of the initial session(s). To address these concerns, a follow-up study was conducted after the main experiment, so as to examine the eventual influence of learning on the results obtained.

At the beginning of the test, the subject was asked to adjust the stool so that their head was properly located at the viewing position, at a height of 1.2 m above the floor and at a distance of 0.6 m from the visual fixation point (i.e., the diffusive screen). These measurements were carefully controlled by the experimenter using a tape measure, yet tolerating minimal head movements by the subjects. A clear set of instructions was then given, including a definition of discomfort glare, a description of the four benchmarks of Glare Sensation Vote (GSV) that the subjects were requested to use ('Just Perceptible', 'Just Noticeable', 'Just Uncomfortable', and 'Just Intolerable'; these are discussed in detail in Section 3.1.2.4), and an illustration of how the experiment would run. To confirm that subjects had a proper understanding of the GSV criteria, they were initially asked to look at the reference glare source and to adjust its brightness (by a dimmer located by their seat) so that the glare sensation produced was judged by them as, progressively, 'Just Perceptible', 'Just Noticeable', 'Just Uncomfortable' and 'Just Intolerable'. The reference glare source was then turned off.

The subject was then asked to direct their gaze towards a fixation point located at the centre of the small diffusive screen and to imagine that this represented a visual task. At this point, the luminance of the screen started to be increased, at a controlled and constant pace, by the experimenter via the computer connected to the projector. Progressive intensification of the luminous stimulus controlled on the computer caused the luminance of the screen to increase based on a polynomial function. Throughout the procedure, the subject was asked to vocally indicate when the sensation of discomfort glare due to the increasing luminance of the diffusive screen became, respectively, 'Just Perceptible', 'Just Noticeable', 'Just Uncomfortable' and 'Just Intolerable'. The photometric values at which each GSV was reported by the subject were recorded. Each test session lasted around 10 minutes.

3.1.2.3. Photometric Measurements and Glare Indices

Before the start of the experiments, photometric measurements were taken from the position of the test subjects' eye using a spot-point Minolta LS-110 luminance meter mounted on a tripod. The mean background luminance was calculated from 17 independent measurements (Figure 3-4) taken on a regular grid symmetrical about the central fixation point, and extending across the width of the cubicle (Tuaycharoen, 2006). Table 3-1 presents the resultant spot-point luminance measurements (cd/m^2) associated with the points located on the cubicle walls (Figure 3-4). These points were mean averaged, resulting in a background luminance that was held at constant 65cd/m^2 (SD= 23.59) throughout the experiment. This value is within the range commonly associated with office interior spaces (CIBSE, 1996).

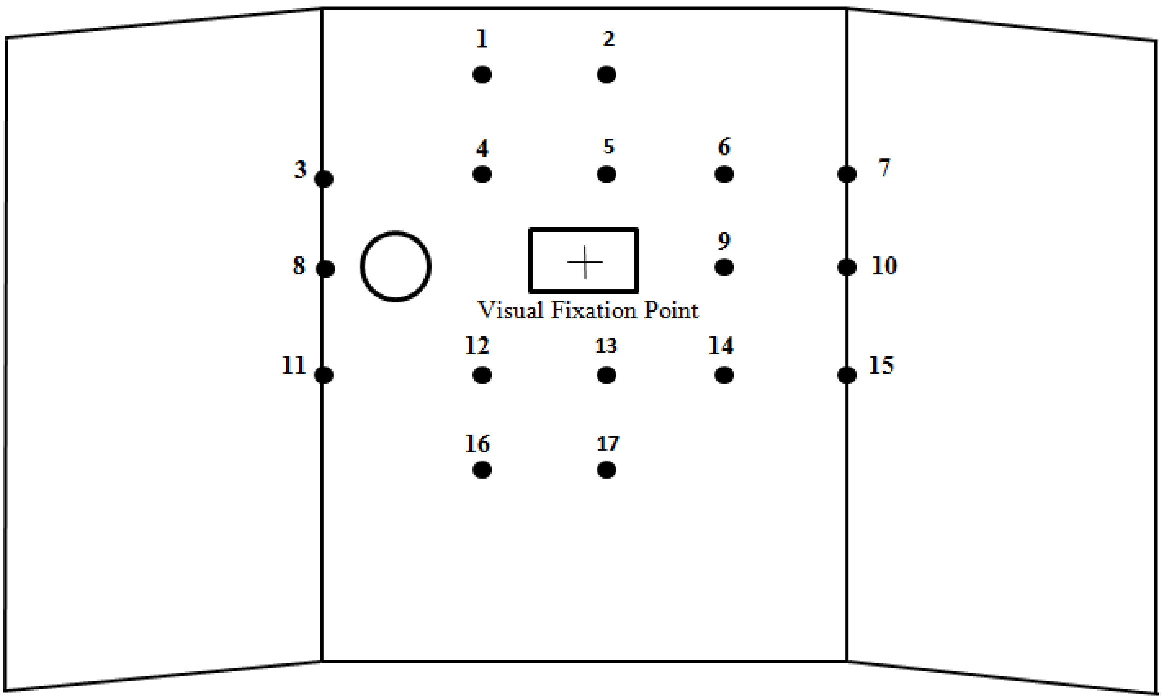


Figure 3-4 Background luminance spot-point measurements (points 1 to 17). The figure also shows the position of the 60W LED lamp (circle) and of the diffusive screen (rectangle) containing the visual fixation point (cross)

Table 3-1 Spot-point luminance measurements and mean luminance values

Point Measurement	**Spot-point Luminance (cd/m^2)**
1	103.60
2	114.70
3	66.88
4	79.10
5	98.00
6	85.24
7	68.80
8	53.06
9	59.00
10	52.99
11	40.62
12	46.62
13	46.28
14	49.27
15	47.34
16	42.55
17	41.15
Average (mean)	64.42

The luminance of the diffusive screen was also evaluated by spot-point measurements. The source luminance was directly manipulated through operation of the computer connected to the projector, using the relative brightness function of an image editing software (Figure 3-5).

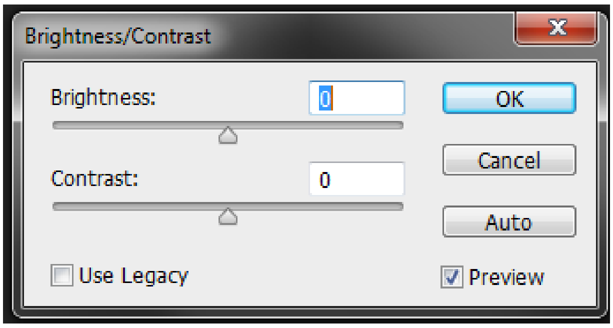


Figure 3-5 Relative brightness function used to adjust the luminance output of the projector

In order to achieve precise luminance outputs in repeated procedures, the projector had to be calibrated to the image editing software. The relative brightness function was adjusted at evenly distributed intervals, and three independent spot-point luminance measurements were taken at each interval (Figure 3-6).

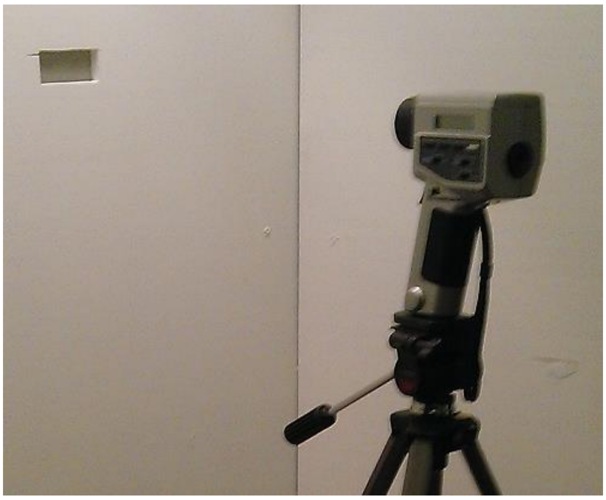


Figure 3-6 Minolta LS-110 luminance meter mounted on a tripod perpendicular to the diffusive screen used to take independent spot-point luminance measurements

The independent spot-point measurements were mean averaged and then interpolated using a polynomial function to obtain values between the calibrated luminance points.

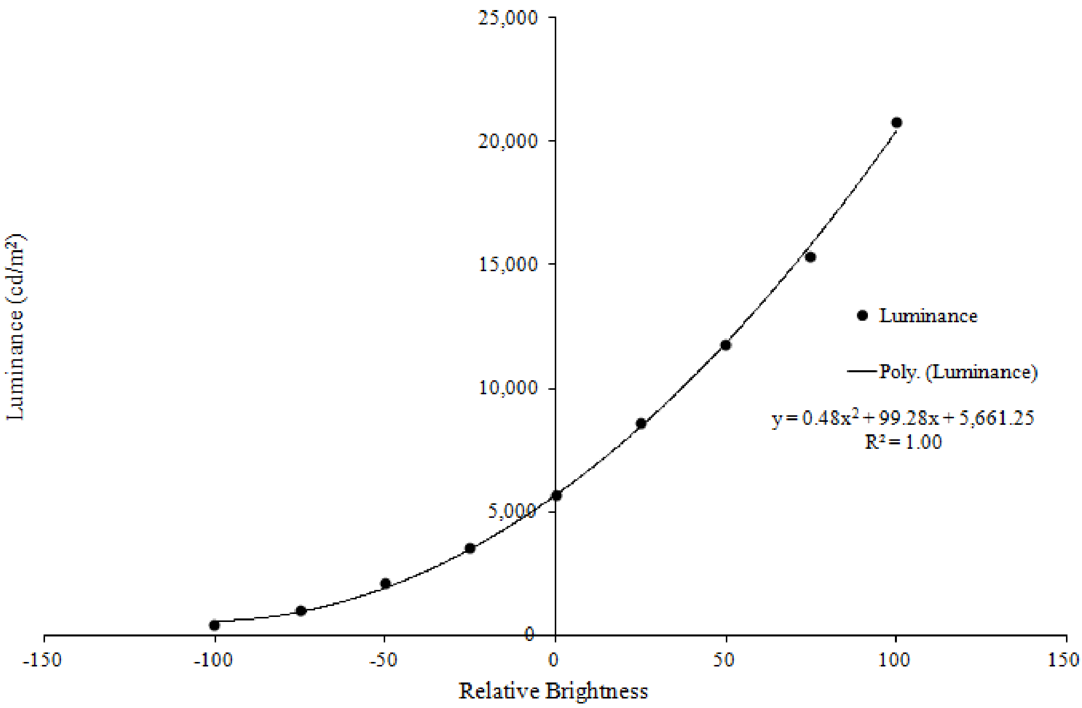


Figure 3-7 Comparison between relative brightness values (x-axis) and measured luminance (y-axis) of the diffusive screen at equal intervals. The points correspond to mean average luminance scores (N= 3) for each interval with a polynomial fit

Figure 3-7 plots the relative brightness function corresponding to the recorded spot-point measurements (x-axis), and the resultant mean average luminance values (y-axis). A polynomial interpolation, with its associated curve equation and coefficient of determination (R^2), is also displayed.

To minimise the influence of the glare source luminance on the background luminance of the central partition, the diffusive screen was placed at the back of the cubicle wall, which had a depth of 12 mm. The reveal was painted matte white to aid diffusing light from the glare source, and preventing concentrations of high luminance in this area.

To be consistent with the literature (Tregenza & Wilson, 2011) and the experimental design described in Tuaycharoen & Tregenza (2005), glare indices were used in this study as the primary evaluation parameter. The IES-GI – a glare formula commonly found in lighting codes (Equation 3-1) – was used to objectively quantify subjective assessments of glare sensation from the artificial lighting source. Results were calculated also using the UGR glare index formula (Equation 3-2), and they were found to follow the same trends of the IES-GI. For this reason, results obtained using the UGR have been included in Appendix F.

$$IES-GI = 10Log_{10}0.478\sum\left(\frac{L_S^{1.6}\cdot\omega^{0.8}}{L_b\cdot P^{1.6}}\right) \qquad (3\text{-}1)$$

$$UGR = 8Log_{10}\left(\frac{0.25}{L_b}\sum\frac{L_S^2\cdot\omega}{P^2}\right) \qquad (3\text{-}2)$$

Where:

- L_s = Source luminance (cd/m^2);
- L_b = Background luminance (cd/m^2);
- ω = Subtended size of the source (sr);
- P = Position index (-).

Both the IES-GI and the UGR carry high, although differently weighted, exponentials for source luminance, a parameter that has been strongly correlated to the experience of glare (Rodriguez & Pattini, 2014). The sensitivity of these glare indices to changes in source luminance made them suitable to detect the effect of time of day in this experiment. In fact, since these glare indices are well-tested tools, any variability within the index corresponding to the GSVs provided by subjects gives an indication that there is a factor, other than the parameters embedded within the formula itself, which is influencing glare response.

3.1.2.4. Subjective Assessments

Based on the literature, discomfort glare is a personal sensation that requires subjective methods of evaluation (Velds, 2002). In accordance with this, during the sessions, subjects were requested to make judgements of glare sensation utilising as benchmarks adaptations of the Glare Sensation Votes (GSVs) used by Iwata *et al.* (1992a; 1992b), Iwata & Tokura (1998), and Mochizuki *et al.* (2009). GSVs correspond to thresholds of visual discomfort linked to increases in magnitude of glare sensation that a test subject experiences: 'Just Perceptible', 'Just Noticeable', 'Just Uncomfortable', and 'Just Intolerable'. Since it was considered that these criteria might be open to self-interpretation (e.g., due to the abstraction caused by the assessment method), to aid subjects providing more meaningful judgements, the GSV criteria were linked to time-span descriptors (MacGowan, 2010; Velds, 2002):

- 'Just Perceptible' (GSV= 0). This would be the borderline between imperceptible and perceptible glare, as represented by the point where glare is first perceived;
- 'Just Noticeable' (GSV= 1). This would be the glare that the subject could tolerate for approximately 1 day, when working in someone else's room. However, if the subject

had to work under these lighting conditions in their own room, they would likely use visual protection controls (e.g., blinds);

- 'Just Uncomfortable' (GSV= 2). This would be the glare that the subject could tolerate for approximately 15 to 30 minutes, if the task took this amount of time to complete. After this period, adjustments to the lighting environment would be made;
- 'Just Intolerable' (GSV= 3). This would be the glare that the subjects could no longer withstand, immediately acting to modify the lighting environment.

Hopkinson (1950) stated that if a criterion is not well informed, subjects could express doubts as to their own perceived judgement of the glare source, this potentially increasing scatter in the responses collected. On these bases, it was considered that time-span descriptors would provide a clearer explanation of the GSV criteria, so that these could be understood by the subjects according to the intended interpretation of the experimenter (Fotios, 2015).

3.1.2.5. Statistical Analysis

To analyse the data, a Friedman's analysis of variance (ANOVA) was performed to compare calculated values of IES-GI for all four test sessions against each other. The inferential tests adopted to explore temporal variation in glare response for each criterion of reported GSV were based on the following null (Equation 3-3) and alternative hypotheses (Equation 3-4):

$$H_0: R_{Morning} = R_{Midday} = R_{Afternoon} = R_{Evening} \quad \text{(Equation 3-3)}$$

$$H_1: R_{Morning} \neq R_{Midday} \neq R_{Afternoon} \neq R_{Evening} \quad \text{(Equation 3-4)}$$

Where R is the cumulative score of the ranks associated with an independent variable (test session).

The Friedman's ANOVA was used since graphical (e.g., Quantile-Quantile (Q-Q) plots) and statistical inspection (e.g., Shapiro-Wilk (S-W) and Kolmogorov-Smirnov (K-S) tests) of the data revealed that sampling distributions were not normally distributed around the mean, hence violating one of the assumptions associated with a parametric test (Heiman, 2013).

Figure 3-8 presents selected Q-Q plots related to consideration of the 'Just Perceptible' criterion in the morning (left), and the 'Just Intolerable' GSV in the evening sessions (right). For purposes of graphical clarity, the y-axis and x-axis of both displays are not to scale.

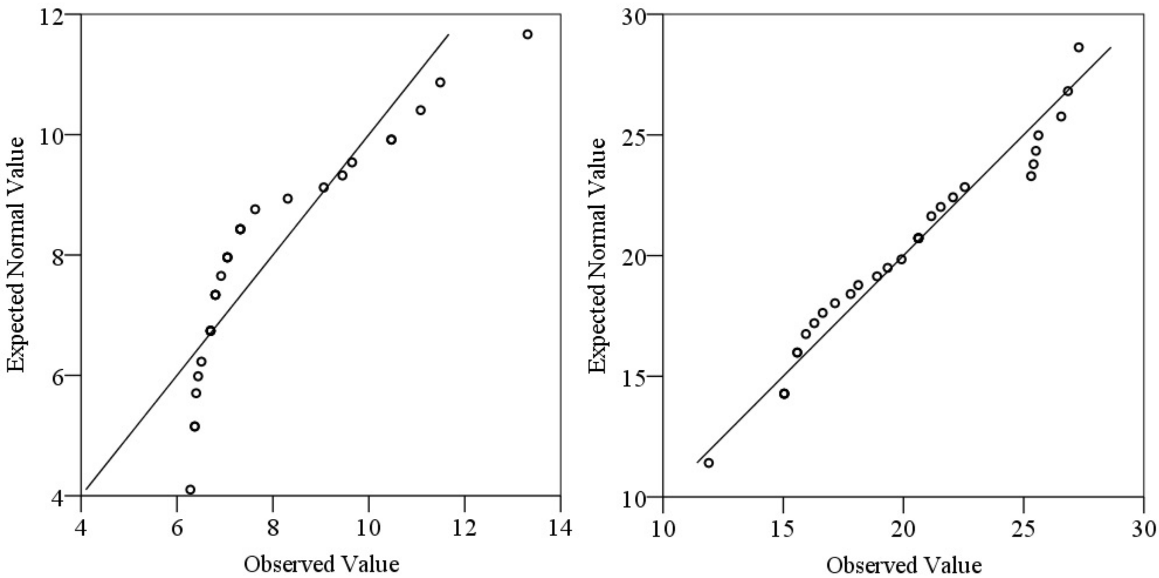


Figure 3-8 Q-Q plots representing the distribution of data for the GSV criteria of 'Just Perceptible' at the morning session (left), and 'Just Intolerable' at the evening session (right)

The figure presents, on the x-axis, the IES-GI calculated at the point in which test subjects reported the GSV criteria and, on the y-axis, values from a population that is representative of data normally distributed around its mean (Field, 2013). The points that fall close to the straight line represent a normal distribution (Wilk & Gnanadesikan, 1968). The figure shows a progressive pattern of deviation from a normal distribution, both at high and low values (more markedly) of IES-GI, a typical characteristic of negatively skewed data (Field, 2013).

Table 3-2 presents the test statistic, the degrees of freedom (df), and the statistical significance (*p*-value) of the results of the K-S and S-W tests for each time of day and GSV criterion. Both the K-S and S-W tests are suitable for analysing whether the distribution of the data under examination is statistically different from a population that is representative of a normal distribution (Shapiro & Wilk, 1965).

Table 3-2 Results of the Kolmogorov-Smirnov and Shapiro-Wilk tests

GSV	Time of Day	Kolmogorov-Smirnov			Shapiro-Wilk		
		Statistic	df	*p*-value	Statistic	df	*p*-value
Just Perceptible	Morning	0.29	30	0.00***	0.78	30	0.00***
	Midday	0.18	30	0.01**	0.81	30	0.00***
	Afternoon	0.24	30	0.00***	0.80	30	0.00***
	Evening	0.20	30	0.00**	0.80	30	0.00***
Just Noticeable	Morning	0.23	30	0.00***	0.84	30	0.00***
	Midday	0.19	30	0.01**	0.89	30	0.00**
	Afternoon	0.17	30	0.02*	0.91	30	0.01**
	Evening	0.21	30	0.00**	0.88	30	0.00**
Just Uncomfortable	Morning	0.16	30	0.05 n.s.	0.93	30	0.06 n.s.
	Midday	0.13	30	0.20 n.s.	0.98	30	0.74 n.s.
	Afternoon	0.16	30	0.06 n.s.	0.94	30	0.09 n.s.
	Evening	0.12	30	0.20 n.s.	0.94	30	0.10 n.s.
Just Intolerable	Morning	0.09	30	0.20 n.s.	0.98	30	0.72 n.s.
	Midday	0.08	30	0.20 n.s.	0.98	30	0.89 n.s.
	Afternoon	0.10	30	0.20 n.s.	0.97	30	0.50 n.s.
	Evening	0.13	30	0.20 n.s.	0.95	30	0.17 n.s.

* weakly significant; ** significant; *** highly significant; n.s. not significant

The results demonstrate that, for the K-T tests, the differences detected were highly significant in 3 cases, significant in 4 cases, weakly significant in 1 case, and not significant in 8 out of 16 cases. The differences detected by the S-W tests were highly significant in 5 cases, significant in 3 cases, and not significant in 8 out of 16 cases.

The results from the graphical (Q-Q plots) and statistical (K-S and S-W) tests, therefore, provided evidence that the data were not normally distributed around the mean parameter.

The Friedman's ANOVA tests account for non-normally distributed data across independent variables by ranking the data (Friedman, 1937; 1940). The tests allocate the independent variable (test session) corresponding to the largest value of calculated IES-GI with the highest rank (i.e., 4 – this being the number of test sessions), and then simply add up the ranks for each independent variable (denoted as R_i, where i represents the independent group).

For all independent variables (test sessions) and for each GSV criterion under consideration, a non-parametric Levine's test of homogeneity of variance – a test suitable for non-Gaussian distributions (Nordstokke & Zumbo, 2010; Nordstokke *et al.*, 2011) – was conducted to verify that this assumption was not violated, examining whether the differences associated with the variances at each test session were statistically significant and practically relevant. This required testing the differences between four variables (test sessions) across different independent variables. As an example, the null (Equation 3-5) and alternative hypotheses (Equation 3-6) for consideration of the relationship that exists between the variances associated with each test session were as follows:

$$H_0: \sigma_{Morning} = \sigma_{Midday} = \sigma_{Afternoon} = \sigma_{Evening} \qquad \text{(Equation 3-5)}$$

$$H_1: \sigma_{Morning} \neq \sigma_{Midday} \neq \sigma_{Afternoon} \neq \sigma_{Evening} \qquad \text{(Equation 3-6)}$$

Where σ is the variance associated with the independent variable under examination.

The non-parametric Levine's test inherits the limitations of null hypothesis significance testing (NHST). In fact, in NHST, the statistical significance (*p*-value) is dependent on both the size of the effect and on the size of the sample (Ellis, 2010). Therefore, in interpreting the results of the tests, emphasis was placed not solely on their statistical significance (which may be confounded particularly with large sample sizes) but also on the effect size (i.e., a standardised measure of the difference between independent groups). The effect size shows if the predictor variable has any practical significance, and provides a more rigorous support to inferences (Cohen, 1965; Schiavon & Altomonte, 2014). In this analysis, the effect size was calculated by extrapolating both the sum of squares of the effect (SS_{Effect}) and the sum of squares total (SS_{Total}) from the one-way ANOVA – utilised to perform the non-parametric Levine's test – according to the following formula (Field & Hole, 2013) (Equation 3-7):

$$\eta^2 = \frac{SS_{Effect}}{SS_{Total}} \qquad \text{(Equation 3-7)}$$

The interpretation of the outcome was derived from the tables provided by Ferguson (2009), where conventional benchmarks have been proposed for 'small', 'moderate', and 'large' effect sizes ($\eta^2 \geq 0.04$, 0.25, and 0.64, respectively). Values of $\eta^2 < 0.04$ were considered as non-substantive (i.e., 'negligible') and, thus, not presenting any practically relevant effect.

Table 3-3 presents the *F* test statistic and degrees of freedom, the statistical significance (*p*-value) and the effect size (η^2) of the non-parametric Levene's test of homogeneity of variance.

Table 3-3 Non-parametric Levene's test of homogeneity of variance

Glare Sensation Vote	***F***(3, 116)	*p*-value	η^2
Just Perceptible	1.98	0.12 n.s.	0.05
Just Noticeable	0.99	0.40 n.s.	0.03
Just Uncomfortable	1.43	0.24 n.s.	0.04
Just Intolerable	0.66	0.57 n.s.	0.02

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

The inferential statistics indicate that the differences are, for all comparisons, not statistically significant. Calculation of effect sizes (η^2) demonstrates that the practical relevance of the differences detected for the dependent variable (IES-GI) is negligible ($\eta^2<0.04$) in 2 cases, and small ($0.04\leq\eta^2<0.25$) in 2 cases out of 4. However, in both cases where $\eta^2\geq0.04$, the effect sizes are just on the borderline of practical relevance. These results suggest that any significant and practically relevant influence of violating the assumption of homogeneity of variance can be safely excluded from subsequent inferential tests.

To isolate the main effects within the initial Friedman's ANOVA, follow-up pairwise comparisons were made utilising the Wilcoxon matched-pairs (or signed-rank) test, which is suitable when examining two sets of scores. The Wilcoxon matched-pairs test also accounts for non-normality by ranking the data (Wilcoxon, 1945). The difference in scores (i.e., ΔIES-GI) for each test subject is primarily calculated making note of the sign (positive or negative) and excluding comparisons with a difference of zero. The differences are then ranked (regardless of their sign) from smallest to largest, whereas larger differences (ΔIES-GI) are assigned a higher rank. Finally, the ranks are added up. Grouping the scores by the sign of the difference and the test statistic, standard errors and significant levels (*p*-value) can be obtained. When the assumptions of parametric tests (e.g., normality) are not met, non-parametric counterparts lead to much more powerful inferences (Siegel, 1957).

3.1.3. Results and Analysis

Figure 3-9 plots, on the y-axis, the IES-GI calculated when test subjects reported the various GSV criteria. On the x-axis, the graph displays the votes of glare sensation and the test sessions corresponding to the time of the day at which assessments were given. In Figure 3-9, the outliers (circles) and extreme scores (asterisks) are also plotted, so as to represent scores of calculated IES-GI that extended beyond the box-and-whisker plots (Field, 2013). The visual displays reveal a tendency for the statistical values (e.g., 25th percentile, median, 75th percentile) to correspond to higher levels of calculated IES-GI as the day progresses for all reported GSV criteria.

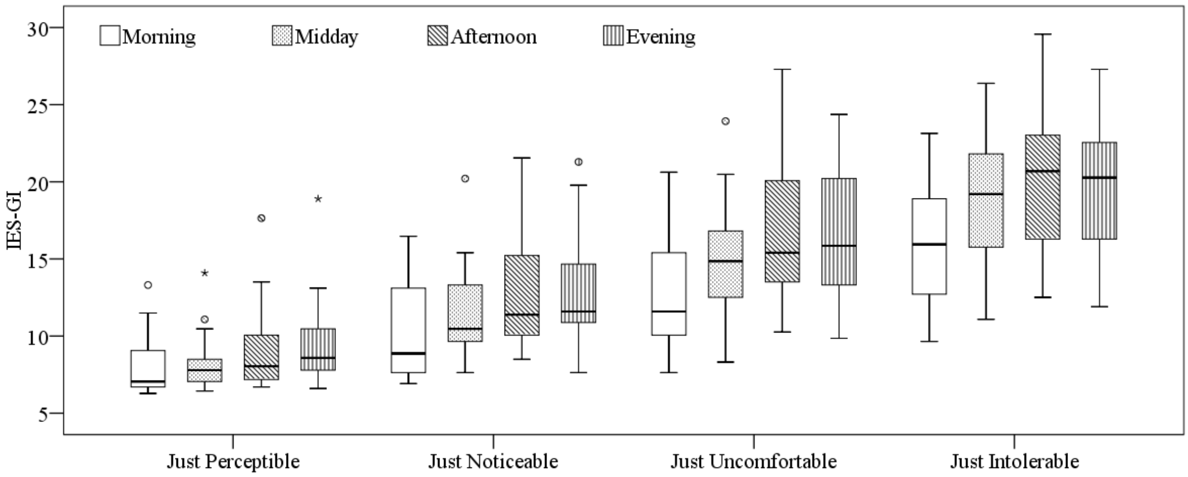


Figure 3-9 Boxplots of IES-GI statistical parameters for the four criteria of GSV

Graphical inspection of the boxplots, therefore, suggests that, under all criteria of GSV, there is a consistent trend for central tendencies to correspond to higher levels of IES-GI at later test sessions. Since – when controlling for background luminance, source size, and position of glare source relative to the observer – subjects provided the same vote of glare sensation ('Just Perceptible', 'Just Noticeable', 'Just Uncomfortable', and 'Just Intolerable') at increasing levels of source luminance as the day progresses, visual analysis of the data leads to hypothesise that, under the experimental controlled setting, participants became more tolerant to the glare source at later test sessions for all levels of GSV.

For the IES-GI, the Friedman's ANOVA detected highly significant differences across the four test periods (i.e., Morning vs. Midday vs. Afternoon vs. Evening): 'Just Perceptible', $\chi^2(2)= 29.60$, $p \leq 0.001$; 'Just Noticeable', $\chi^2(2)= 31.79$, $p \leq 0.001$; 'Just Uncomfortable', $\chi^2(2)= 30.97$, $p \leq 0.001$; and 'Just Intolerable', $\chi^2(2)= 19.79$, $p \leq 0.001$.

To isolate the main effects between variables, *post-hoc* contrasts – utilising pairwise comparisons – were performed, whereby all permutations between times of the day were compared against each other (Bird & Hadzi-Pavlovic, 2013). The statistical difference of the pairwise comparisons was calculated using a one-tailed Wilcoxon matched-pairs test to determine where the variations in IES-GI were within the Friedman's ANOVA under each GSV criterion. Directionality of the one-tailed hypothesis was informed by graphical observation (Figure 3-9) and inspection of the descriptive statistics of central tendencies (Hauschke & Steinijans, 1996). Since all statistical values (25th percentile, median, 75th percentile) – for all times of the day and reported GSVs – presented a consistent trend to increase as the day progresses, one-tailed hypothesis testing was adopted (Ruxton & Neuhauser, 2010). Bonferroni corrections were applied in consideration of the experiment-wise error rate caused by the alpha level (0.05, 0.01, 0.001) inflating across multiple tests carried out with the same hypothesis (Cabin & Mitchell, 2000). This was calculated as $1 - 0.95^n = 0.26$ (thus, risking a 26% probability of making at least one Type I error), where n= 6, i.e. the number of Wilcoxon matched-pairs tests carried out with the same hypothesis (Field, 2013). For each pairwise comparison, the effect size was also calculated by making use of equivalence between the standardised measure of the observed difference and the Pearson's coefficient *r*, by extrapolating the Z_{score} test statistic from the Wilcoxon matched-pairs tests, according to the following formula (Field, 2013) (Equation 3-8):

$$r = \frac{Z_{Score}}{\sqrt{N}} \qquad \text{(Equation 3-8)}$$

Where N is the number of test subjects within each independent variable.

Interpretation of the outcome was derived from the tables provided by Ferguson (2009), where benchmarked values representative of 'small', 'moderate', and 'large' effect sizes ($r \geq$ 0.20, 0.50, and 0.80, respectively) have been proposed. Table 3-4 to Table 3-7 report the median (M_{dn}) and inter-quartile range (IQR) for the IES-GI calculated on the lighting values recorded when test subjects reported their glare sensation (GSV) at all times of the day, the difference (ΔM_{dn}) between the medians for each pairwise comparison and the associated statistical significance (NHST, *p*-value), the ranks for (positive) and against (negative) the hypothesis, the ties, and the effect size (r).

Table 3-4 Wilcoxon matched-pairs tests and effect sizes for 'Just Perceptible'

Time of Day	M_{dn} (IQR)	M_{dn} (IQR)	ΔM_{dn}^{NHST}	Positive	Negative	Ties	Effect Size (r)
Mid. vs. Morn.	7.79 (1.41)	7.05 (2.17)	0.74 n.s.	21	8	1	0.27
Aft. vs. Morn.	8.05 (2.55)	7.05 (2.17)	1.00*	23	6	1	0.43
Even. vs. Morn.	8.59 (2.63)	7.05 (2.17)	1.54**	26	4	0	0.56
Aft. vs. Mid.	8.05 (2.55)	7.79 (1.41)	0.26*	18	8	4	0.44
Even. vs. Mid.	8.59 (2.63)	7.79 (1.41)	0.80***	24	5	1	0.66
Even. vs. Aft.	8.59 (2.63)	8.05 (2.55)	0.54 n.s.	18	10	2	0.33

With Bonferroni Correction: * 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

Table 3-5 Wilcoxon matched-pairs tests and effect sizes for 'Just Noticeable'

Time of Day	M_{dn} (IQR)	M_{dn} (IQR)	ΔM_{dn}^{NHST}	Positive	Negative	Ties	Effect Size (r)
Mid. vs. Morn.	10.47 (3.31)	8.87 (4.79)	1.60*	24	5	1	0.48
Aft. vs. Morn.	11.39 (5.07)	8.87 (4.79)	2.52***	25	5	0	0.68
Even. vs. Morn.	11.60 (3.54)	8.87 (4.79)	2.73***	25	5	0	0.69
Aft. vs. Mid.	11.39 (5.07)	10.47 (3.31)	0.92*	21	9	0	0.53
Even. vs. Mid.	11.60 (3.54)	10.47 (3.31)	1.13*	23	6	1	0.52
Even. vs. Aft.	11.60 (3.54)	11.39 (5.07)	0.20 n.s.	19	9	2	0.23

With Bonferroni Correction: * 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

Table 3-6 Wilcoxon matched-pairs tests and effect sizes for 'Just Uncomfortable'

Time of Day	M_{dn} (IQR)	M_{dn} (IQR)	ΔM_{dn}^{NHST}	Positive	Negative	Ties	Effect Size (r)
Mid. vs. Morn.	14.85 (4.20)	11.60 (4.87)	3.26***	25	5	0	0.66
Aft. vs. Morn.	15.40 (6.56)	11.60 (4.87)	3.81***	26	4	0	0.75
Even. vs. Morn.	15.85 (6.51)	11.60 (4.87)	4.26***	26	4	0	0.74
Aft. vs. Mid.	15.40 (6.56)	14.85 (4.20)	0.55*	21	9	0	0.46
Even. vs. Mid.	15.85 (6.51)	14.85 (4.20)	1.00 n.s.	18	10	2	0.35
Even. vs. Aft.	15.85 (6.51)	15.40 (6.56)	0.45 n.s.	17	13	0	0.08

With Bonferroni Correction: * 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

Table 3-7 Wilcoxon matched-pairs tests and effect sizes for 'Just Intolerable'

Time of Day	M_{dn} (IQR)	M_{dn} (IQR)	ΔM_{dn}^{NHST}	Positive	Negative	Ties	Effect Size (r)
Mid. vs. Morn.	19.20 (5.74)	15.94 (6.05)	3.26***	23	7	0	0.62
Aft. vs. Morn.	20.69 (6.56)	15.94 (6.05)	4.75***	24	6	0	0.75

Even. vs. Morn.	20.27 (6.05)	15.94 (6.05)	4.33***	24	6	0	0.72
Aft. vs. Mid.	20.69 (6.56)	19.20 (5.74)	1.49*	19	11	0	0.46
Even. vs. Mid.	20.27 (6.05)	19.20 (5.74)	1.07 n.s.	16	13	1	0.27
Even. vs. Aft.	20.27 (6.05)	20.69 (6.56)	-0.41 n.s.	12	16	2	0.14

With Bonferroni Correction: * 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

The results indicate that the median differences (ΔM_{dn}) for the pairwise comparisons are consistently positive, signalling a tendency for the values of calculated IES-GI to increase as the day progresses. This is coherent with the adoption of a one-tailed hypothesis, since a prevailing trend can be observed for all times of the day and GSV criteria. The Wilcoxon matched-pairs tests provide evidence that the differences between glare indices calculated at different times of the day are highly significant in 9 cases, significant in 1 case, weakly significant in 7 cases, and not significant in 7 out of 24 cases. The differences detected have a generally substantive (i.e., of practically relevant magnitude) effect size, mostly ranging between ‘moderate’ ($0.50\leq r<0.80$ in 12 cases out of 24) and ‘small’ ($0.20\leq r<0.50$ in 10 cases out of 24). For all GSV criteria, the data indicate that the differences in IES-GI become greater as the time interval between test sessions increases (e.g., Midday vs Morning compared to Evening vs. Morning). This observation is evident through inspection of both the difference in IES-GI median values (ΔM_{dn}) and effect sizes (r). Therefore, from statistical analysis of the data, it can be inferred that the effect of time of the day on glare sensation becomes more substantive as the period of time between test sessions increases.

Figure 3-10 and Figure 3-11 plot the logarithmic luminance (cd/m^2) of the small diffusive screen recorded at the point in which test subjects reported, respectively, a glare sensation of ‘Just Perceptible’, ‘Just Noticeable’, ‘Just Uncomfortable’, and ‘Just Intolerable’. Specifically, the graphical displays present comparisons of recorded screen luminance between the Morning and Midday (Figure 3-10), and the Morning and Evening sessions (Figure 3-11).

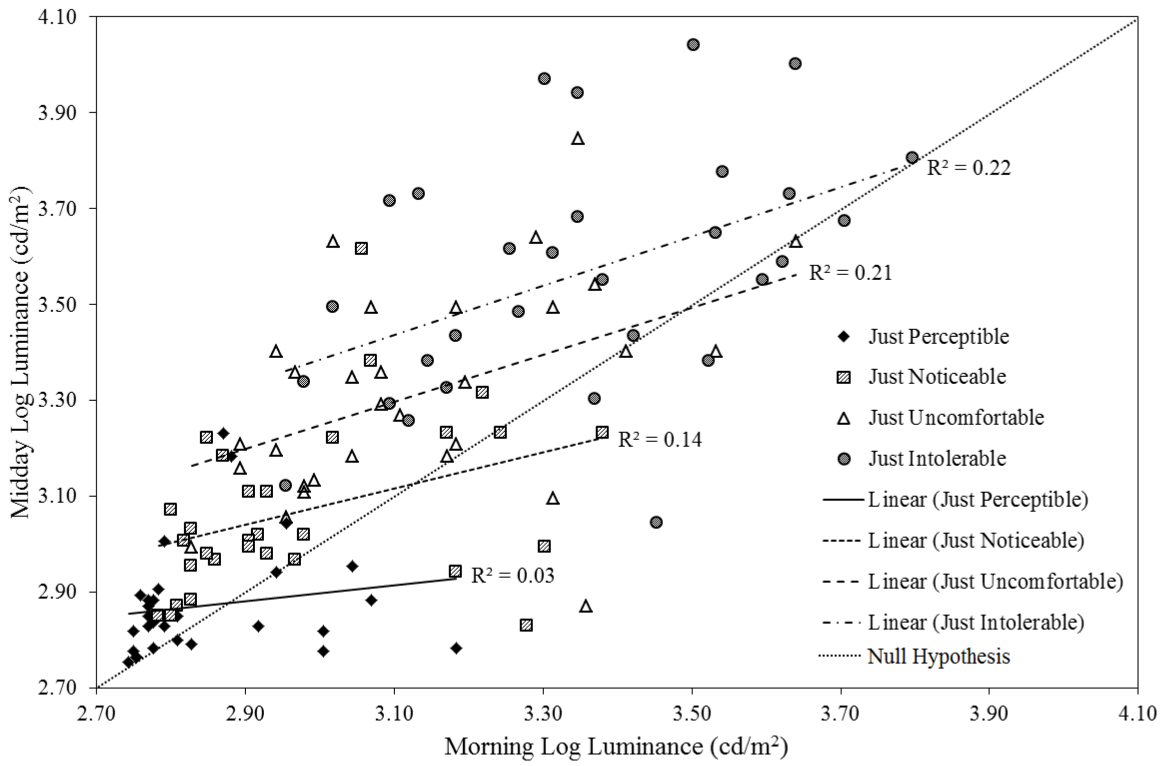


Figure 3-10 Comparison between log luminance of the diffusive screen for each GSV criterion at the Morning (x-axis) and Midday (y-axis) test sessions

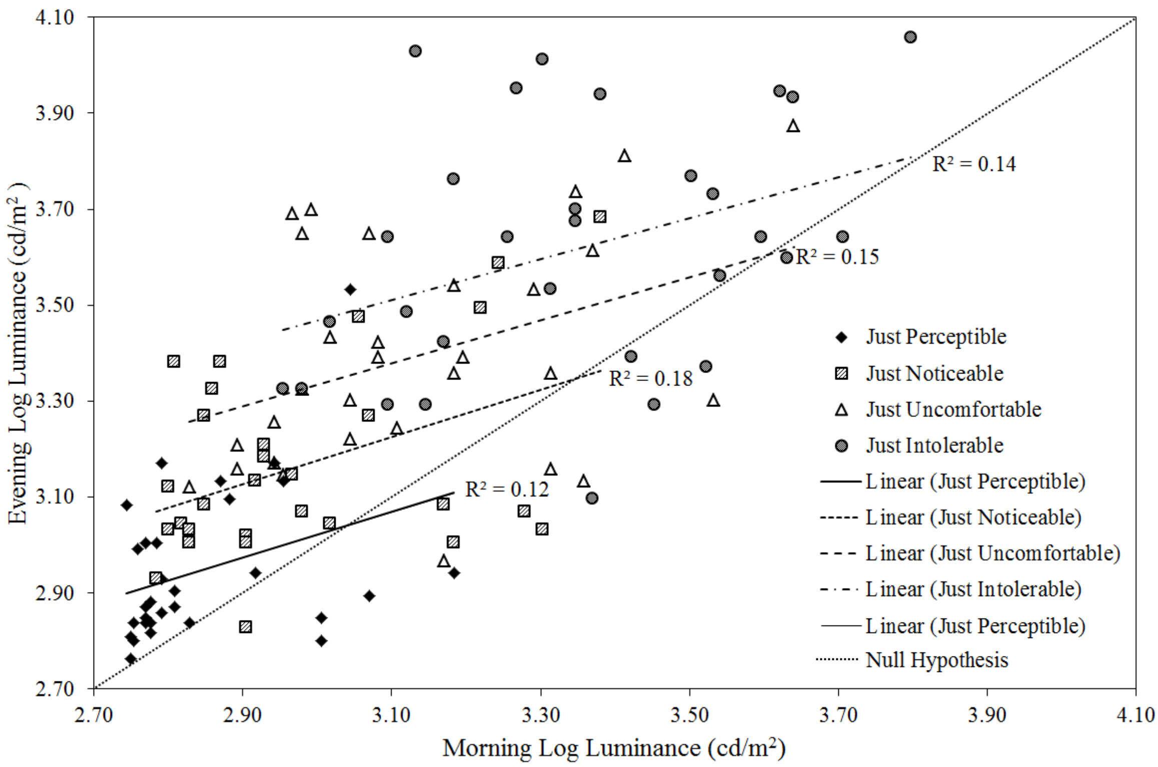


Figure 3-11 Comparison between log luminance of the diffusive screen for each GSV criterion at the Morning (x-axis) and Evening (y-axis) test sessions

In both figures, the null hypothesis line is plotted along the diagonal, representing no difference between the source luminance corresponding to each level of GSV reported by test subjects at each session. Graphical inspection of Figure 3-11 (Morning vs. Evening) reveals that, although some GSV points are located below the diagonal (indicating higher luminance values for test subjects in the morning session), the linear interpolations for each GSV are above the null hypothesis line. This confirms that test subjects were able to tolerate higher levels of source luminance in the Evening than in the Morning session for the same reported vote of glare sensation. The difference detected appears to be much larger than when examining the comparison between the Morning and Midday sessions (Figure 3-10) hence supporting the conclusions from the inferential statistical tests.

3.2. Time of the Day and Learning

3.2.1. Introduction

The first experiment primarily aimed at detection of the effect of interest, but it was not fully protected from any potential influence of learning.

As a consequence, it would be plausible to hypothesise that test subjects became more tolerant to glare with experience due to the sequence of the sessions. For example, during the Morning test, subjects could have felt more anxious about the experiment, this making them more sensitive to the luminous stimulus. This implies that the detected effect of time of day on glare sensation might have been due to the learning/experience gained throughout the experimental procedure rather than to the time intervals between test sessions as the day progressed. In this context, Hopkinson (1950) posited that a less experienced subject may not be able to interpret the meaning of glare descriptors in a consistent manner. In addition, Peterbridge & Hopkinson (1950) stated that the technique of subjective appraisal used to evaluate glare sensation is dependent on the confidence and competence of test subjects. A study by Hopkinson & Peterbridge (1954) also found that an 'experienced team' – who had previously made extensive series of observations of glare sources (N= 20) – rated discomfort glare differently when compared to the 'general population' – i.e., observers that were naïve and had no previous experiences of glare source stimuli (N= 50).

Although the pre-test procedure that each participant undertook before any data were recorded – where test subjects looked at a reference glare source, adjusting its luminance output by a dimmer – aimed at reducing the influence of experience, a follow-up experiment was designed to investigate whether the 3-hour interval between test sessions (under systematic variation) was masked from the effect of learning, and to determine if differences in IES-GI could be detected over 2 consecutive days using the same test subject.

3.2.2. Experimental Procedure

To be consistent with the first experiment, the same setting and test procedure were adopted. Each subject was requested to attend the original test sequence of four sessions equally spaced at 3-hour intervals on a chosen day, and then an identical procedure was repeated on a successive day. Test subjects could elect whether to take part to the sequence of sessions starting at 09:00 or rather at 09:30. The timing of the sessions was as follows:

- Morning: 09:00 or 09:30;
- Midday: 12:00 or 12:30;
- Afternoon: 15:00 or 15:30;
- Evening: 18:00 or 18:30.

To account for any potential influence of uncontrolled variables between the two days, fatigue, sleepiness, and metabolic stimulation (e.g., ingestion of food or caffeine) were recorded through the self-reported scales and questionnaires. The repeated-measures design also counterbalanced the influence of subjects' demographic features (e.g., gender, age, cultural background, nationality, visual acuity, etc.) (Field & Hole, 2013). On the basis of the analogous results obtained using two glare indices (IES-GI and UGR) in the first experiment, only the IES-GI was calculated as the evaluation parameter for this investigation.

The IES-GI values were calculated for each GSV reported by test subjects at four times of the day and in both days. To examine the potential influence of learning on the effect of experimental interest, for all test subjects the differences in IES-GI between sessions were calculated (i.e., Afternoon vs. Morning) on each day, and then these differences were compared across the two days.

3.2.3. Statistical Analysis

A repeated-measures analysis of covariance (ANCOVA) was used to compare the calculated differences in IES-GI over each session across both days, using the level of fatigue reported by test subjects as covariate. In fact, qualitative comments provided by participants had led to believe that there were differences in perceived levels of fatigue across the two days. The ANCOVA was adopted since, unlike the first experiment, statistical inspection of the data (K-S and S-W tests) revealed that the assumptions of normality of the sampling distributions and of homogeneity of variance (Levine's test) were not violated (Heiman, 2013). In addition, tests of homogeneity of regression slopes between the dependent variable (IES-GI) and the covariate (fatigue) – that is, the covariate has to have the same correlation with the dependent variable over the independent variable (time of the day) (Field, 2013) – showed that also this assumption was not violated.

Table 3-8 reports the results of both the K-S and S-W tests, displaying the test statistic, the degrees of freedom (df), and the statistical significance (*p*-value) for the IES-GI values calculated based on the lighting values recorded when a GSV was reported by test subjects for all times of the day.

Table 3-8 Results of the Kolmogorov-Smirnov and Shapiro-Wilk tests

GSV	Time of Day	Kolmogorov-Smirnov			Shapiro-Wilk		
		Statistic	df	*p*-value	Statistic	df	*p*-value
Just Perceptible	Morning	0.21	8	0.20 n.s.	0.91	8	0.33 n.s.
	Midday	0.24	8	0.20 n.s.	0.92	8	0.40 n.s.
	Afternoon	0.18	8	0.20 n.s.	0.90	8	0.29 n.s.
	Evening	0.30	8	0.04*	0.86	8	0.12 n.s.
Just Noticeable	Morning	0.26	8	0.21 n.s.	0.83	8	0.07 n.s.
	Midday	0.21	8	0.20 n.s.	0.88	8	0.21 n.s.
	Afternoon	0.21	8	0.20 n.s.	0.95	8	0.66 n.s.
	Evening	0.28	8	0.07 n.s.	0.82	8	0.04*
Just Uncomfortable	Morning	0.33	8	0.01**	0.67	8	0.001***
	Midday	0.15	8	0.20 n.s.	0.92	8	0.47 n.s.
	Afternoon	0.21	8	0.20 n.s.	0.90	8	0.27 n.s.
	Evening	0.25	8	0.14 n.s.	0.87	8	0.14 n.s.

Just Intolerable	Morning	0.25	8	0.14 n.s.	0.84	8	0.07 n.s.
	Midday	0.28	8	0.08 n.s.	0.83	8	0.06 n.s.
	Afternoon	0.22	8	0.20 n.s.	0.88	8	0.17 n.s.
	Evening	0.22	8	0.20 n.s.	0.84	8	0.07 n.s.

* weakly significant; ** significant; *** highly significant; n.s. not significant

The K-S tests are significant in 1 case, weakly significant in 1 case, and not significant in 14 cases out of 16. These results provide evidence that the distribution of data around the mean parameter is not statistically different from a sample population with a normal distribution. Also for the S-W tests, the results indicate that the differences are highly significant in 1 case, weakly significant in 1 case, and not significant in 14 cases out of 16. Unlike the first experiment (Section 3.1.2.2), statistical inspection of the data revealed that the assumption of normality of the sampling distributions around the mean parameter was not violated.

Table 3-9 reports the results for the Levine's test of homogeneity of variance, displaying the degrees of freedom (df), the *F* test statistics, and the statistical significance (*p*-value) for the variance associated with the IES-GI values calculated at the point in which test subjects reported GSV for each comparison between independent variables (i.e., test sessions).

Table 3-9 Levene's test of homogeneity of variance

GSV	Time of Day	df	***F***	*p*-value
Just Perceptible	Mid. vs. Morn.	1, 14	0.14	0.14 n.s.
	Aft. vs. Mid.	1, 14	0.34	0.57 n.s.
	Even. vs. Aft.	1, 14	0.07	0.80 n.s.
Just Noticeable	Mid. vs. Morn.	1, 14	2.56	0.13 n.s.
	Aft. vs. Mid.	1, 14	0.10	0.76 n.s.
	Even. vs. Aft.	1, 14	2.42	0.14 n.s.
Just Uncomfortable	Mid. vs. Morn.	1, 14	2.58	0.13 n.s.
	Aft. vs. Mid.	1, 14	2.68	0.12 n.s.
	Even. vs. Aft.	1, 14	2.40	0.14 n.s.
Just Intolerable	Mid. vs. Morn.	1, 14	1.98	0.18 n.s.
	Aft. vs. Mid.	1, 14	2.66	0.13 n.s.
	Even. vs. Aft.	1, 14	1.75	0.21 n.s.

* weakly significant; ** significant; *** highly significant; n.s. not significant

The results indicate that no statistically significant differences were detected for all 12 cases. These data show that the variances associated with each independent group were not different from each other, hence the assumption of homogeneity of variance was not violated.

Table 3-10 reports the results for the test of homogeneity of regression slopes, displaying the *F* test statistics with degrees of freedom (2, 16), the statistical significance (*p*-value), and the effect size ($p\eta^2$).

Table 3-10 Homogeneity of regression slopes

GSV	Time of Day	***F*(2,16)**	*p*-value	$p\eta^2$
Just Perceptible	Mid vs. Morn.	0.31	0.79 n.s.	0.03
	Aft. vs. Mid	0.60	0.63 n.s.	0.07
	Even. vs. Aft	0.66	0.54 n.s.	0.09
Just Noticeable	Mid vs. Morn.	0.26	0.77 n.s.	0.03
	Aft. vs. Mid	0.07	0.94 n.s.	0.01
	Even. vs. Aft.	0.07	0.94 n.s.	0.01
Just Uncomfortable	Mid vs. Morn.	0.03	0.97 n.s.	0.01
	Aft. vs. Mid	0.11	0.90 n.s.	0.02
	Even. vs. Aft.	0.03	0.97 n.s.	0.01
Just Intolerable	Mid vs. Morn.	0.25	0.78 n.s.	0.03
	Aft. vs. Mid	0.13	0.88 n.s.	0.02
	Even. vs. Aft.	0.33	0.72 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; $p\eta^2\geq 0.64$ = large

The results provide evidence that the interactions between the dependent variable (IES-GI) and the covariate (fatigue) are not statistically significant for all 12 cases (Field, 2013). To provide more rigorous support to inferences, the effect size was calculated by standardised equivalence between the observed difference and the partial eta squared ($p\eta^2$), which was extrapolated from the ANCOVA. The effects detected are small ($0.04\leq p\eta^2<0.25$) in 3 cases, but have mostly negligible ($p\eta^2<0.04$ in 9 cases out of 12) magnitude. These results suggest that any significant and practically relevant influence of violating the assumption of homogeneity of regression slopes can be excluded from the analysis.

3.2.4. Results

A total of 8 subjects volunteered to take part to this study, all varying in age, cultural background, and nationality. These participants had not taken part to the previous experiment. The study took place between the months of April and May. Figure 3-12 and Figure 3-13 present the estimated marginal means of the differences in IES-GI calculated for every level of GSV in each of the two days – adjusted for the effect of fatigue across Day 1 and Day 2 – at which subjects reported the GSV of 'Just Perceptible' (left) and 'Just Noticeable' (right) (Figure 3-12), and the criteria of 'Just Uncomfortable' (left) and 'Just Intolerable' (Figure 3-13), and their associated 95% upper and lower confidence intervals. On the x-axis, the figures present the comparisons between tests sessions – later versus earlier (e.g., Midday vs. Morning) – plotted for Day 1 and Day 2. In the displays, the variance associated with the covariate (fatigue) is removed from the analysis, allowing for a more accurate assessment of the effect of experimental interest (Field, 2013). The figures reveal statistical values (means) that consistently correspond to positive levels, indicating increased tolerance to glare as the day progresses (i.e., higher IES-GI at later sessions). Statistical values appear to be greater for Day 2, suggesting larger (positive) differences in IES-GI in the second day of the study.

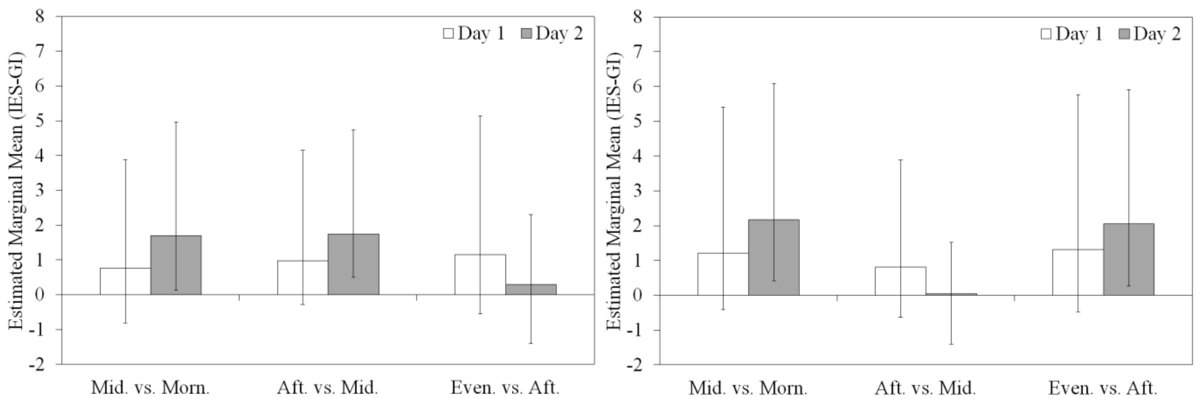


Figure 3-12 Estimated marginal means with 95% confidence intervals of the differences in IES-GI on Day 1 and Day 2 for the GSV criteria of 'Just Perceptible' (left) and 'Just Noticeable' (right)

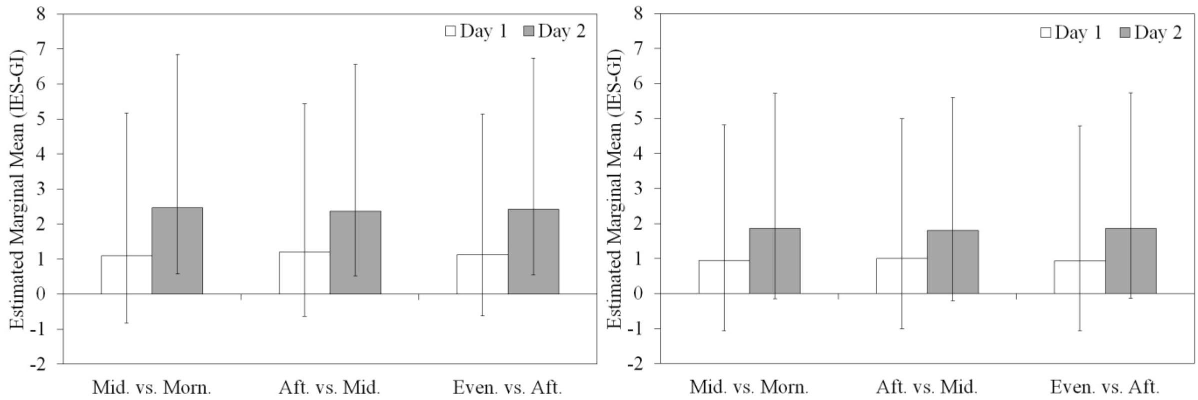


Figure 3-13 Estimated marginal means with 95% confidence intervals of the differences in IES-GI on Day 1 and Day 2 for the GSV criteria of 'Just Uncomfortable' (left) and 'Just Intolerable' (right)

Table 3-11 to Table 3-14 report, for every GSV criterion, the adjusted marginal mean (M) and standard deviation (SD) of the variations in IES-GI between the test sessions in each of the two days controlled for the effect of fatigue, the lower (CI_L) and upper (CI_U) 95% confidence intervals for the adjusted marginal mean, the difference between the estimated marginal means obtained in Day 1 and in Day 2 (ΔM), the *F* test statistic with degrees of freedom (df), the statistical significance of the difference (*p*-value), and the effect size ($p\eta^2$).

Table 3-11 ANCOVA and effect sizes for 'Just Perceptible'

Time of Day	M (SD) Day 1	CI_L, CI_U	M (SD) Day 2	CI_L, CI_U	ΔM (Day 1-2)	***F***(df)	*p*-value	$p\eta^2$
Mid. vs. Morn.	0.77 (2.34)	-0.81, 2.35	1.70 (2.14)	0.12, 3.28	-0.93	0.81(1,13)	0.39 n.s.	0.06
					Covariate	3.01(1,13)	0.11 n.s.	0.19
Aft. vs. Mid.	0.97 (1.64)	-0.28, 2.22	1.75 (0.61)	0.50, 3.00	-0.78	0.88(1, 13)	0.37 n.s.	0.06
					Covariate	1.91(1,13)	0.19 n.s.	0.13
Even. vs. Aft.	1.15 (2.08)	-0.55, 2.84	0.30 (2.21)	-1.40, 2.00	0.85	0.58(1, 13)	0.46 n.s.	0.04
					Covariate	0.05(1, 13)	0.46 n.s.	0.00

* 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 3-12 ANCOVA and effect sizes for 'Just Noticeable'

Time of Day	M (SD) Day 1	CI_L, CI_U	M (SD) Day 2	CI_L, CI_U	ΔM (Day 1-2)	***F***(df)	*p*-value	$p\eta^2$
Mid. vs. Morn.	1.22 (2.60)	-0.41, 2.97	2.17 (1.67)	0.42, 3.93	-0.95	0.69(1, 13)	0.42 n.s.	0.05
					Covariate	1.33(1,13)	0.27 n.s.	0.09

Aft. vs. Mid.	0.81 (2.32)	-0.64, 2.27	0.04 (1.76)	-1.41, 1.50	0.77 Covariate	0.64(1,13) 3.49(1,13)	0.44 n.s. 0.08 n.s.	0.05 0.21
Even. vs. Aft.	1.32 (1.50)	-0.47, 3.12	2.06 (1.58)	0.27, 3.87	-0.74 Covariate	0.39(1, 13) 0.70(1, 13)	0.54 n.s. 0.42 n.s.	0.03 0.05

* 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 3-13 ANCOVA and effect sizes for 'Just Uncomfortable'

Time of Day	M (SD) Day 1	CI_L, CI_U	M (SD) Day 2	CI_L, CI_U	ΔM (Day 1-2)	***F***(df)	*p*-value	$p\eta^2$
Mid. vs. Morn.	1.09 (3.20)	-0.82, 2.99	2.47 (2.51)	0.57, 4.38	-1.23 Covariate	1.23(1,13) 0.00(1, 13)	0.23 n.s. 0.96 n.s.	0.09 0.00
Aft. vs. Mid.	1.20 (2.88)	-0.63, 3.04	2.36 (1.07)	0.52, 4.19	-0.77 Covariate	0.90(1, 13) 1.16(1, 13)	0.36 n.s. 0.30 n.s.	0.07 0.08
Even. vs. Aft.	1.13 (1.75)	-0.62, 2.88	2.43 (1.90)	0.55, 4.30	-0.74 Covariate	1.13(1, 13) 0.53(1, 13)	0.31 n.s. 0.48 n.s.	0.08 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 3-14 ANCOVA and effect sizes for 'Just Intolerable'

Time of Day	M (SD) Day 1	CI_L, CI_U	M (SD) Day 2	CI_L, CI_U	ΔM (Day 1-2)	***F***(df)	*p*-value	$p\eta^2$
Mid. vs. Morn.	0.94 (3.45)	-1.06, 2.94	1.86 (2.19)	-0.14, 3.86	-0.92 Covariate	0.49(1, 13) 0.01(1, 13)	0.50 n.s. 0.95 n.s.	0.04 0.00
Aft. vs. Mid.	1.00 (3.89)	-1.00, 3.00	1.80 (1.65)	-0.20, 3.80	-0.80 Covariate	0.37(1, 13) 0.21(1, 13)	0.56 n.s. 0.65 n.s.	0.03 0.02
Even. vs. Aft.	0.93 (1.56)	-1.07, 2.93	1.87 (2.58)	-0.13, 3.86	-0.94 Covariate	0.51(1, 13) 0.03(1, 13)	0.49 n.s. 0.86 n.s.	0.04 0.00

* 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

Visual inspection of the descriptive statistics reveals that the means (M) consistently present positive values for all GSV criteria, signalling greater tolerance to source luminance as the day progresses for both Day 1 and Day 2. The data indicate that there is a general tendency towards larger variations of glare indices between test sessions on Day 2. This is suggested by predominantly negative mean differences (ΔM) between Day 1 and Day 2, particularly at higher levels of reported glare sensation ('Just Uncomfortable' and 'Just Intolerable').

The only positive values of mean difference (ΔM) over the test sessions that can be observed between Day 1 and Day 2 – signalling a potential influence of learning – are detected for the criterion of 'Just Perceptible' when comparing the Evening and Afternoon test sessions (ΔM= 0.85, p= 0.46 n.s., pη^2= 0.04), and for the 'Just Noticeable' GSV in the comparison between Afternoon and Midday (ΔM= 0.77, p= 0.44 n.s., pη^2= 0.05). In both cases, however, after controlling for the effect of the covariate over the independent variable (i.e., assuming that there is no difference between subjective evaluations of fatigue provided by test subjects on Day 1 and Day 2), the inferential statistics show that the differences are not statistically significant and the effect size is on the borderline of practical relevance (respectively: F(1,13)= 0.58, p= 0.46, pη^2= 0.04; and F(1,13)= 0.64, p= 0.44, pη^2= 0.05). These results lead to conclude that any significant and practically relevant influence of learning can be safely excluded from the experimental results obtained in the first experiment (Section 3.1.3).

3.2.5. Discussion: Discomfort Glare, Time of Day and Learning

The results of the initial experiment provided statistically significant and practically relevant evidence of an effect of time of the day on the level of glare sensation reported by test subjects within a controlled laboratory setting under artificial lighting conditions. Graphical displays (Figure 3-10 and Figure 3-11) also showed trends for Glare Sensation Votes (GSV) to correspond to higher levels of IES-GI as the day progresses, hence confirming the conclusions made from inspection of both descriptive (i.e., central tendencies) and inferential (i.e., effect size) statistics. Further investigation verified that the effect of time of day on reported glare sensation was not likely to be brought on by any influence of learning.

When plots of GSV were regressed, a large scattering effect was observed in the data, as confirmed by the low coefficients of determination (R^2) of linear fits (Figure 3-10 and Figure

3-11). This finding is consistent with the literature (Tregenza & Wilson, 2011; Tuaycharoen & Tregenza, 2005: 2007; Velds, 2002; Wienold & Christoffersen, 2006) and suggests that, despite the experimental design had controlled as many variables known to influence glare response as possible, there were further factors not experimentally controlled (also varying with the time of day) that might have caused individual variations in tolerance to luminance increases. Although it may not be feasible to completely eliminate the scatter commonly associated with subjective evaluations of glare sensation, individual differences may be mitigated if these are systematically identified (Rodriquez & Pattini, 2014).

For the purpose of this investigation, several temporal variables (fatigue, caffeine and food ingestion, sky condition, prior daylight exposure) and personal factors (age, gender, ethnicity, visual acuity, chronotype, photosensitivity) were also measured at the beginning of each test session. As revealed by the literature (Section 2.5.5), these temporal variables and personal factors might also be hypothesised to have an influence on glare response. By isolating potential causes of scatter within each test session as the day progresses (i.e., at a between-subject level (Field & Hole, 2013)), it can be assumed that any large differences across the variables isolated can aid in the explanation of the large scattering effect observed in the results. For this reason, the following section investigates the influence of several temporal variables and personal factors on the subjective evaluation of glare sensation at different times of the day.

3.3. Temporal Variables and Personal Factors in Glare Sensation

3.3.1. Introduction

The laboratory experiments described in Section 3.1 and Section 3.2, conducted under controlled laboratory conditions, revealed a tendency towards greater tolerance to luminance increases in artificial lighting as the day progressed. The postulated temporal effect was found to be directly linked to the independent variable (i.e., time of day); that is, as the time interval between sessions of artificial light exposure increased, so did the difference in IES-GI (and, therefore, source luminance) corresponding to reported levels of glare sensation. In the experiment, all variables known to potentially influence glare response were controlled. Nevertheless, when votes of glare sensation were compared across the times of the day (Figure 3-10 and Figure 3-11), linear regression of the data showed a low coefficient of determination (R^2), signalling a large scatter in the results. This is consistent with the wide variation in individual judgements of visual discomfort already postulated by Hopkinson (1955), and the suggestions of Osterhaus (2005) and Boyce (2014) that emphasized that the physiological and psychological mechanisms underlying glare sensation still require a more complete characterisation. This led to hypothesise that, in the experiments conducted, there could have been other variables, not controlled through experimental manipulation, which influenced glare response.

3.3.2. Method

3.3.2.1. Experimental Procedure

Data from the laboratory experiment described in Section 3.1.2.1.1were utilised in this analysis. At the beginning of the test procedure (Section 3.1.2.2), subjects were asked to fill in a short questionnaire featuring demographic information (e.g., age, gender, ethnicity, visual acuity) and requiring them to provide self-evaluations of personal factors, such as chronotype and photosensitivity. From a review of the literature, chronotype can be defined as a personal attribute that reflects individual circadian phases, providing an indication of the time of the day when physiological processes (e.g., hormonal secretion, minimum/maximum core body temperature, etc.) are triggered and preferences for periods of diurnal activity, sleep habits, etc. (Levandovski *et al.*, 2013). Since it was considered that this term may have not been fully understood by all test subjects, the Munich Chronotype Questionnaire (MCTQ) (Roenneberg *et al.*, 2003) was used to aid participants in providing informed judgements. To assess photosensitivity, subjects were required to provide self-assessments of their sensitivity to natural and artificial light (measured on a dichotomous scale (yes/no)), for example answering questions related to their use of solar protection devices (e.g., sunglasses), their luminous preferences at work (e.g., bright or dim conditions), and their frequency of interaction with environmental controls (e.g., blinds) in indoor environments (Aries, 2005). Finally, after all GSV judgements were provided at each experimental sessions, participants were asked to complete one additional questionnaire reporting their perceived level of fatigue, their caffeine and food ingestion prior to the test, the prevailing sky conditions and the light exposure (artificial or natural) that they had experienced between test sessions.

3.3.2.2. Statistical Analysis

Data from all the 30 test subjects who volunteered to take part to the experiment were used (Section 3.1.2.2.) When feasible, the data related to temporal variables and personal factors were reduced into dichotomous categories, an approach that is consistent with several other glare studies (Rodriquez & Pattini, 2014; Tuaycharoen & Tregenza, 2005; Wienold & Christoffersen, 2006). This was done to avoid making large comparisons across multiple independent groups, whilst also reducing the possibility of occurrence of Type II errors (i.e., failure to reject a false null hypothesis while examining a substantive effect (Field, 2013)).

Variables that could take on more values (e.g., fatigue and chronotype) were measured on ordered 7-point Likert scales, since it was considered that forcing these factors into discrete (dichotomous) categories may have risked underestimating their true effect on glare sensation. Since all tests were conducted under the same controlled photometric conditions (background luminance, subtended size of the source, position index) – while the luminance of the small diffusive screen placed on the line of sight of the observer was gradually increased – for all temporal variables and personal factors, the log luminance of the glare source (instead of the IES-GI) was used in this analysis as the primary parameter to evaluate individual differences associated with the subjective evaluation of glare sensation (Stone & Harker, 1973; Tuaycharoen & Tregenza, 2005: 2007; Velds, 2002; Wienold & Christoffersen, 2006).

For all temporal variables and personal factors under investigation, graphical (Q-Q plots) and statistical inspection (K-S and S-W tests) revealed that data were not normally distributed around the mean, thus violating one of the assumptions for a parametric test (Heiman, 2013). Moreover, the non-parametric Levine's test of homogeneity of variance returned non-statistically significant (p>0.05) nor practically relevant (η^2<0.04) differences between variables (Nordstokke & Zumbo, 2010; Nordstokke *et al.*, 2011).

The statistical significance of the differences in log luminance considered under dichotomous variables was evaluated using the non-parametric Mann-Whitney U test (also called Mann–Whitney–Wilcoxon or Wilcoxon rank-sum), a test suitable for comparing two independent samples. The Mann-Whitney U test accounts for non-normal distributions by ranking the data, from lowest to highest, and then summing up the totals across independent groups (Mann & Whitney, 1947). When ranking the data for the two groups, the test compares the ranks of one group to the average ranks of both groups to determine if the ranks of each of the two samples are significantly different. In so doing, the test takes the lowest sum of ranks as the test statistic U, which is then converted to a Z_{score} for the calculation of the effect size. Since the test statistic U is set to equal the smaller of the two adjusted sums of ranks (and the lowest sum of ranks will be evidently lower than the average ranks), in a one-tailed test the sign of the Z_{score} will always be negative (as will the effect size, which therefore in a uni-directional Mann-Whitney U test has to be interpreted in its absolute value) (Sawilowsky, 2007). Conversely, statistical significance of priori ordering effects with respect to ordered variables was analysed using the Jonckheere-Terpstra test, which requires independent groups divided into ranked orders so as to evaluate the strength of trends within the data (Jonckheere, 1954).

When estimating the statistical significance (*p*-value) for both the Mann-Whitney U and the Jonckheere-Terpstra tests, it was considered that the sample size was too small to base calculations on the default (asymptotic) method utilised by the SPSS statistical package, which gives an approximation of the statistical significance based on large sample groups (Field, 2013). Conversely, the Monte Carlo method was used to estimate the 'exact' statistical significance, whereas asymptotic assumptions were not met (i.e., without relying on asymptotic distributions by taking multiple samples (10,000) from the sample distribution (North, 2002)). Inferential testing was performed for all temporal variables and personal factors measured during the experimental procedure, analysing the differences and trends in the log source luminance corresponding to each criterion of reported GSV for all times of the day. Directionality of the hypothesis was informed by inspection of descriptive statistics and graphical displays of the data (Hauschke & Steinijans, 1996). However, if no consistent directionality could be determined through preliminary exploratory analysis, two-tailed hypothesis was applied (Ruxton & Neuhauser, 2010).

To support findings from NHST and provide more robust inferences, the effect size was extrapolated from both the Mann-Whitney U and the Jonckheere-Terpstra tests to evaluate if the predictor variable had any practical relevance (Cohen, 1965; Schiavon & Altomonte, 2014). The effect size was calculated by equivalence with the Pearson's r, according to the following formula (Equation 3-9) (Field, 2013; Rosenthal & DiMatteo, 2001):

$$Effect\ Size = \frac{Z_{score}}{\sqrt{N}} \qquad \text{(Equation 3-9)}$$

Where N is the number of observations, and the test statistic (Z_{score}) is extracted from the Mann-Whitney U and Jonckheere-Terpstra tests.

3.3.3. Results and Discussion

Initial analysis of the data revealed that no statistically significant differences across the independent groups for all levels of reported GSV and times of the day could be detected based on the gender, ethnicity, and photosensitivity of test subjects. Likewise, due to the tightly controlled membership to the test sample, the age of participants was regarded as masked from the experiment, while analysis of the variation of reported food ingestion revealed no statistically significant or practically relevant differences between independent groups. Hence, consideration of these factors was excluded from the subsequent analysis.

3.3.3.1. Fatigue

Figure 3-14 and Figure 3-15 present selected boxplots related to consideration of the temporal variable fatigue, for all test sessions and for GSVs of particular interest. Specifically, the figures provide boxplots for the log source luminance (y-axis) recorded at the point where test subjects reported a GSV of 'Just Perceptible' (Figure 3-14) and 'Just Intolerable' (Figure 3-15) at different times of day (x-axis). For each boxplot, fatigue is ordered in accordance to the Samn-Perelli seven-point scale, which divides the dependent variable into categories of increasing magnitude (Samn & Perelli, 1982). Table 3-15 provides the numerical values related to the Samn-Perelli ordinal scale (from 1 to 7), their associated descriptions, and a distribution of test subjects based on their self-assessed level of fatigue at each test session. For all times of the day, the 7th point of the scale has not been included in the figures since none of the test subjects reported its corresponding level of fatigue.

Table 3-15 Descriptors from the Samn-Perelli ordinal scale and distribution of test subjects based on their self-assessed fatigue

Ordinal scale	1	2	3	4	5	6	7
Description	Fully alert, wide awake	Very lively, responsive, but not at peak	Okay, somewhat fresh	A little tired, less than fresh	Moderately tired, let down	Extremely tired, very difficult to concentrate	Completely exhausted, unable to function effectively
Morning	1	5	11	8	5	0	0
Midday	8	10	7	5	0	0	0
Afternoon	5	6	7	7	4	1	0
Evening	1	9	4	7	8	1	0

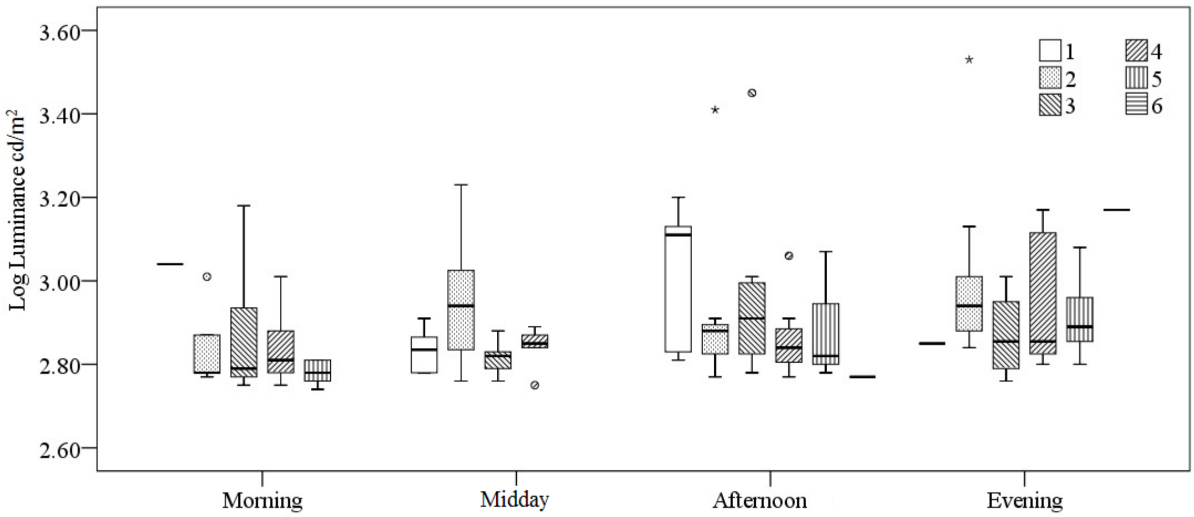


Figure 3-14 Boxplots of Log Luminance statistical parameters at each test session for the variable Fatigue and the ‘Just Perceptible’ GSV criterion

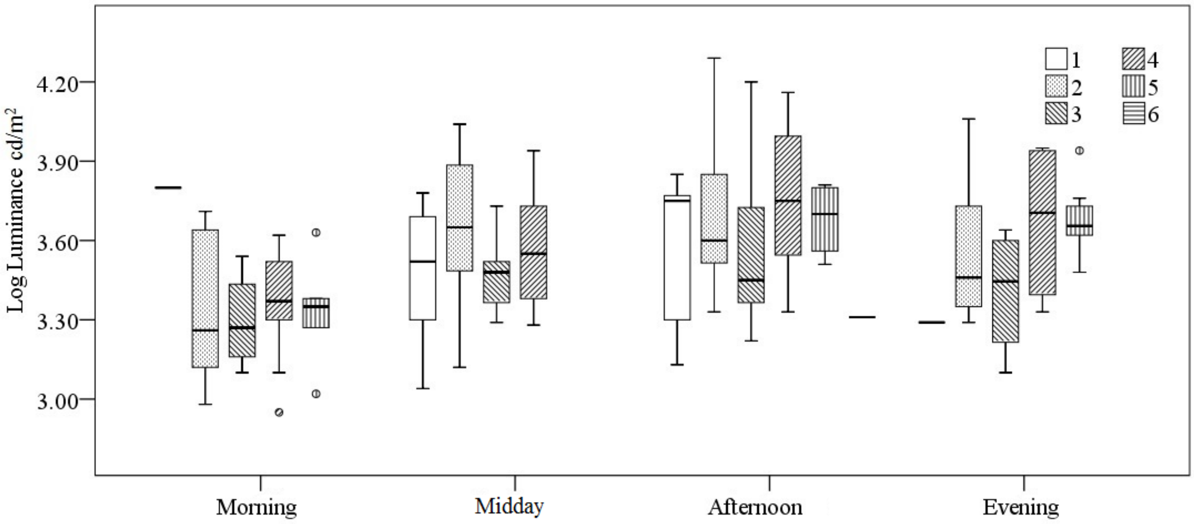


Figure 3-15 Boxplots of Log Luminance statistical parameters at each test session for the variable Fatigue and the 'Just Intolerable' GSV criterion

From graphical inspection of the data, it is difficult to make any univocal interpretation of the distributions of statistical values (minimum, 25th percentile, median, 75th percentile, and maximum), particularly for the criterion of 'Just Perceptible' (Figure 3-14), and to recognise any consistent tendency for votes of GSV to be reported by subjects at higher or lower source luminance as the magnitude of self-assessed fatigue increases on the Samn-Perelli scale.

Table 3-16 to Table 3-19 report the J-T value, the test statistic (Z_{score}), the Monte Carlo simulated lower (CI_L) and upper (CI_U) 95% confidence intervals for the statistical significance (*p*-value), and the effect size (r) for the Jonckheere-Terpstra tests performed for each level of GSV. Since initial exploratory analysis of central tendencies and graphical displays did not reveal any univocal prevailing directionality, a two-tailed alternative hypothesis supported either a direct or an inverse relationship between variables (i.e., as fatigue increases the luminance respectively increases or decreases). In the tables, the effect size provides a measure of the magnitude of these relationships, and its sign indicates whether the trend is direct (positive) or inverse (negative).

Table 3-16 Jonckheere-Terpstra test and effect sizes for 'Just Perceptible'

Time of the Day	J-T value	Test statistic	CI_L	CI_U	*p*-value (two-tailed)	Effect size (r)
Morning	131.00	-1.65	0.09	0.10	0.10 n.s.	-0.30
Midday	162.00	-0.48	0.62	0.64	0.63 n.s.	-0.09
Afternoon	136.00	-1.34	0.03	0.04	0.04*	-0.36
Evening	192.00	1.19	0.75	0.78	0.76 n.s.	0.05

*** highly significant; ** significant; * weakly 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

Table 3-17 Jonckheere-Terpstra test and effect sizes for 'Just Noticeable'

Time of the Day	J-T value	Test statistic	CI_L	CI_U	*p*-value (two-tailed)	Effect size (r)
Morning	131.50	-1.62	0.09	0.10	0.10 n.s	-0.30
Midday	164.00	-0.48	0.62	0.65	0.63 n.s.	-0.08
Afternoon	154.50	-1.34	0.17	0.19	0.18 n.s.	-0.24
Evening	217.50	1.19	0.22	0.24	0.23 n.s.	0.22

*** highly significant; ** significant; * weakly 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

Table 3-18 Jonckheere-Terpstra test and effect sizes for 'Just Uncomfortable'

Time of the Day	J-T value	Test statistic	CI_L	CI_U	*p*-value (two-tailed)	Effect size (r)
Morning	138.50	-1.37	0.17	0.18	0.17 n.s.	-0.25
Midday	166.00	-0.34	0.72	0.74	0.73 n.s.	-0.06
Afternoon	179.50	-0.47	0.62	0.64	0.63 n.s.	-0.09
Evening	236.00	1.85	0.05	0.06	0.06 n.s.	0.34

*** highly significant; ** significant; * weakly 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

Table 3-19 Jonckheere-Terpstra test and effect sizes for 'Just Intolerable'

Time of the Day	J-T value	Test statistic	CI_L	CI_U	*p*-value (two-tailed)	Effect size (r)
Morning	173.50	-0.12	0.90	0.92	0.90 n.s.	-0.02
Midday	171.00	-0.16	0.86	0.88	0.87 n.s.	-0.03
Afternoon	197.00	0.33	0.74	0.76	0.75 n.s.	0.06
Evening	245.50	2.18	0.02	0.03	0.02*	0.40

*** highly significant; ** significant; * weakly 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

The two-tailed Jonckheere-Terpstra tests show evidence of only two weakly statistically significant influences. Specifically, for the 'Just Perceptible' GSV during the Afternoon session, the results (J-T= 136.00, *p*= 0.04, r= -0.36 (small)) indicate a significant and practically relevant inverse relationship between variables, i.e. a higher tolerance to luminance levels for test subjects reporting lower levels of fatigue. Conversely, for the 'Just

Intolerable' GSV during the Evening session, a significant and substantive higher tolerance to increases in source luminance was detected for test subjects reporting higher levels of perceived fatigue (J-T= 245.50, p= 0.02, r= 0.40 (small)). Interestingly, although the detection of statistical significance was limited to only 2 cases out of 16, substantive trends in effect sizes ($0.20 \leq r < 0.50$) were detected in 8 cases. These tendencies suggest inverse relationships between variables at earlier tests sessions (negative effect sizes at the Morning, Midday and Afternoon), and a direct trend between fatigue and luminance in the later part of the day.

3.3.3.2. Chronotype

Figure 3-16 and Figure 3-17 present selected boxplots referred to consideration of chronotype for the 'Just Perceptible' (Figure 3-16) and 'Just Intolerable' (Figure 3-17) GSVs. Unlike other variables, the characteristics of chronotype depend on sleep phases (i.e., when subjects typically fall asleep and wake up within their circadian cycles) (Roenneberg *et al.*, 2003; Levandovski *et al.*, 2013). Hence, this factor can be considered an intrinsic subjective trait ('fixed-effect' (Field, 2013)) not depending on time of the day. To characterise this personal attribute, test subjects were requested to complete the MCTQ (Roenneberg *et al.*, 2003), which categorises chronotype according to preferred sleep patterns using an ordinal scale ranging from 'extremely early'= Type 0 to 'extremely late'= Type 6. Table 3-20 provides the numerical values related to the MCTQ ordinal scale and their associated descriptions.

Table 3-20 Descriptors from the Munich Chronotype Questionnaire (MCTQ) scale

Ordinal scale	0	1	2	3	4	5	6
Description	Extreme early type	Moderate early type	Slight early type	Normal type	Slight late type	Moderate late type	Extreme late type

The distribution of participants according to their self-reported chronotype was: Type 0= 0, Type 1= 5, Type 2= 11, Type 3= 3, Type 4= 4, Type 5= 7, and Type 6= 0.

Figure 3-16 and Figure 3-17 plot, on the y-axis, the log luminance (cd/m^2) of the diffusive screen at the points in which test subjects gave the GSVs of 'Just Perceptible' (Figure 3-16) and 'Just Intolerable' (Figure 3-17). On the x-axis, the figures provides the times of the day when GSVs were reported. Since no test subjects were self-assessed as extremely early (Type 0) or extremely late (Type 6), these have not been included within the graphical displays.

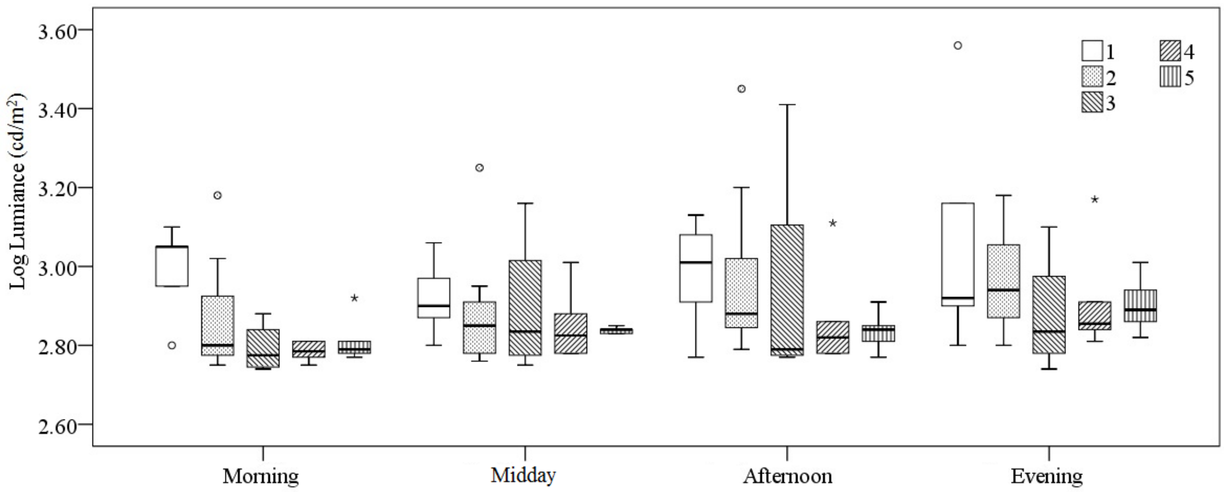


Figure 3-16 Boxplots of Log Luminance statistical parameters at each test session for the variable Chronotype and the 'Just Perceptible' GSV criterion

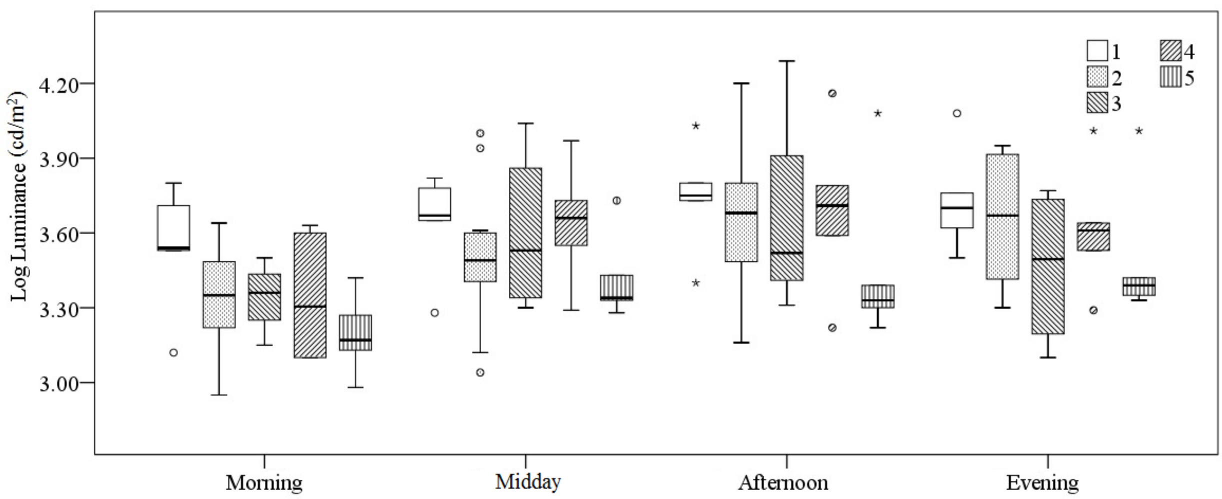


Figure 3-17 Boxplots of Log Luminance statistical parameters at each test session for the variable Chronotype and the 'Just Intolerable' GSV criterion

Graphical inspection of statistical values suggests a general prevailing tendency for both criteria of GSV to be reported at lower levels of source luminance as the chronotype of test subjects progresses from early to late type. Considering results for all levels of glare sensation votes, this led to hypothesise that test subjects self-assessed as earlier chronotypes were associated with a higher tolerance to source luminance.

Table 3-21 to Table 3-24 provide the J-T value, the test statistic (Z_{score}), the Monte-Carlo simulated lower (CI_L) and upper (CI_U) 95% confidence intervals for the statistical significance (*p*-value), and the effect size (r) for the Jonckheere-Terpstra tests at each GSV. Initial exploratory analysis of descriptive statistics and graphical displays suggested the presence of an inverse relationship between variables (i.e., earlier chronotypes tended to report GSVs at higher levels of source luminance). Thus, the alternative hypothesis examined the influence of an uni-directional (one-tailed) inverse effect, and the effect size (r) measured the magnitude of the trend detected.

Table 3-21 Jonckheere-Terpstra test and effect sizes for 'Just Perceptible'

Time of the Day	J-T value	Test statistic	CI_L	CI_U	*p*-value (one-tailed)	Effect size (r)
Morning	127.00	-2.03	0.01	0.03	0.02*	-0.37
Midday	151.50	-1.17	0.11	0.13	0.12 n.s.	-0.21
Afternoon	125.50	-2.08	0.01	0.02	0.01**	-0.38
Evening	147.50	-1.30	0.08	0.10	0.09 n.s.	-0.24

*** highly significant; ** significant; * weakly 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

Table 3-22 Jonckheere-Terpstra test and effect sizes for 'Just Noticeable'

Time of the Day	J-T value	Test statistic	CI_L	CI_U	*p*-value (one-tailed)	Effect size (r)
Morning	127.00	-2.03	0.02	0.03	0.02*	-0.37
Midday	150.00	-1.22	0.10	0.12	0.11 n.s.	-0.22
Afternoon	126.00	-2.06	0.01	0.02	0.01**	-0.37
Evening	128.00	-1.99	0.01	0.02	0.02*	-0.36

*** highly significant; ** significant; * weakly 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

Table 3-23 Jonckheere-Terpstra test and effect sizes for ‘Just Uncomfortable’

Time of the Day	J-T value	Test statistic	CI_L	CI_U	*p*-value (one-tailed)	Effect size (r)
Morning	115.50	-2.43	0.00	0.01	0.01**	-0.44
Midday	160.50	-0.85	0.18	0.20	0.20 n.s.	-0.16
Afternoon	128.00	-1.99	0.02	0.03	0.02*	-0.36
Evening	138.00	-1.64	0.04	0.06	0.05*	-0.30

*** highly significant; ** significant; * weakly 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

Table 3-24 Jonckheere-Terpstra test and effect sizes for ‘Just Intolerable’

Time of the Day	J-T value	Test statistic	CI_L	CI_U	*p*-value (one-tailed)	Effect size (r)
Morning	125.00	-2.10	0.01	0.02	0.01**	-0.38
Midday	163.00	-0.76	0.09	0.12	0.11 n.s.	-0.14
Afternoon	144.00	-1.43	0.06	0.08	0.07 n.s.	-0.26
Evening	139.00	-1.60	0.04	0.06	0.05*	-0.29

*** highly significant; ** significant; * weakly 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

Observation of the consistent negative sign of the test statistic (Z_{score}) and of the effect size (r) is coherent with the adoption of a one-tailed alternative hypothesis. The Jonckheere-Terpstra tests provide evidence that the inverse relationship between chronotype and the source log luminance is statistically significant in 4 cases and weakly significant in 6 cases out of 16. The trends demonstrate a consistently substantive (i.e., practically relevant) magnitude of the effect of chronotype for most reported levels of GSVs and test sessions (0.20≤r<0.50 in 14 cases out of 16). Inferential analysis of the data, therefore, verifies the initial exploratory postulation, providing statistically significant and practically relevant evidence to support the hypothesis that earlier chronotype test subjects were able to tolerate higher levels of source luminance for all levels of reported glare sensation and at all times of the day.

3.3.3.3. Caffeine Ingestion

Figure 3-18 and Figure 3-19 present selected boxplots related to the temporal variable caffeine ingestion. This was determined from the questionnaire administered before each test, where subjects were asked to report whether they had ('Caffeine') or had not ('No Caffeine') ingested caffeine before or between sessions. The figures plot, on the y-axis, the source log luminance (cd/m^2) at which test subjects gave the 'Just Perceptible' (Figure 3-18) and 'Just Intolerable' glare sensation votes (Figure 3-19). On the x-axis, the displays provide the times of the day along with the boxplots corresponding to the dichotomous outcome of the temporal variable: 'No Caffeine'= white, and 'Caffeine'= shaded.

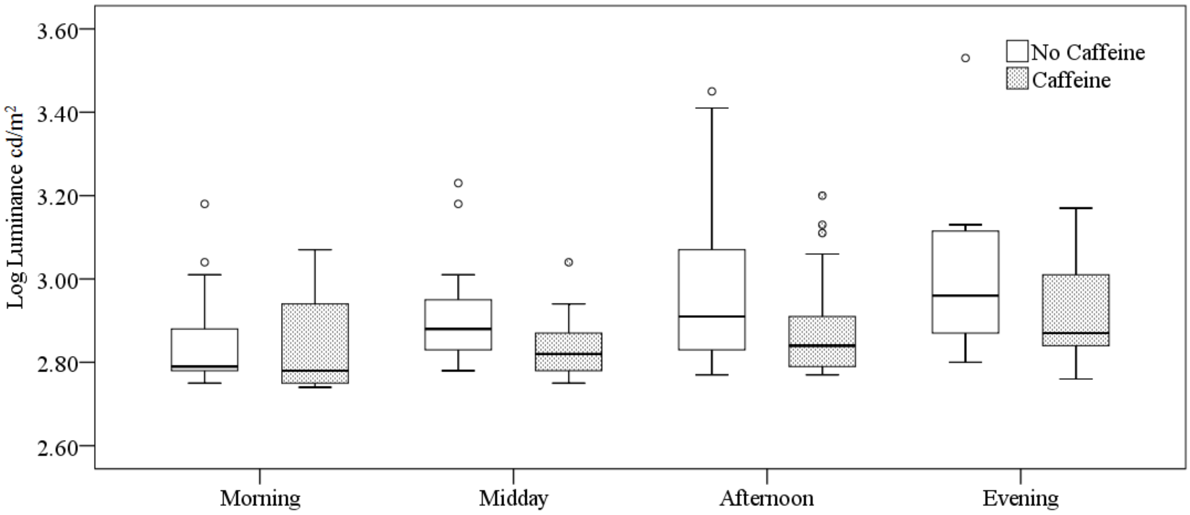


Figure 3-18 Boxplots of Log Luminance statistical parameters at each test session for the variable Caffeine Ingestion and the 'Just Perceptible' GSV criterion

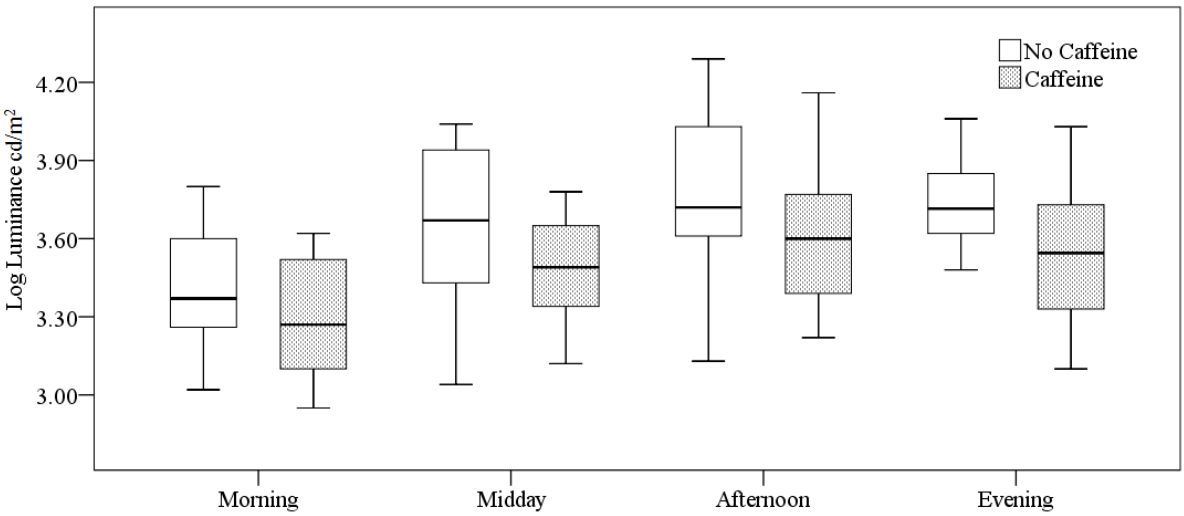


Figure 3-19 Boxplots of Log Luminance statistical parameters at each test session for the variable Caffeine Ingestion and the 'Just Intolerable' GSV criterion

Graphical inspection of the distributions of data revealed a tendency for the statistical values to correspond to higher levels of source log luminance for the 'No Caffeine' group at each reported level of glare sensation and for all times of the day.

Table 3-25 to Table 3-28 report, for every GSV, the sample size (N) of the dichotomous groups (x_0 corresponding to the 'No Caffeine' and x_1 to the 'Caffeine' group), the medians (M_{dn}) and interquartile ranges (IQR) of the log luminance at each session for both groups, the median difference (ΔM_{dn}, 'No Caffeine' vs. 'Caffeine') between the groups and its associated statistical significance (NHST, *p*-value calculated using a one-tailed hypothesis), the mean ranks for each group, the Mann-Whitney test statistic (U), and the effect size (r).

Table 3-25 Mann-Whitney U test and effect sizes for 'Just Perceptible'

Time of the day	N ($\mathbf{x}_0$, $\mathbf{x}_1$)	No caffeine M_{dn} (IQR)	Caffeine M_{dn} (IQR)	ΔM_{dn}^{NHST}	Mean rank($\mathbf{x}_o$)	Mean rank($\mathbf{x}_1$)	U	Effect size (r)
Morning	(17, 13)	2.79 (0.11)	2.76 (0.19)	0.02 n.s.	17.06	13.46	82.00	-0.20
Midday	(13, 17)	2.88 (0.13)	2.82 (0.09)	0.06**	19.46	12.47	59.00	-0.38
Afternoon	(9, 21)	2.91 (0.24)	2.84 (0.11)	0.07 n.s.	18.89	14.05	60.50	-0.26
Evening	(8, 22)	2.96 (0.23)	2.87 (0.15)	0.09 n.s.	19.25	14.14	62.00	-0.24

*** highly significant; ** significant; * weakly 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

Table 3-26 Mann-Whitney U test and effect sizes for 'Just Noticeable'

Time of the day	N (x_0, x_1)	No caffeine M_{dn} (IQR)	Caffeine M_{dn} (IQR)	ΔM_{dn}^{NHST}	Mean rank($\mathbf{x}_o$)	Mean rank($\mathbf{x}_1$)	U	Effect size (r)
Morning	(17, 13)	2.93 (0.15)	2.83 (0.36)	0.10*	18.06	12.15	67.00	-0.33
Midday	(13, 17)	3.11 (0.23)	2.99 (0.09)	0.12*	18.92	12.88	66.00	-0.34
Afternoon	(9, 21)	3.26 (0.30)	3.01 (0.35)	0.25 n.s.	18.33	14.29	69.00	-0.21
Evening	(8, 22)	3.24 (0.25)	3.07 (0.17)	0.17*	19.88	13.91	53.00	-0.30

*** highly significant; ** significant; * weakly 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

Table 3-27 Mann-Whitney U test and effect sizes for 'Just Uncomfortable'

Time of the day	N (x_0, x_1)	No caffeine M_{dn} (IQR)	Caffeine M_{dn} (IQR)	ΔM_{dn}^{NHST}	Mean rank($\mathbf{x}_o$)	Mean rank($\mathbf{x}_1$)	U	Effect size (r)
Morning	(17, 13)	3.17 (0.25)	2.98 (037)	0.19*	17.85	12.42	70.50	-0.31

Midday	(13, 17)	3.40 (0.36)	3.20 (0.23)	0.21*	19.04	12.79	64.50	-0.35
Afternoon	(9, 21)	3.51 (0.48)	3.30 (0.42)	0.21 n.s.	18.61	14.17	67.50	-0.24
Evening	(8, 22)	3.48 (0.24)	3.28 (0.40)	0.20*	20.38	13.73	49.00	-0.33

*** highly significant; ** significant; * weakly 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

Table 3-28 Mann-Whitney U test and effect sizes for ‘Just Intolerable’

Time of the day	N (x_0, x_1)	No caffeine M_{dn} (IQR)	Caffeine M_{dn} (IQR)	ΔM_{dn}^{NHST}	Mean rank($\mathbf{x}_o$)	Mean rank($\mathbf{x}_1$)	U	Effect size (r)
Morning	(17, 13)	3.37 (0.34)	3.27 (0.43)	0.10 n.s.	17.50	12.88	76.50	-0.26
Midday	(13, 17)	3.67 (0.51)	3.49 (0.31)	0.18*	18.73	13.03	68.50	-0.32
Afternoon	(9, 21)	3.73 (0.42)	3.60 (0.38)	0.13 n.s.	17.89	14.48	73.00	-0.18
Evening	(8, 22)	3.72 (0.18)	3.55 (0.39)	0.17*	19.94	13.89	52.50	-0.30

*** highly significant; ** significant; * weakly 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

Analysis of descriptive statistics reveals that the variation of statistical values across the independent variable (ΔM_{dn}, ‘No Caffeine’ versus ‘Caffeine’) is consistently positive for all GSVs and test sessions. This signals a higher tolerance to source luminance for test subjects who declared not to have ingested caffeine. This supports the directionality of the alternative hypothesis (ranks ’No Caffeine’ > ranks ‘Caffeine’) and justifies the adoption of a one-tailed test. The inferential results indicate that the differences across the independent variable are significant in 1 case and weakly significant in 8 out of 16 cases. The differences are generally substantive, with practically relevant effect sizes (0.20≤r<0.50) detected in 15 cases. In the comparison that did not display a substantive influence (‘Just Intolerable’, Afternoon session), the effect size is on the borderline of practical relevance (r= -0.18). To remind here that in a one-tailed Mann Whitney U test, the effect size has to be interpreted in its absolute value.

3.3.3.4. Sky Condition

Figure 3-20 and Figure 3-21 show selected boxplots related to the variable sky condition. This was obtained from a question where subjects were asked to self-report the prevailing sky conditions prior to or between sessions. In the boxplots, responses are binary coded as 'Clear Sky'= white and 'Overcast'= shaded. Subjects were advised to include partially cloudy skies under the 'Clear Sky' option. Reported sky conditions corresponding to sessions are plotted against the source log luminance (cd/m^2) related to the point at which test subjects gave the glare sensation votes of 'Just Perceptible' (Figure 3-20) and 'Just Intolerable' (Figure 3-21).

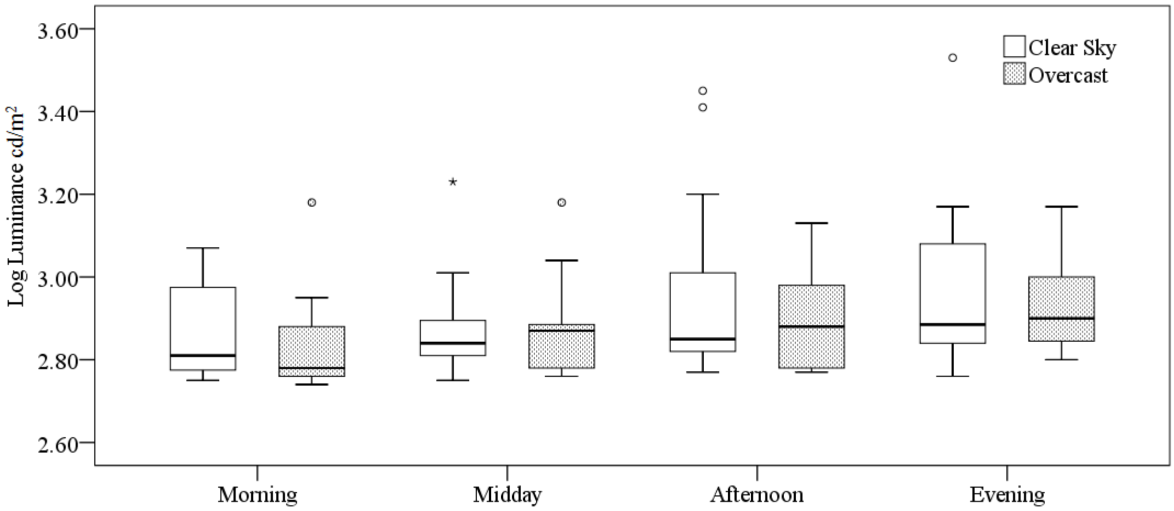


Figure 3-20 Boxplots of Log Luminance statistical parameters at each test session for the variable Sky Condition and the 'Just Perceptible' GSV criterion

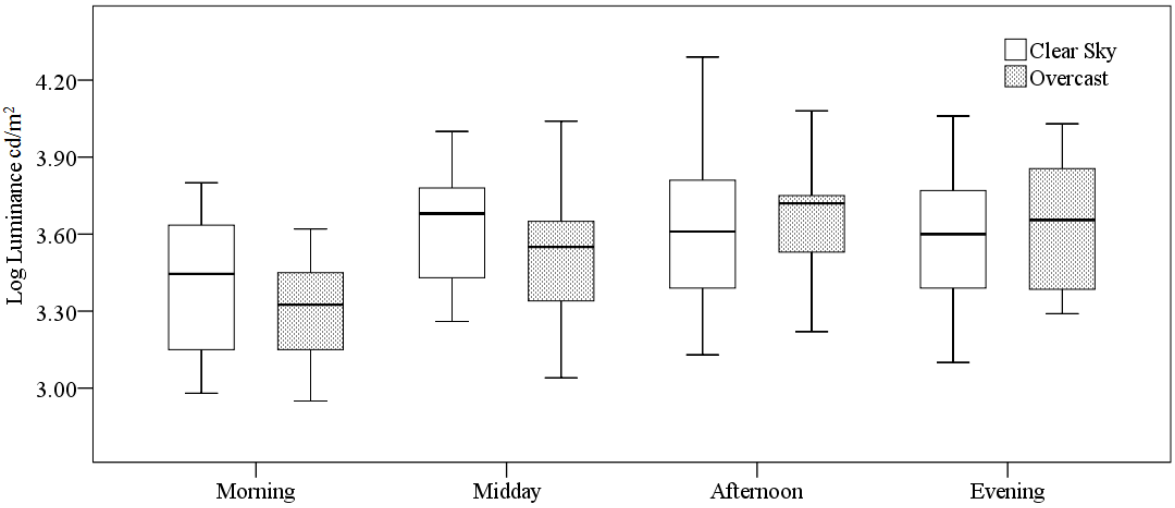


Figure 3-21 Boxplots of Log Luminance statistical parameters at each test session for the variable Sky Condition and the 'Just Intolerable' GSV criterion

Graphical inspection of the statistical values showed that, for all times of the day, no univocal interpretation of a prevailing tendency could be made across the independent variable.

Table 3-29 to Table 3-32 report the sample size (N) of the dichotomous groups (x_0 referring to the 'Clear Sky' and x_1 to the 'Overcast' group), the medians (M_{dn}) and interquartile ranges (IQR) of the log luminance, the median difference over the independent variable (ΔM_{dn}, 'Clear Sky' versus 'Overcast') and its associated statistical significance (NHST, *p*-value calculated with a two-tailed test), the mean rank for the x_0 and x_1 groups, the test statistic (U), and the effect size (r), calculated for all criteria of reported GSV and times of the day.

Table 3-29 Mann-Whitney U test and effect sizes for 'Just Perceptible'

Time of the day	N (x_0, x_1)	Clear Sky M_{dn} (IQR)	Overcast M_{dn} (IQR)	ΔM_{dn}^{NHST}	Mean rank($\mathbf{x}_o$)	Mean rank($\mathbf{x}_1$)	U	Effect size (r)
Morning	(12, 18)	2.81 (0.18)	2,78 (0.10)	0.03 n.s.	17.31	13.43	83.00	0.25
Midday	(19, 11)	2.84 (0.08)	2.87 (0.10)	-0.03 n.s.	15.37	15.73	102.00	-0.02
Afternoon	(21, 9)	2.85 (0.19)	2.88 (0.20)	-0.03 n.s.	16.00	14.33	84.00	0.09
Evening	(18, 12)	2.89 (0.23)	2.90 (0.15)	-0.01 n.s.	15.50	15.50	108.00	0.00

*** highly significant; ** significant; * weakly 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

Table 3-30 Mann-Whitney U test and effect sizes for 'Just Noticeable'

Time of the day	N (x_0, x_1)	Clear Sky M_{dn} (IQR)	Overcast M_{dn} (IQR)	ΔM_{dn}^{NHST}	Mean rank($\mathbf{x}_o$)	Mean rank($\mathbf{x}_1$)	U	Effect size (r)
Morning	(12, 18)	2.95 (0.27)	2.85 (0.14)	0.09 n.s.	17.96	13.86	78.50	0.24
Midday	(19, 11)	3.02 (0.23)	3.01 (0.23)	0.01 n.s.	15.79	15.00	99.00	0.04
Afternoon	(21, 9)	3.01 (0.33)	3.08 (0.21)	-0.08 n.s.	15.43	15.67	93.00	-0.01
Evening	(18, 12)	3.08 (0.24)	3.10 (0.24)	0.03 n.s.	14.53	16.96	90.50	-0.14

*** highly significant; ** significant; * weakly 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

Table 3-31 Mann-Whitney U test and effect sizes for 'Just Uncomfortable'

Time of the day	N (x_0, x_1)	Clear Sky M_{dn} (IQR)	Overcast M_{dn} (IQR)	ΔM_{dn}^{NHST}	Mean rank($\mathbf{x}_o$)	Mean rank($\mathbf{x}_1$)	U	Effect size (r)
Morning	(12, 18)	3.18 (0.25)	3.01 (0.14)	0.18 n.s.	18.00	13.83	73.00	0.24
Midday	(19, 11)	3.34 (0.28)	3.21 (0.35)	0.13 n.s.	17.12	14.26	89.50	0.16
Afternoon	(21, 9)	3.27 (0.43)	3.48 (0.30)	-0.22 n.s.	14.90	16.89	82.00	-0.10

Evening	(18, 12)	3.33 (0.29)	3.40 (0.35)	-0.07 n.s.	14.94	16.33	98.00	-0.08

*** highly significant; ** significant; * weakly 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

Table 3-32 Mann-Whitney U test and effect sizes for 'Just Intolerable'

Time of the day	N (x_0, x_1)	Clear Sky M_{dn} (IQR)	Overcast M_{dn} (IQR)	ΔM_{dn}^{NHST}	Mean rank($\mathbf{x}_o$)	Mean rank($\mathbf{x}_1$)	U	Effect size (r)
Morning	(12, 18)	3.45 (0.39)	3.29 (0.28)	0.16 n.s.	17.79	13.97	80.50	0.22
Midday	(19, 11)	3.59 (0.32)	3.49 (0.33)	0.09 n.s.	17.77	13.76	81.00	0.23
Afternoon	(21, 9)	3.61 (0.42)	3.72 (0.22)	-0.11 n.s.	15.24	16.11	89.00	-0.05
Evening	(18, 12)	3.60 (0.35)	3.66 (0.37)	-0.06 n.s.	14.94	16.33	98.00	-0.08

*** highly significant; ** significant; * weakly 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

The two-tailed Mann-Whitney U tests did not detect statistically significant differences across any of the 16 cases under consideration. However, the differences over the independent variable are practically relevant (small effect size, 0.20≤r<0.50) in 5 out of 16 cases. In particular, the Morning session corresponds to substantive effect sizes for differences between independent groups across all criteria of GSV. In fact, for the Morning session: 'Just Perceptible': ΔM_{dn}= 0.03, p= 0.19, r= 0.25; 'Just Noticeable': ΔM_{dn}= 0.09, p= 0.19, r= 0.24; 'Just Uncomfortable': ΔM_{dn}= 0.18, p= 0.19, r= 0.24; and, 'Just Intolerable'; ΔM_{dn}= 0.16, p= 0.23, r= 0.22. In all these cases, the median differences (ΔM_{dn}) and effect sizes correspond to positive values, suggesting higher tolerance to source luminance for test subjects that had reported clear sky conditions before the start of the test sessions.

3.3.3.5. Prior Light Exposure

Figure 3-22 and Figure 3-23 present selected boxplots referred to consideration of prior light exposure for the 'Just Perceptible' (Figure 3-22) and 'Just Intolerable' (Figure 3-23) GSVs. To measure this temporal variable, subjects were requested to provide information on the lighting conditions they had experienced before or between test sessions, for example reporting whether their luminous environment was predominantly artificially or naturally lit, or whether they had sat next to a window with access to daylight. Prior light exposure is displayed in dichotomous groups: ‘Artificial Light’= white and ‘Daylight’= shaded.

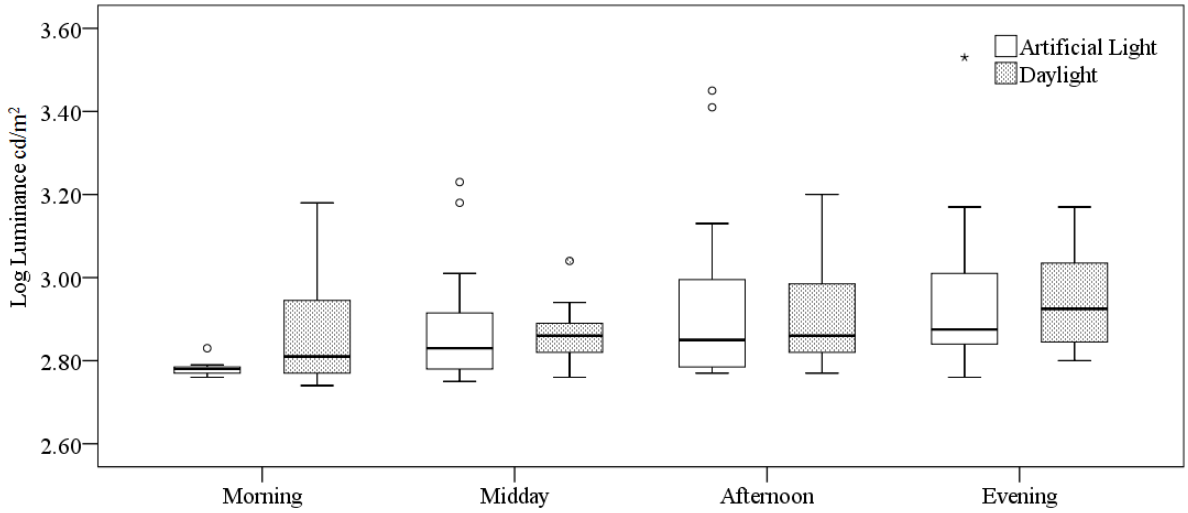


Figure 3-22 Boxplots of Log Luminance statistical parameters at each test session for the variable Prior Light Exposure and the ‘Just Perceptible’ GSV criterion

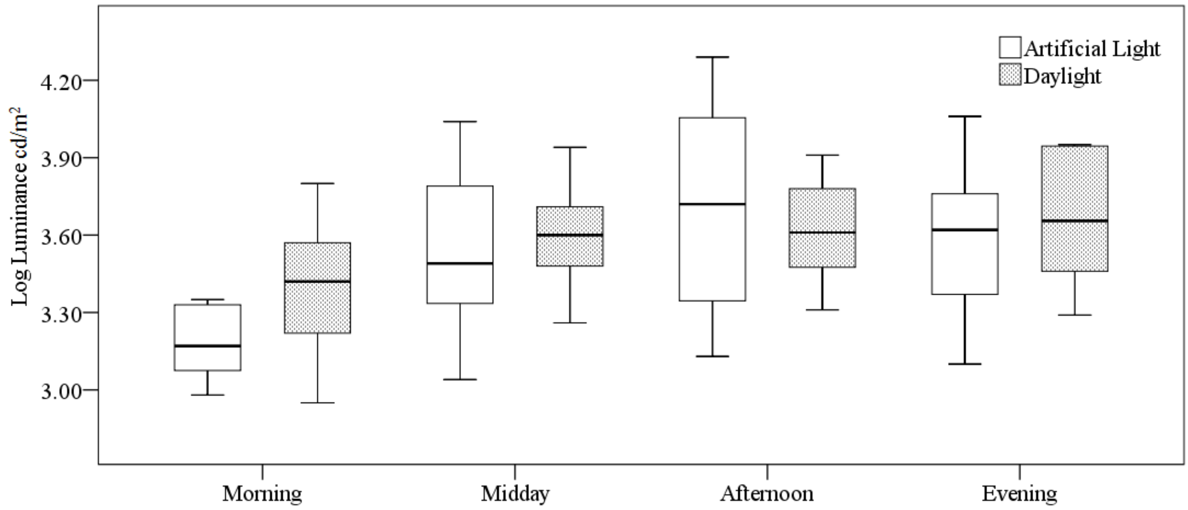


Figure 3-23 Boxplots of Log Luminance statistical parameters at each test session for the variable Prior Light Exposure and the ‘Just Intolerable’ GSV criterion

Graphical inspection of statistical values across all GSVs did not lead to the identification of a convincing univocal prevailing tendency, although the distribution of data presented in Figure 3-22 and Figure 3-23 seems to suggest that, for several test sessions, subjects who had previously been exposed to predominantly daylit conditions reported votes of glare sensation at higher levels of log luminance.

Table 3-33 to Table 3-36 present, for all GSVs, the sample size (N) of the dichotomous groups (x_0 corresponding to the 'Artificial Light' and x_1 to the 'Daylight' group), the medians (M_{dn}) and interquartile ranges (IQR) of the log luminance at each session, the median difference (ΔM_{dn}, 'Artificial Light' vs. 'Daylight') across the independent variable and its associated statistical significance (NHST, *p*-value calculated with a two-tailed Mann-Whitney U test), the mean ranks for the x_0 and x_1 groups, the test statistic (U), and the effect size (r).

Table 3-33 Mann-Whitney U test and effect sizes for 'Just Perceptible'

Time of the day	N (x_0, x_1)	Artificial M_{dn} (IQR)	Daylight M_{dn} (IQR)	ΔM_{dn}^{NHST}	Mean rank($\mathbf{x_o}$)	Mean rank($\mathbf{x_1}$)	U	Effect size (r)
Morning	(7, 23)	2.78 (0.01)	2.81 (0.18)	-0.03 n.s.	11.71	16.65	54.00	-0.26
Midday	(16, 14)	2.83 (0.12)	2.86 (0.07)	-0.03 n.s.	14.03	17.18	88.50	-0.19
Afternoon	(15, 15)	2.85 (0.20)	2.86 (0.16)	-0.01 n.s.	14.73	16.27	101.00	-0.07
Evening	(22, 8)	2.88 (0.17)	2.92 (0.17)	-0.05 n.s.	14.95	17.00	76.00	-0.06

*** highly significant; ** significant; * weakly significant; n.s. = not significant
$r<0.20$ = negligible; $0.20\leq r<0.50$ = small; $0.50\leq r<0.80$ = moderate; $r\geq0.80$ = large

Table 3-34 Mann-Whitney U test and effect sizes for 'Just Noticeable'

Time of the day	N (x_0, x_1)	Artificial M_{dn} (IQR)	Daylight M_{dn} (IQR)	ΔM_{dn}^{NHST}	Mean rank($\mathbf{x_o}$)	Mean rank($\mathbf{x_1}$)	U	Effect size (r)
Morning	(7, 23)	2.86 (0.07)	2.97 (0.34)	-0.11 n.s.	10.29	17.09	44.00	-0.30
Midday	(16, 14)	2.99 (0.27)	3.05 (0.21)	-0.06 n.s.	13.91	17.32	86.50	-0.16
Afternoon	(15, 15)	3.06 (0.37)	3.08 (0.28)	-0.03 n.s.	14.93	16.07	104.00	-0.06
Evening	(22, 8)	3.08 (0.23)	3.13 (0.23)	-0.06 n.s.	14.66	17.81	69.50	-0.16

*** highly significant; ** significant; * weakly significant; n.s. = not significant
$r<0.20$ = negligible; $0.20\leq r<0.50$ = small; $0.50\leq r<0.80$ = moderate; $r\geq0.80$ = large

Table 3-35 Mann-Whitney U test and effect sizes for ‘Just Uncomfortable’

Time of the day	N (x_0, x_1)	Artificial M_{dn} (IQR)	Daylight M_{dn} (IQR)	ΔM_{dn}^{NHST}	Mean rank($\mathbf{x}_o$)	Mean rank($\mathbf{x}_1$)	U	Effect size (r)
Morning	(7, 23)	3.02 (0.11)	3.18 (0.34)	-0.17*	9.07	17.46	35.50	-0.40
Midday	(16, 14)	3.24 (0.37)	3.34 (0.20)	-0.10 n.s.	14.69	16.43	99.00	-0.11
Afternoon	(15, 15)	3.51 (0.54)	3.30 (0.35)	0.21 n.s.	16.20	14.80	102.00	0.08
Evening	(22, 8)	3.34 (0.37)	3.41 (0.37)	-0.07 n.s.	14.43	18.44	64.50	-0.20

*** highly significant; ** significant; * weakly 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

Table 3-36 Mann-Whitney U test and effect sizes for ‘Just Intolerable’

Time of the day	N (x_0, x_1)	Artificial M_{dn} (IQR)	Daylight M_{dn} (IQR)	ΔM_{dn}^{NHST}	Mean rank($\mathbf{x}_o$)	Mean rank($\mathbf{x}_1$)	U	Effect size (r)
Morning	(7, 23)	3.17 (0.26)	3.42 (0.35)	-0.25*	9.29	17.39	37.00	-0.39
Midday	(16, 14)	3.49 (0.45)	3.60 (0.21)	-0.10 n.s.	15.19	15.86	107.00	-0.04
Afternoon	(15, 15)	3.72 (0.71)	3.61 (0.30)	0.11 n.s.	15.83	15.17	107.50	0.04
Evening	(22, 8)	3.62 (0.37)	3.66 (0.44)	-0.04 n.s.	14.93	17.06	75.50	-0.11

*** highly significant; ** significant; * weakly 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

Analysis of descriptive statistics reveals that the median differences (ΔM_{dn}, ‘Artificial Light’ vs. ‘Daylight’) are predominantly negative, signalling that test subjects who self-reported to have been exposed to daylight before or between sessions gave the GSVs at a higher source luminance across the various times of the day. The only exceptions were detected in the Afternoon for the ‘Just Uncomfortable’ (ΔM_{dn}= 0.21) and ‘Just Intolerable’ (ΔM_{dn}= 0.11) GSV criteria. However, the inferential results show that the differences are (weakly) significant only in 2 cases out of 16, both corresponding to the Morning session for the two highest levels of glare sensation vote (Table 3-35 and Table 3-36). The differences across the independent variable detected 5 practically relevant effect sizes (0.20≤r<0.50). In all these cases, both the median differences (ΔM_{dn}) and the effect sizes suggest an inverse relationship between variables, indicating a higher tolerance to source luminance for test subjects who had been previously exposed to daylight. Similar to the results obtained for sky condition, the Morning test session consistently corresponds to practically relevant effect sizes, although for the GSV criteria of 'Just Perceptible' and 'Just Noticeable' no statistical significance was detected (respectively: p= 0.23, r= -0.26; and, p= 0.08, r= -0.30).

3.3.4. Discussion

In interpreting the results of the graphical, descriptive, and inferential statistics, it should be considered that the analysis of the data was strictly limited to between-subject tests (Jonckheere-Terpstra and Mann-Whitney U). These methods of analysis are less likely to detect statistical significance than repeated-measures tests (e.g., Friedman's ANOVA and Wilcoxon matched-pairs). This is generally due to the individual differences between subjects that can be controlled (albeit, not totally eliminated) by homogenous samples (Field & Hole, 2013). Conversely, practical relevance (effect size) is not dependent on sampling, although it is not always a direct reflection of statistical significance (Rosenthal & Rosnow, 2000). In fact, it is possible for a difference across the independent variable under examination to be statistically significant (hence, leading to rejection of the null hypothesis) but trivial in its magnitude ($r<0.20$). On the other hand, a result may be substantive ($r\geq0.20$) even if outside the conventional boundaries of statistical significance ($p>0.05$) (Cohen, 1994; Ellis, 2010). Since the effect size provides a standardised measure of the magnitude of the difference between sample groups, it can be considered as a credible evaluation parameter of the scatter that was previously observed (Figure 3-10 and Figure 3-11) when the source luminance at the point when test subjects reported a GSV was regressed over the time of the day.

Figure 3-24 provides a visualisation of the influences (effect sizes) of each temporal variable and personal factor (right) on the reported levels of GSV (left) at all times of the day (x-axis).

It is important to note that the sign of the effect size is retained in this figure to account for the two-tailed directionality of the effects in the cases of the variables fatigue, sky condition, and prior light exposure. Conversely, for chronotype and caffeine ingestion, the consistently negative sign of the effect sizes substantiates the application of a one-tailed alternative hypothesis, as resulting from the initial graphical inspection of statistical parameters and supported by the analysis of descriptive and inferential statistics.

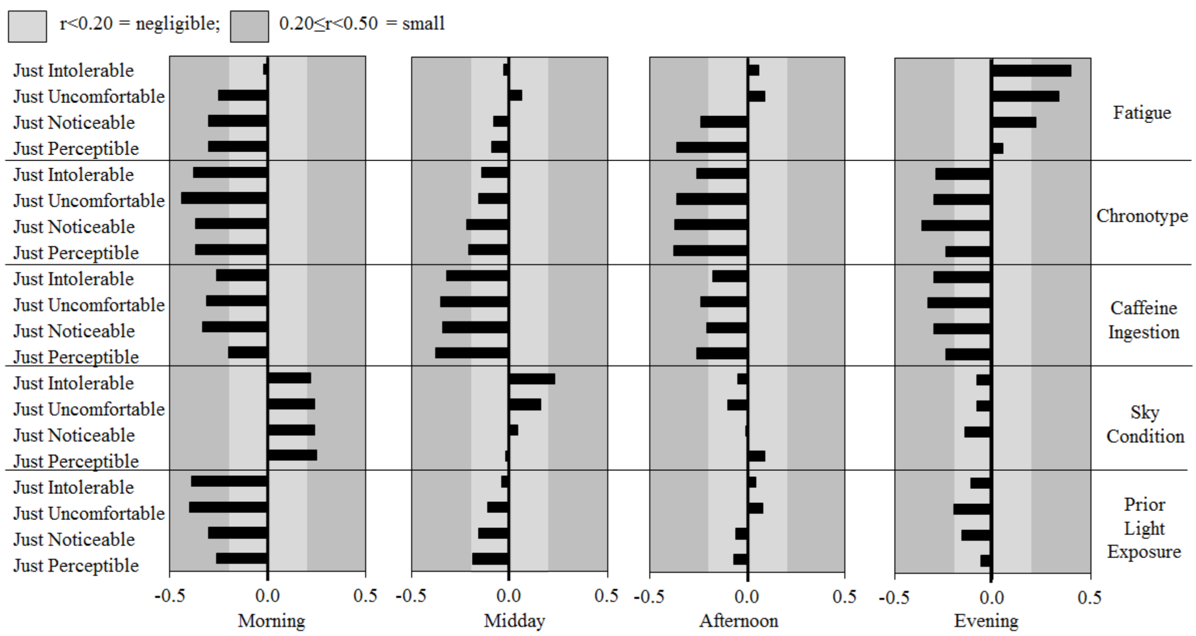


Figure 3-24 Effect sizes (r) detected in the inferential tests for all the GSV criteria (y-axis, left), personal factors and temporal variables (y-axis, right), and test sessions (x-axis)

In Figure 3-24, the temporal trends (represented by the magnitude and directionality of the effects across the time of the day) for both fatigue and chronotype follow a pattern similar to that found for other proxies, such as cognitive performance and body temperature. This is typically exemplified by the ‘post-lunch dip’ that is often apparent from 12:00 to 14:00, and is characterised by a drop in individuals’ performance (Monk, 2005; van Bommel & van den Beld, 2004). In Figure 3-24, this can be observed by the evident decrease in the effect sizes for both fatigue and chronotype – across all levels of glare sensation – at Midday (12:00 or 12:30) with respect to the Morning (previous) and Afternoon (following) sessions.

In consideration of chronotype, the inferential analysis provided statistically significant and practically relevant evidence that earlier chronotypes were able to tolerate higher levels of source luminance at all reported levels of GSV and times of day. Inspection of the variation of effect sizes across the time of the day indicates that the influence of chronotype on perceived levels of glare sensation reduces around Midday, before increasing again in the Afternoon with the same directionality of the effect (that is, earlier chronotypes seem to tolerate higher levels of source luminance for the same reported level of glare sensation).

The effects related to the temporal variable caffeine ingestion are relatively consistent across the day, its negative sign suggesting that test subjects featured in the 'No Caffeine' group were invariably more tolerant to source luminance at all levels of GSV. However, the display indicates that this influence appears to be less substantive during the Afternoon session. In fact, at this time (15:00 or 15:30), the practical relevance of the effect of caffeine ingestion over source luminance for all GSV criteria reduces with respect to other times of the day.

Finally, both sky condition and prior light exposure present practically relevant effect sizes ($r \geq 0.20$) almost exclusively in the Morning session. From a review of the literature (Section 3.1.3), there is evidence to suggest that early morning daylight exposure from 7 am to 8 am is important to induce alerting effects (Okamoto *et al.*, 2014). Moreover, bright light exposure in the first hours of the day is known to contribute to melatonin suppression (van Bommel, 2006). The results collected seem to suggest that, given the ranges of the independent variable under consideration (time of day), the early morning would likely be the period where access to natural light (which, likely, is strongly correlated to prevailing sky conditions) would be most influential on the variations in reported levels of glare sensation.

Reflecting on the findings from Sections 3.1.3, 3.2.4, and 3.3.3, it must be considered that all results were obtained under experimental conditions in which test subjects were observing an abstract artificial lighting glare source (small diffusive screen), whose luminance was strictly controlled by the experimenter. In this context, a study by Fotios (2010) found that, when test subjects were presented with different ranges of the same stimulus, the mean preferred illuminance had a tendency to lie in the middle of the range under examination. Since illuminance preferences varied depending on the available range determined by the experimental procedure, these findings suggested an effect of stimulus range bias. Coherent with these results, it might be plausible that the outcomes illustrated in from Sections 3.1.3, 3.2.4, and 3.3.3 could also be related to a stimulus range bias. That is, the reported levels of glare sensation associated with the incremental increases in screen brightness might have been influenced by the available range in luminance values presented to test subjects.

In addition, Osterhaus (2005) argued that little attention has been dedicated to collecting evaluations of glare sensation while an observer is presented with a visual task. Regardless of physical and photometric parameters, in fact, the levels of visual discomfort experienced might dramatically change if the sight of an observer focuses, for example, on a computer display or rather is not engaged in any visual task.

In response to the literature, a further laboratory experiment was designed to explore the potential relationships between visual tasks of various difficulties, influences of temporal variables, and glare responses as the day progresses. To fully counterbalance the effect of systematic confounding variables (Field & Hole, 2013), and protect from any influence of learning or experience, the new experiment was designed embracing a randomised order of presentation of the independent variables. This is illustrated in the following section.

3.4. Visual Task Difficulty and Temporal Influences in Glare Response

3.4.1. Introduction

The literature often refers to glare as 'unwanted light' – similarly to what noise is to sound (Tregenza & Wilson, 2011) – that causes overstimulation of the retina. The sensation of discomfort glare has also been defined as a form of distraction that hinders the observer's attention, vigilance, and communication, while performing a visual task (Lynes, 1974: 1977).

The laboratory experiments described in Section 3.1 and Section 3.3 analysed glare sensation along the day by requesting the subjects to look directly at the glare source rather than focusing on a visual stimulus. Therefore, to further the exploration of the effect of interest to this investigation, on the basis of the literature, a new experimental procedure was developed including the presence of a visual task, so that its impact on the reported level of glare sensation as the day progresses could be studied, along with the effects of temporal variables.

3.4.2. Method

3.4.2.1. Experimental Setup

In the context of this analysis, a systematic experimental design approach was adopted to respond to three aims:

- The first aim consisted in searching for temporal variations in the perceived level of glare sensation when subjects performed visual tasks of various difficulties at different times of the day. In this case, no comparisons between responses across visual tasks were conducted but, individually for each visual task, differences between the glare sensation reported by test subjects at each time of the day were analysed.
- The second aim of the experiment consisted in comparing the temporal differences observed in glare response across groups of visual tasks at various times of the day.

Therefore, the influence of task manipulation and difficulty over the effect of time of the day on reported visual discomfort was directly studied.

- The third aim intended to analyse the effect of several temporal variables (fatigue, food intake, caffeine ingestion, mood, prior daylight exposure, sky condition,) on the glare response provided by subjects while engaging with visual tasks.

In relation to the general research hypothesis of this investigation – i.e., whether an effect of time of day can be detected on the reported level of glare sensation in the presence of daylight from a window – this experiment essentially aimed at substantiating or challenging the previously detected variation in visual discomfort at different times of day (Section 3.1 and Section 3.1) and the influence of temporal variables on glare sensation as the day progresses (Section 3.1), within a procedure where subjects were presented with a visual task.

In addition, it was expected that the results from this experiment could provide evidence to infer if visual task difficulty is a variable potentially linked to the scatter associated with subjective evaluations of glare sensation (Tregenza & Wilson, 2011). Detection of such influence would give an indication as to whether task difficulty (other than the variables explored in Section 3.3) needs to be controlled in the subsequent stage of this investigation, when the effect of time of day on glare sensation is explored in the presence of daylight.

The setup from the previous experiments was retained to analyse the testable hypotheses. The design of the apparatus (Figure 3-25) and of the experimental procedure was informed by the studies conducted by Tuaycharoen & Tregenza (2005) and by Flannagan *et al.* (1990) to investigate the effect, respectively, of view interest and task difficulty on perceived discomfort glare. The additional stimulus used in the new experimental design was constituted by the presence of visual tasks of various difficulty that subjects were requested to perform at different times of the day while providing their rating of glare sensation.

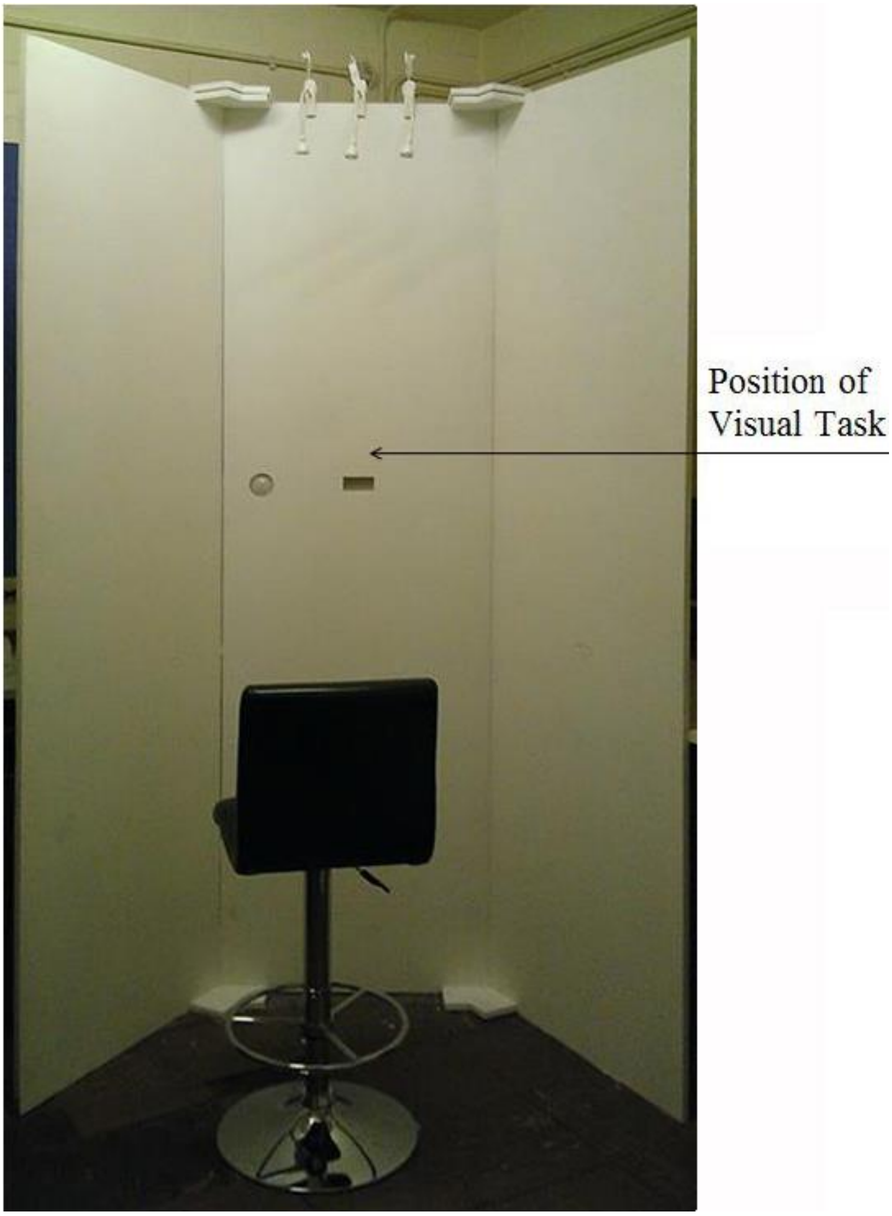


Figure 3-25 Internal view of the experimental lighting chamber and position of the visual task

3.4.2.2. Visual Task Difficulty

Based on the literature, 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 (2014), increased task difficulty can cause excessive pressure on the lens of the eye as it seeks to intensify its optical power and keep the retinal image sharp, therefore fostering muscular fatigue and possibly triggering enhanced symptoms of visual discomfort.

The measureable dependencies of visual task difficulty were expressed in this experiment by the size and contrast of the characters contained within the stimulus (Boyce, 1974; Tregenza & Loe, 2014; Tregenza & Wilson, 2011). A selection of numerical-based stimuli was adopted in the visual tasks (Clear & Berman, 1989; Rea, 1981; Slater *et al.*, 1983), since it was considered that abstract values held no real meaning for test subjects to interpret. This helped to mask the experiment from confounding variables (e.g., learning, interest, etc.) as compared with reading-based tasks.

A total of 12 visual tasks were used in this study, divided into three groups of four tasks each: 1) font size; 2) internal contrast; and, 3) character contrast (Table 3-37). The font of the characters was always set at Verdana (Rodriguez & Pattini, 2014).

For the first group, Task 1 to 4 varied the *font size*, which was set respectively at 10 (Task 1), 8 (Task 2), 6 (Task 3), and 4 (Task 4) points. In the second group, Task 5 to 8 varied the *internal contrast* (Tregenza & Loe, 2014; Boyce 1974), saturating the background of the visual stimulus (which contained the numerical information) by 20% (Task 5) and 40% (Task 6), and distorting the task information utilising vertical (Task 7) and horizontal (Task 8) lines (Markus, 1967). In the third group, Task 9 to 12 directly manipulated the information contained within each visual stimulus varying 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 presented on a strip of matt white paper (Slater & Boyce, 1990), with dimensions of 13.7 cm x 1.7 cm, subtending a solid angle at the eye of 0.006 steradians. This was purposely designed so that subjects were able to fully accommodate the visual task, reducing the likelihood of their changing head position while completing the test. Numerical characters were arranged in 1 row and 9 columns, and were spaced at an equal interval of 1.3cm from each other, so as to maximise the clarity of the task.

Table 3-37 Task number and visual tasks

Task number	Visual task
1	4 6 1 7 9 8 2 5 3
2	6 9 3 4 5 1 2 8 7
3	9 2 3 7 5 6 1 4 8
4	2 7 1 9 4 6 8 3 5
5	6 7 5 4 9 1 8 3 2
6	6 9 4 7 3 8 1 5 2
7	4 9 5 6 8 3 2 1 7
8	5 7 3 6 2 4 9 1 8
9	6 5 9 1 7 8 3 2 4
10	7 2 9 3 4 8 1 6 5
11	3 1 2 6 9 8 7 5 4
12	3 9 5 4 1 7 6 2 8

3.4.2.3. Pilot Tests

To refine the experimental setup, a series of pilot tests (N= 3) was run to determine an appropriate position for the visual task and to set a constant luminance for the diffusive screen, under which the experimental procedure (described in Section 2.2.2.4) would be conducted. Research by Iwata & Tokura (1997) 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. A study by Kim *et al.* (2009) recorded luminance values at the point in which subjects reported the BCD (Luckiesh & Guth, 1949) – discussed in Section 2.2.2 – when a glare source was placed at various positions in the visual field. The observers were instructed to look at centre of a screen in which a luminance source (0.0011 sr in size) was randomly located at designated 36 positions separated by 10^{o} displacement angles. For each location of the glare source, the luminance was increased and observers were requested to report glare sensation when discomfort was reached. Findings indicated that subjects were more sensitive to deviations in glare source position at angular displacements corresponding to 10^{o} from the line of sight (i.e., the centre of the screen). These results suggest that subjects were less tolerant to discomfort glare when the glare source was located in the area surrounding the visual fixation point. Unfortunately, neither of these studies utilised visual tasks, whose influence could have been considered in evaluating glare response.

Based on these findings from the literature and on the results of the pilot tests (N= 3), the visual task was placed above the diffusive screen, with a displacement angle of 2.87^{o} from the subject's line of sight. In fact, a series of trials performed at lower displacement angles ($<2.87^{o}$) revealed that the differences between the votes of perceived glare sensation reported by tests subjects when varying visual tasks did not allow any robust statistical analysis.

The pilot tests also contributed to define the set luminance of the diffusive screen, which was kept at a constant value of 10,541.31 cd/m^2 throughout the experimental procedure. This luminance value was obtained by taking three independent spot-point measurements using a Minolta LS-110 luminance meter mounted on a tripod, which were then mean averaged to obtain the source luminance of the small diffusive screen (Table 3-38).

Table 3-38 Spot-point measurements of the diffusive screen and mean luminance values

Point measurement	Screen luminance (cd/m^2)
1	10,520.18
2	10,556.90
3	10,546.86
Mean	10,541.31

This value of source luminance corresponds to an IES-GI of 22 that, on the Hopkinson scale (Hopkinson, 1971), describes the glare source as being ‘Just Uncomfortable’. In fact, no substantive variations in glare response between visual tasks were reported by pilot test subjects when the source luminance of the small diffusive screen was set at a constant level corresponding to the glare sensation votes of ‘Just Perceptible’ and ‘Just Noticeable’.

Coherent with the procedure adopted by Tuaycharoen & Tregenza (2005), before the start of the experiments, the twelve visual tasks were presented in a randomised order to an independent sample of 20 test subjects, different from those used in the experiment. These observers, which did not participate in the collection of glare responses nor had been involved in any of the previous tests, were asked to evaluate the perceived level of difficulty of each visual task. Subjects were selected by convenience sampling among University doctoral students, and varied in age, nationality, and cultural background. The independent observers were requested to provide scores of perceived visual task difficulty on an ascending 4-point rating scale (ranging from 1= ‘Least Difficult’ to 4= ‘Most Difficult’), for each of the three groups of tasks (i.e. font size, internal contrast, and character contrast). The visual tasks were presented to subjects with the same size, font, and contrast manipulations as those used in the experiment. To aid in the interpretation of the outcome, all subjects were informed that the visual difficulty of the task was to be intended as “*the ability to verify the numerical information contained on the paper, depending on its size and contrast*” (Boyce, 2014).

For each of the visual tasks under examination (Tasks 1-12), Table 3-39 presents the mean and standard deviation (SD) scores from these independent assessments.

Table 3-39 Visual tasks and scores of perceived difficulty (mean and standard deviation)

Task number	Perceived difficulty – Mean (SD)
1	1.05 (0.22)
2	1.95 (0.22)
3	3.00 (0.00)
4	4.00 (0.00)
5	1.00 (0.00)
6	2.75 (0.79)
7	2.80 (0.70)
8	3.45 (0.83)
9	1.00 (0.00)
10	2.00 (0.00)
11	3.00 (0.00)
12	4.00 (0.00)

3.4.2.4. Experimental Procedure

Coherent with previous experiments (Sections 3.1 and 3.2), subjects were requested to participate to four test sessions. However, in the new design for this experiment, sessions were held on different days (one session per day), each at a different time, and under a randomised sequence. The sessions were scheduled as follows:

- Morning: 09:00 or 09:30
- Midday: 12:00 or 12:30
- Afternoon: 15:00 or 15:30
- Evening: 18:00 or 18:30

The randomised sequence of test sessions was pre-assigned to each participant using a balanced Latin-squared design, while the Randomiser software (via MATLAB) was used to shuffle the order of the visual tasks presented to test subjects. By counterbalancing the sequence of sessions and tasks, systematic behavioural effects on subjects were created (i.e., learning and experience), thus making the experiment more sensitive in uncovering the effect of experimental interest (Field and Hole, 2013).

Table 3-40 presents the participant number, the order in which each participant completed the test sessions, and the day when each test session took place. Based on the principle of Latin-squared design, the sequence was repeated every four participants.

Table 3-40 Experimental sequence of test sessions for each participant (Latin-squared design)

Participant	**Order of test sessions**			
	Day 1	Day 2	Day 3	Day 4
1	A	B	C	D
2	B	D	A	C
3	D	C	B	A
4	C	A	D	B
5	A	B	C	D
6	B	D	A	C
7	D	C	B	A
8	C	A	D	B
9	A	B	C	D
10	B	D	A	C
11	D	C	B	A
12	C	A	D	B
13	A	B	C	D
14	B	D	A	C
15	D	C	B	A
16	C	A	D	B
17	A	B	C	D
18	B	D	A	C
19	D	C	B	A
20	C	A	D	B

A = “Morning’, B = “Midday”, C = “Afternoon”, D = “Evening”

At the beginning of the experiment, participants were asked to adjust the stool so that their head was positioned at a height of 1.2 m from the floor and at the correct viewing position in front of the small diffusive screen. A detailed set of instructions was then given, including a definition of discomfort glare, the meaning of the GSV criteria (described in Section 3.1.2.4) linked to the time-span descriptors, and an explanation of how the experiment would run. At this point, subjects were requested to sign a consent form and to provide general demographic information (e.g., age, gender, ethnicity, etc.) on a paper questionnaire.

To confirm that participants had a proper understanding of the GSV criteria, and to allow them to familiarise themselves with the experimental setup, before the start of their first session, each subject was asked to perform a pre-test procedure using a reference visual task. This contained 9 numbers ranging from 1 to 9 in ascending order from left to right; their font was black Verdana, their size was set at 10 points, and they were placed within a clear white background. The subject was instructed to look at the visual task positioned above the glare source – whose luminance was kept at constant 10,541.31cd/m^2 – and, using a 10 cm long continuous GSV visual analogue scale (VAS) (Figure 3-26), to place a mark on the scale corresponding to their perceived level of glare sensation. 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. In addition, test subjects were asked to mark their self-assessed level of fatigue on another VAS ranging from 'Not Fatigued' to 'Very Fatigued'. Each VAS was handed out by the experimenter. Both VAS scales were presented to the subject with the same size (10 cm). Subjects were informed that fatigue can be intended as "*a symptom of discomfort or dislike to a given activity, which can be related to tiredness or loss of energy*" (Sharpes & Wilks, 2002).

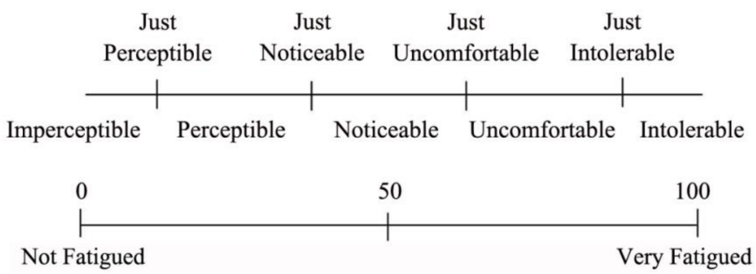


Figure 3-26 Continuous Visual Analogue Scales (VAS) for Glare Sensation Votes (GSV) and self-assessed fatigue

After the subject provided their assessments of glare sensation and fatigue, a relaxation period of two minutes was allowed and the real experiment began. The subject was asked to direct their gaze towards a fixation point marked above the glare source. The luminance of the screen was kept at a constant luminance of 10,541.31 cd/m^2 controlled by the experimenter via a computer connected to the projector. After a short period 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 3 random numbers (ranging between 1 and 9 – for example 1, 3, and 7) within the visual task. Then, using a pen provided, the subject was requested to mark a clear vertical line that ranged anywhere between ‘Imperceptible’ and ‘Intolerable’ on a GSV visual analogue scale to indicate their perceived level of glare sensation. The following task was then presented, the GSV was collected, and the procedure was repeated with all visual tasks. At the end of each session, participants were requested to provide self-assessments of several temporal variables (i.e., factors that vary with the time of day) measured on continuous visual analogue scales presented with the same size of the GSV one. This allowed any confounding effect of temporal variables on visual discomfort judgements to be accounted for through statistical adjustment (i.e., covariate testing). The visual analogue scales and relative descriptors for the temporal variables that were developed for this experiment 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 test. Each session lasted around 15 minutes.

3.4.2.5. Photometric Values and Glare Indices

In order to allow an easy comparison between votes of glare sensation and conventional glare indexes, Tuaycharoen & Tregenza (2007) proposed that Glare Sensation Votes (GSVs) reported by test subjects could be converted into a more meaningful measure by 'scaling' their values to an equivalent Glare Response Vote (GRV). The index resulting from their work – GRV(DGI) scaled to the DGI (Daylight Glare Index) (Chauvel *et al.*, 1982) – is applicable to large luminance sources such as daylight from windows. However, since the experiment here described was conducted in a laboratory setting, it was considered that the values obtained by Tuaycharoen & Tregenza (2007) would have not been suitable for a small artificial lighting source. Therefore, the data obtained by Hopkinson (1971) and the tables by Chauvel *et al.* (1982) – which allow to relate GSVs to corresponding values of IES-GI (Table 3-41) – were used to convert the scale of GSVs into a GRV(IES-GI) score, thereby creating a Glare Response Vote (GRV) formula scaled to the IES-GI glare index that is suitable for assessing perceived glare sensation from small artificial lighting sources.

Table 3-41 Comparison between Glare Sensation Votes (GSV) and IES-GI glare index for similar criteria of glare sensation

	GSV	IES-GI
Just (im)perceptible	0	10
Just acceptable	1	16
Borderline between comfort and discomfort	1.5	19
Just uncomfortable	2	22
Just intolerable	3	28

Utilising the data of Table 3-41, displaying the glare descriptors, the numerical values of GSVs, and the equivalent values of IES-GI, a GRV formula was derived (Equation 3-10):

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

3.4.2.6. Statistical Analysis

A total of 20 subjects volunteered to take part to the experiment (all postgraduate students, 6 male and 14 female). Test 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 (Osterhaus and Bailey, 1992; Rodriquez & Pattini, 2014; Tuaycharoen & Tregenza, 2005; Sivak *et al.*, 1989). The criterion adopted for the selection of subjects was purposive sampling, and all participants were recruited via an online advertisement addressed to all postgraduate students in architecture at the University of Nottingham. All tests were performed during the months of May and June, a period of mixed weather, varying from overcast to clear skies.

For all variables under examination, graphical inspection (e.g., Q-Q plot) and statistical analysis (e.g., K-S and S-W tests) revealed that the data were normally distributed around the mean. Furthermore, the Levine's test of homogeneity of variance showed no statistically significant evidence to suggest different variances associated with each variable. Table 3-42 presents the test statistic, the degrees of freedom (df), and the statistical significance (*p*-value) for both the K-S and S-W tests performed on the GRV(IES-GI) calculated on the votes of glare sensation provided by subjects, for all visual tasks and for all times of the day.

Table 3-42 Results of the Kolmogorov-Smirnov and Shapiro-Wilk tests

Visual Task	Session	Kolmogorov-Smirnov			Shapiro-Wilk		
		Statistic	df	*p*-value	Statistic	df	*p*-value
Task1	Morning	0.14	20	0.20 n.s.	0.93	20	0.15 n.s.
	Midday	0.09	20	0.20 n.s.	0.95	20	0.36 n.s.
	Afternoon	0.15	20	0.20 n.s.	0.94	20	0.24 n.s.
	Evening	0.16	20	0.15 n.s.	0.92	20	0.11 n.s.
Task2	Morning	0.09	20	0.20 n.s.	0.97	20	0.90 n.s.

	Midday	0.10	20	0.20 n.s.	0.98	20	0.92 n.s.
	Afternoon	0.09	20	0.20 n.s.	0.97	20	0.84 n.s.
	Evening	0.12	20	0.20 n.s.	0.97	20	0.89 n.s.
Task3	Morning	0.13	20	0.20 n.s.	0.95	20	0.40 n.s.
	Midday	0.15	20	0.20 n.s.	0.95	20	0.49 n.s.
	Afternoon	0.11	20	0.20 n.s.	0.97	20	0.79 n.s.
	Evening	0.08	20	0.20 n.s.	0.96	20	0.66 n.s.
Task4	Morning	0.10	20	0.20 n.s.	0.96	20	0.53 n.s.
	Midday	0.13	20	0.20 n.s.	0.94	20	0.35 n.s.
	Afternoon	0.14	20	0.20 n.s.	0.94	20	0.25 n.s.
	Evening	0.18	20	0.06 n.s.	0.94	20	0.34 n.s.
Task5	Morning	0.14	20	0.20 n.s.	0.95	20	0.40 n.s.
	Midday	0.12	20	0.20 n.s.	0.96	20	0.58 n.s.
	Afternoon	0.08	20	0.20 n.s.	0.98	20	0.99 n.s.
	Evening	0.15	20	0.20 n.s.	0.94	20	0.27 n.s.
Task6	Morning	0.15	20	0.20 n.s.	0.94	20	0.35 n.s.
	Midday	0.16	20	0.15 n.s.	0.95	20	0.38 n.s.
	Afternoon	0.16	20	0.16 n.s.	0.94	20	0.33 n.s.
	Evening	0.25	20	0.00**	0.88	20	0.01**
Task7	Morning	0.08	20	0.20 n.s.	0.98	20	0.99 n.s.
	Midday	0.10	20	0.20 n.s.	0.95	20	0.53 n.s.
	Afternoon	0.14	20	0.20 n.s.	0.94	20	0.23 n.s.
	Evening	0.21	20	0.02 n.s.	0.94	20	0.27 n.s.
Task8	Morning	0.13	20	0.20 n.s.	0.96	20	0.65 n.s.
	Midday	0.16	20	0.17 n.s.	0.92	20	0.14 n.s.
	Afternoon	0.17	20	0.12 n.s.	0.92	20	0.11 n.s.
	Evening	0.12	20	0.20 n.s.	0.96	20	0.62 n.s.
Task9	Morning	0.12	20	0.20 n.s.	0.96	20	0.55 n.s.
	Midday	0.11	20	0.20 n.s.	0.97	20	0.78 n.s.
	Afternoon	0.13	20	0.20 n.s.	0.95	20	0.37 n.s.
	Evening	0.12	20	0.20 n.s.	0.96	20	0.59 n.s.
Task 10	Morning	0.16	20	0.17 n.s.	0.97	20	0.74 n.s.
	Midday	0.11	20	0.20 n.s.	0.98	20	0.99 n.s.
	Afternoon	0.10	20	0.20 n.s.	0.95	20	0.46 n.s.
	Evening	0.10	20	0.20 n.s.	0.96	20	0.69 n.s.
Task 11	Morning	0.11	20	0.20 n.s.	0.98	20	0.94 n.s.
	Midday	0.11	20	0.20 n.s.	0.96	20	0.57 n.s.
	Afternoon	0.18	20	0.08 n.s.	0.91	20	0.06 n.s.
	Evening	0.13	20	0.20 n.s.	0.96	20	0.56 n.s.
Task 12	Morning	0.08	20	0.20 n.s.	0.96	20	0.67 n.s.
	Midday	0.11	20	0.20 n.s.	0.93	20	0.21 n.s.
	Afternoon	0.11	20	0.20 n.s.	0.95	20	0.49 n.s.
	Evening	0.20	20	0.02*	0.92	20	0.11 n.s.

*** highly significant; ** significant; * weakly significant; n.s. = not significant

The K-S tests indicated that the differences detected are statistically significant – i.e., the distribution of the data collected around the mean parameter is different from normal – in 1 case, weakly significant in 1 case, and non-significant in 46 out of 48 cases. For the S-W tests, the results showed only 1 statistically significant difference out of 48 cases.

Table 3-43 presents the degrees of freedom (df), the *F* test statistic, and the statistical significance (*p*-value) for the Levine's tests of homogeneity of variance. The results indicate that the differences in variances associated with each independent variable (time of the day) for each visual task are not statistically significant in all twelve cases.

Table 3-43 Levine's test of homogeneity of variance

Visual Task	df	***F***	*p*-value
Task1	3, 76	1.63	0.18 n.s.
Task2	3, 76	0.44	0.72 n.s.
Task3	3, 76	0.28	0.83 n.s.
Task4	3, 76	0.21	0.88 n.s.
Task5	3, 76	0.61	0.60 n.s.
Task6	3, 76	2.44	0.07 n.s.
Task7	3, 76	0.75	0.52 n.s.
Task8	3, 76	0.84	0.47 n.s.
Task9	3, 76	0.21	0.88 n.s.
Task10	3, 76	1.36	0.26 n.s.
Task11	3, 76	0.62	0.60 n.s.
Task12	3, 76	0.39	0.75 n.s.

*** highly significant; ** significant; * weakly significant; n.s. = not significant

Since the basic assumptions of normality and homogeneity of variance were not violated, parametric tests were adopted for the analysis of the data (Field, 2013; Siegel, 1957).

3.4.3. Results

3.4.3.1. Temporal Variation of Glare Response for Each Visual Task

Among the groups of visual tasks, Figure 3-27 presents the results related to variation of font size (Tasks 1 to 4). The y-axis displays the GRV(IES-GI) calculated on the votes of glare sensation reported by test subjects when presented with Task 1, 2, 3, and 4 (Section 3.4.2.2), which are displayed on the x-axis. For each boxplot, the graph groups the statistical parameters by the time of the day in which the votes of glare sensation were provided.

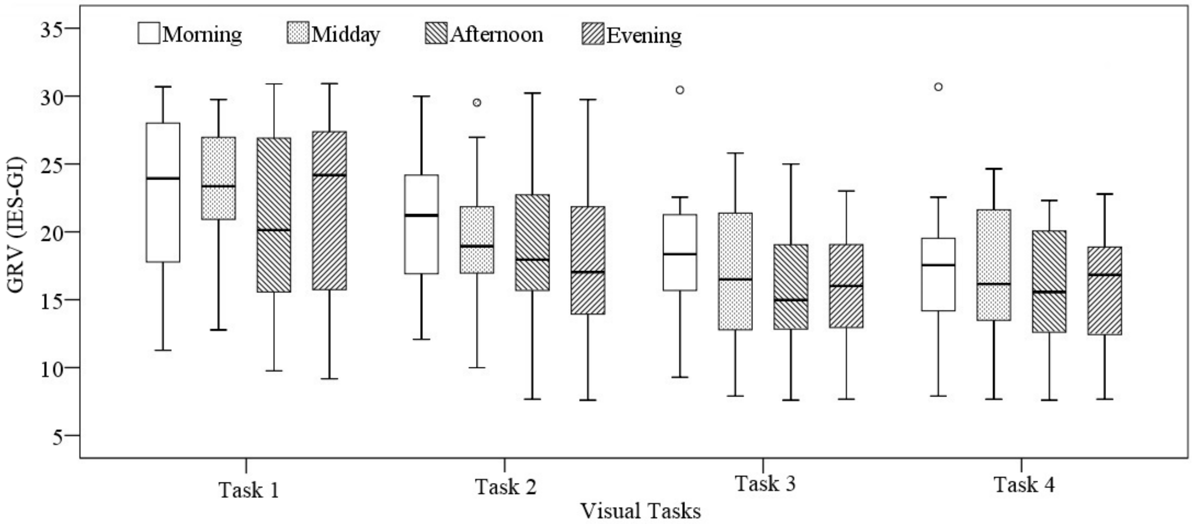


Figure 3-27 Boxplots of GRV(IES-GI) statistical parameters at each test session for Tasks 1 to 4 (variation of font sizes)

Consistent with Figure 3-27, inspection of the boxplots for all 12 visual tasks suggested a tendency for statistical parameters (minimum, 25th percentile, median, 75th percentile, and maximum) to correspond to lower levels of GRV(IES-GI) as the day progresses. That is, when test subjects were presented with the same visual task at later times of the day, the constant source luminance was reported as being less visually discomfortable.

To analyse the data, 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. With reference to the latter, in fact, a review of the literature and findings from previous experiments (Section 3.2) provided reasons to suspect that glare response could be influenced by fatigue.

Tests of homogeneity of regression slopes revealed that the correlations between the dependent variable (GRV(IES-GI)) and the covariate (fatigue) were not statistically different from each other: $F(48, 248.57)= 0.53$, $p= 0.53$ n.s., Wilk's $\Lambda= 0.52$, $p\eta^2= 0.15$ (small). Since this assumption was not violated, the parametric MANCOVA was selected for this analysis (Siegel, 1957) to guard against the potential confounding variable (fatigue) not experimentally controlled over the independent variables, while making the tests more sensitive to detect the effect of experimental interest (Field, 2013).

The results from the MANCOVA showed no statistically significant differences: $F(36, 189.82)= 0.94$, $p= 0.58$ n.s., Wilk's $\Lambda= 0.62$, $p\eta^2= 0.15$ (small). However, when controlling for the effect of fatigue, a statistically significant difference was detected: $F(12, 64)= 2.00$, $p= 0.04$*, Wilk's $\Lambda= 0.74$, $p\eta^2= 0.27$ (moderate). An ANCOVA was then performed to search for significant differences between the independent variables (time of the day) for each visual task. Table 3-44 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, the table shows the *F* test statistic and the degrees of freedom (df), the statistical significance (*p*-value), and the effect size ($p\eta^2$) for every visual task, which was calculated by making use of equivalence between the standardised measure of the detected difference between independent groups and the partial-eta squared ($p\eta^2$). The interpretation of the outcome was again derived from the tables provided by Ferguson (2009), where benchmarks are given for small, moderate, and large effect sizes ($p\eta^2 \geq 0.04$, 0.25, and 0.64 respectively).

Table 3-44 ANCOVA and effect sizes for each visual task

	Fatigue			Time of Day		
Task	*F* (df)	*p*-value	Effect Size	*F* (df)	*p*-value	Effect Size
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

*** highly significant; ** significant; * weakly significant; n.s. = not significant
$p\eta^2<0.04$= negligible; $p\eta^2\geq0.04$= small; $p\eta^2\geq0.25$= moderate; $p\eta^2\geq0.64$= large

The 'Fatigue' column shows that the covariate has 2 weakly significant (Tasks 10 and 12) and 2 significant (Task 5 and 11) effects on the dependent variable. In general, the influence of the covariate appears to be substantive, displaying small but practically relevant effect sizes 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 2 highly significant (Tasks 5 and 7), 5 significant (Tasks 3, 4, 8, 10, and 11), and 3 weakly significant (Tasks 6, 9, and 12) differences. The significant differences detected have consistently a substantive magnitude, all corresponding to a practically relevant effect size ($0.04\leq p\eta^2<0.25$ in 10 out of 12 cases). Also in this case, for Tasks 1 and 2, the values of effect size were below the recommended minimum for practical relevance ($p\eta^2<0.04$). To isolate the main effects detected, contrasts were made using pairwise comparisons (Bird & Hadzi-Pavlovic, 2014). The ANCOVA showed no statistically significant or practically relevant differences for Tasks 1 and 2; therefore, these tasks were excluded from the subsequent *post-hoc* analysis.

In consideration of the experiment-wise error rate caused by the alpha-level inflating across multiple pairwise comparisons carried out with the same hypotheses – 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 (as if these were the only to be compared), and declares statistically significant any detected difference that is larger than the LSD (Williams & Abdi, 2010). The LSD makes no attempt to control the Type I error, but requires the initial

ANCOVA to be statistically significant (Field, 2013). This test has more statistical power compared with other post-hoc correction methods since the alpha-level (p-value) is not adjusted (Williams & Abdi, 2010). In fact, due to the number of independent variables and its over-conservative nature, the use of Bonferroni correction (utilised in Section 3.1.3) would have significantly inflated the risk of occurrence of a Type II error (Cabin & Mitchell, 2000; Nakagawa, 2004). As a result, the interpretation of the findings would have become strongly dependent on the number of tests performed (Perneger, 1998).

The outcome of each contrast was measured by equivalence between the effect size and the Cohen's d coefficient (Equation 3-11):

$$Effect\ size = \frac{m_1 - m_2}{\sigma} \qquad \text{(Equation 3-11)}$$

Where m_1 and m_2 are the mean values of GRV(IES-GI) associated with the pairwise comparisons corresponding to the sessions in which subjects reported their votes of glare sensation, and σ is the pooled variance between independent groups. Both statistical parameters (mean and variance) are adjusted for the effect of the covariate (fatigue).

The effect size was considered small, moderate, and large for, respectively, $d \geq 0.41$, 1.15, and 2.70 (Ferguson, 2009).

Based on initial graphical inspection of data, and informed by the results of previous experiments, the tests adopted the one-tailed alternative hypothesis that the constant source luminance is reported by subjects as being less visually discomfortable as the day progresses.

Table 3-45 presents 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) scores calculated at all test sessions, the difference between the means (ΔM) and the outcome of its statistical significance (NHST, p-value calculated with a one-tailed test), the lower (CI_L) and upper (CI_U) 95% confidence intervals for the mean difference (ΔM), and the effect size (d).

Table 3-45 Pairwise comparisons between test sessions and effect sizes for each visual task

Task	Test Sessions	Mean (SD)	Mean (SD)	ΔM^{NHST}	$[CI_L, CI_U]$		Effect Size (d)
3	Morn. vs. Mid	21.20 (2.18)	19.87 (2.16)	1.33 n.s.	-1.79	4.42	0.61
	Morn. vs. Aft.	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
	Mid vs. Aft.	19.87 (2.16)	18.22 (2.13)	1.65 n.s.	-1.50	4.84	0.77
	Mid vs. Even.	19.87 (2.16)	16.73 (2.22)	3.14*	-0.14	6.44	1.43
	Aft. vs. Even.	18.22 (2.13)	16.73 (2.22)	1.49 n.s.	-1.65	4.62	0.68
4	Morn. vs. Mid	24.13 (2.54)	24.43 (2.51)	-0.30 n.s.	-3.92	3.33	-0.12
	Morn. vs. Aft.	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
	Mid vs. Aft.	24.43 (2.51)	20.62 (2.49)	3.81*	0.11	7.52	1.52
	Mid vs. Even.	24.43 (2.51)	20.50 (2.59)	3.93*	3.34	7.77	1.57
	Aft. vs. Even.	20.62 (2.49)	20.50 (2.59)	0.12 n.s.	-1.52	1.62	0.05
5	Morn. vs. Mid	20.55 (1.61)	19.78 (1.60)	0.77 n.s.	-1.54	3.05	0.48
	Morn. vs. Aft.	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
	Mid vs. Aft.	19.78 (1.60)	16.41 (1.58)	3.38**	1.04	5.73	2.12
	Mid vs. Even.	19.78 (1.60)	16.97 (1.58)	2.82**	0.39	5.25	1.77
	Aft. vs. Even.	16.41 (1.58)	16.97 (1.64)	-0.56 n.s.	-2.87	1.76	-0.35
6	Morn. vs. Mid	22.10 (1.98)	21.59 (1.96)	0.51 n.s.	-2.31	3.34	0.26
	Morn. vs. Aft.	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
	Mid vs. Aft.	21.59 (1.96)	18.73 (1.94)	2.86*	-0.04	5.73	1.47
	Mid vs. Even.	21.59 (1.96)	19.18 (2.02)	2.41*	-0.51	5.40	1.21
	Aft. vs. Even.	18.73 (1.94)	19.18 (2.02)	-0.45 n.s.	-3.28	2.42	-0.23
7	Morn. vs. Mid	22.41 (1.72)	22.20 (1.70)	0.21 n.s.	-2.28	2.62	0.12
	Morn. vs. Aft.	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
	Mid vs. Aft.	22.20 (1.70)	20.20 (1.69)	2.00 n.s.	-0.51	4.50	1.18
	Mid vs. Even.	22.20 (1.70)	19.85 (1.75)	2.35*	-0.23	4.97	1.36
	Aft. vs. Even.	20.20 (1.69)	19.85 (1.75)	0.35 n.s.	-2.08	2.87	0.20
8	Morn. vs. Mid	22.73 (1.97)	22.73 (1.95)	0.00 n.s.	-2.80	2.79	0.00
	Morn. vs. Aft.	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
	Mid vs. Aft.	22.73 (1.95)	20.62 (1.93)	2.11 n.s.	-0.74	4.98	1.09
	Mid vs. Even.	22.73 (1.95)	20.13 (2.01)	2.60*	-0.36	5.57	1.31
	Aft. vs. Even.	20.62 (1.93)	20.13 (2.01)	0.49 n.s.	02.32	3.32	0.25
9	Morn. vs. Mid	17.78 (1.84)	18.59 (1.82)	-0.81 n.s.	3.44	1.80	-0.44
	Morn. vs. Aft.	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
	Mid vs. Aft.	18.59 (1.82)	16.41 (1.80)	2.18*	-0.47	4.86	1.20
	Mid vs. Even.	18.59 (1.82)	16.01 (1.88)	2.58*	-0.19	5.34	1.39
	Aft. vs. Even.	16.41 (1.80)	16.01 (1.88)	0.40 n.s.	-2.23	3.02	0.33
10	Morn. vs. Mid	21.04 (2.18)	20.31 (2.16)	0.73 n.s.	-2.39	3.82	0.34
	Morn. vs. Aft.	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
	Mid vs. Aft.	20.31 (2.16)	18.01 (2.14)	2.30 n.s.	-0.86	5.49	1.07
	Mid vs. Even.	20.31 (2.16)	16.62 (2.22)	3.69**	0.41	7.00	1.68
	Aft. vs. Even.	18.01 (2.14)	16.62 (2.22)	1.39 n.s.	-1.73	4.53	0.62
11	Morn. vs. Mid	21.55 (2.13)	20.73 (2.11)	0.82 n.s.	-2.18	3.85	0.39
	Morn. vs. Aft.	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
	Mid vs. Aft.	20.73 (2.11)	17.55 (2.08)	3.18*	0.12	6.28	1.52
	Mid vs. Even.	20.73 (2.11)	16.22 (2.17)	4.51**	1.32	7.71	2.11
	Aft. vs. Even.	17.55 (2.08)	16.22 (2.17)	1.33 n.s.	-1.69	4.39	0.61
12	Morn. vs. Mid	23.29 (2.55)	23.03 (2.53)	0.26 n.s.	-3.39	3.90	0.10
	Morn. vs. Aft.	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
	Mid vs. Aft.	23.03 (2.53)	19.38 (2.50)	3.65*	-0.08	7.38	1.45
	Mid vs. Even.	23.03 (2.53)	18.43 (2.60)	4.60**	0.75	8.49	1.79
	Aft. 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

Consistent with previous findings (discussed in Section 3.1 and Section 3.2), analysis of descriptive statistics (i.e., means and their differences, ΔM) shows for all 12 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 4 comparisons across visual tasks. In these 4 cases, the differences detected are not statistically significant.

The pairwise comparisons provide evidence that the differences between GRV(IES-GI) scores calculated at different times of day are highly significant in 2 cases, significant in 11 cases, weakly significant in 20 cases, and not significant in 27 cases, out of a total of 60 comparisons. The differences have moderate effect sizes ($1.15\leq d<2.70$) in 34 cases and small magnitudes ($0.41\leq d<1.15$) in 11 cases. Negligible coefficients d ($d<0.41$) were detected in 15 cases. All the negligible effect sizes correspond to non-statistically significant differences.

For 6 visual tasks (Tasks 3, 7, 8, 10, 11 and 12), the largest statistically significant differences as measured by the effect sizes (and by the differences between the adjusted means (ΔM) of GRV(IES-GI)) occurred between the Morning and Evening sessions, corresponding to a gap of 9 hours between tests. Therefore, descriptive and inferential analysis of the data suggests that, for these visual tasks, the length of the time interval between test sessions may have a direct influence on the increased tolerance to artificial lighting from a diffusive screen with constant source luminance as the day progresses. These 6 tasks had an independently-assessed visual task difficulty ranging from 2.00 (0.00) for Task 10 to 4.00 (0.00) for Task 12.

For Tasks 5 and 6, the largest differences in effect size (and ΔMs between GRV(IES-GI) scores) were detected between the Morning and Afternoon sessions (both significant and with, respectively, d= 2.47 and d= 1.72), corresponding to an interval of 6 hours between the tests.

For these tasks, Figure 3-28 and Figure 3-29 plot, on the y-axis, the GRV(IES-GI) calculated from the votes of glare sensation provided by subjects in the Midday, Afternoon, and Evening sessions and, on the x-axis, the glare response votes related to the Morning session.

In both figures, the null hypothesis line is plotted along the diagonal, representing no differences between the calculated GRV(IES-GI) scores at each test session.

Coherent with the literature (Tuaycharoen & Tregenza, 2005; 2007; Wienold & Christoffersen, 2006), 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 indicated by the coefficients of determination).

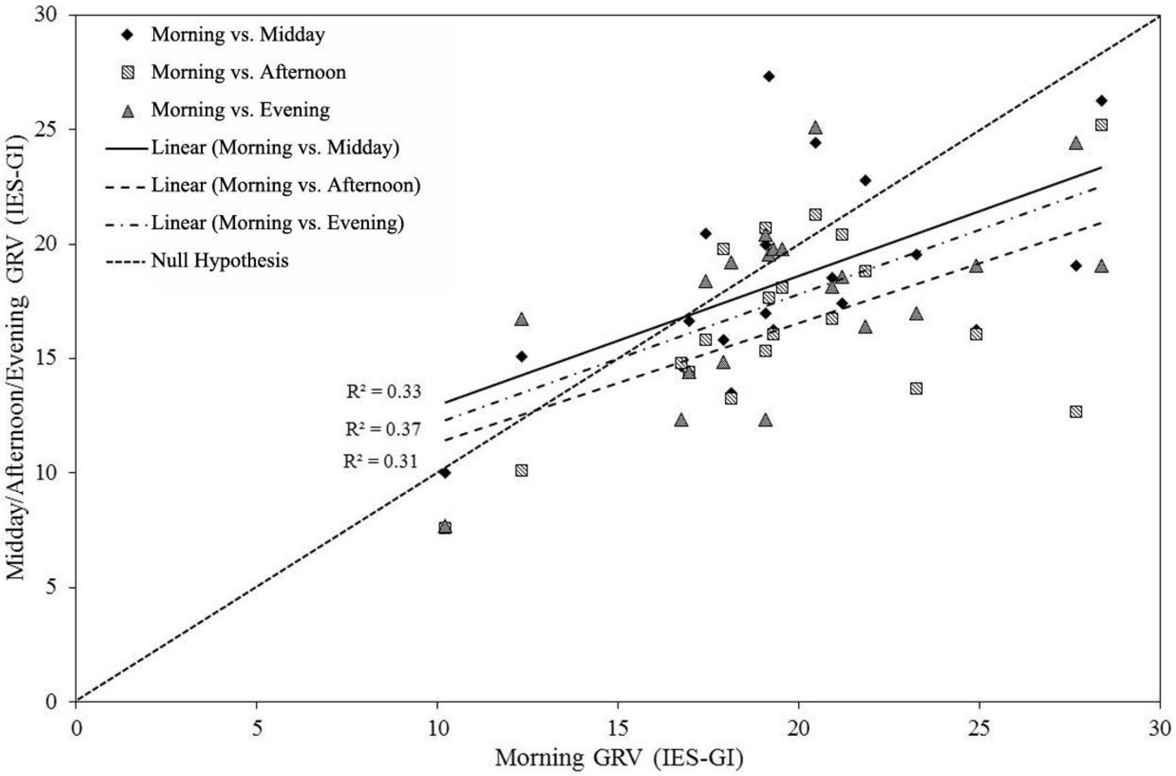


Figure 3-28 Comparisons between GRV(IES-GI) scores at the Morning (x-axis) and Midday/Afternoon/Evening (y-axis) test sessions for Task 5

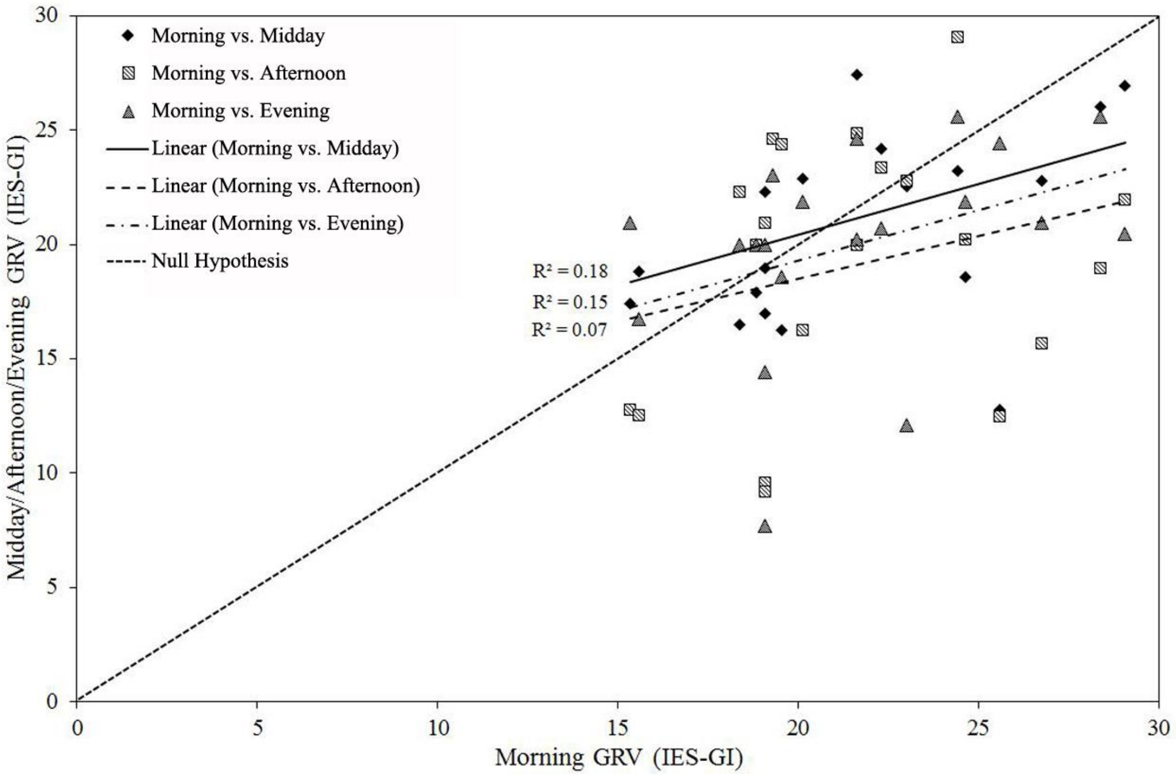


Figure 3-29 Comparisons between GRV(IES-GI) scores at the Morning (x-axis) and Midday/Afternoon/Evening (y-axis) test sessions for Task 6

Although some GRV(IES-GI) scores are above the null hypothesis line, indicating that some subjects provided higher votes of glare sensation in the Morning session, the interpolated lines for each comparison are notably below it. This supports the previous hypothesis (see also Section 3.1 and Section 3.2) that the time interval between sessions may be directly related to an increased tolerance to source luminance along the day. It is important to remind here that evaluations across the independent variable – i.e., the regressions points corresponding to two different sessions – were compared under a constant source luminance. At lower levels of glare response vote, both figures show the interpolations above the null hypothesis line, 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.

When the independent variables were regressed against each other, the presence of individual differences led to a wide scatter in the results, as confirmed by the low coefficients of determination in the linear fits. Interestingly, the spread of the data inflated from Task 5 to Task 6, resulting in considerably lower R^2 for the latter. This may have resulted from increased task 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) for Task 6. However, there is not sufficient evidence to infer whether the increased variability in GRV(IES-GI) scores was caused by visual task difficulty itself, or by the amplification of the effect of other confounding variables not experimentally controlled.

For Tasks 4 and 9, the largest significant differences in effect sizes (and ΔM) were detected between the Midday and Evening sessions, an interval of 6 hours between tests. Tasks 4 and 9 were independently assessed with a task difficulty of, respectively, 4.00 (0.00) and 1.00 (0.00). In the case of Task 9, however, the performed pairwise comparisons detected only two statistically significant differences. Conversely, for Task 4 – with the exception of the differences between the Morning and Midday sessions, and the Midday and Evening tests, which display no statistical nor practical significance – all other differences resulted in a very narrow range of effect sizes ($1.40 \leq d \leq 1.57$). A similar trend is also apparent for the other tasks that had been independently rated as more visually difficult in each group.

For example, 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 comprised between d= 1.45 and d= 1.89 for all the statistically significant differences detected. The means (adjusted for the effect of fatigue) for Tasks 4, 8, and 12 correspond almost invariably to the highest GRV(IES-GI) across all test sessions, a result that is consistent with the literature (Sivak *et al.*, 1989).

This may again suggest that, for more visually demanding stimuli, the effect of time of day on temporal variation of glare sensation may be masked or influenced by other variables and factors, such as the difficulty of the task.

3.4.3.2. Temporal Variation of Glare Response across Visual Tasks

A further analysis was performed to compare temporal variations of glare response across groups of visual tasks as the day progresses, considering the manipulation of their inclusive features (i.e., font size, internal contrast, and font contrast) and their respective independently-assessed task difficulty.

Figure 3-30 to Figure 3-33 plot, on the y-axis, the GRV(IES-GI) calculated at each session and, on the x-axis, the three groups of tasks varying in font size (Tasks 1-4), internal contrast (Tasks 5-8), and character contrast (Tasks 9-12).

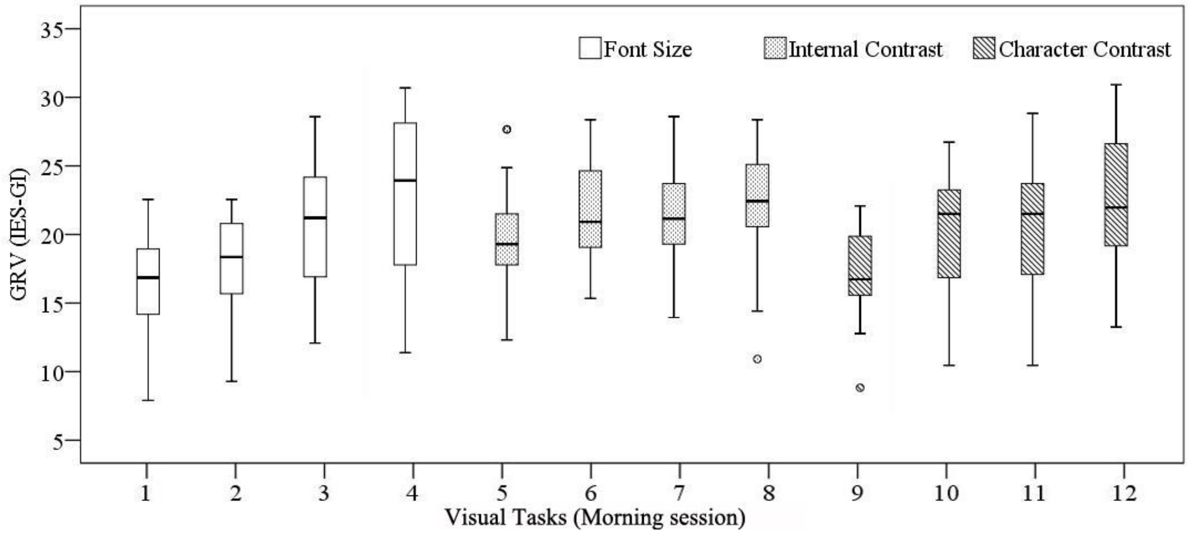


Figure 3-30 Boxplots of GRV(IES-GI) scores for each group of visual tasks in the Morning

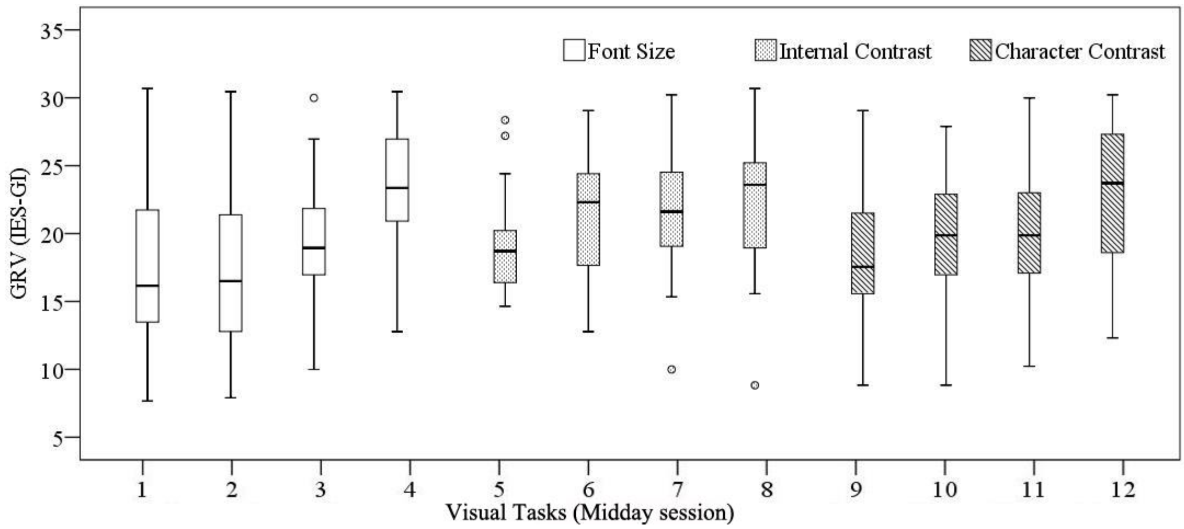


Figure 3-31 Boxplots of GRV(IES-GI) scores for each group of visual tasks at Midday

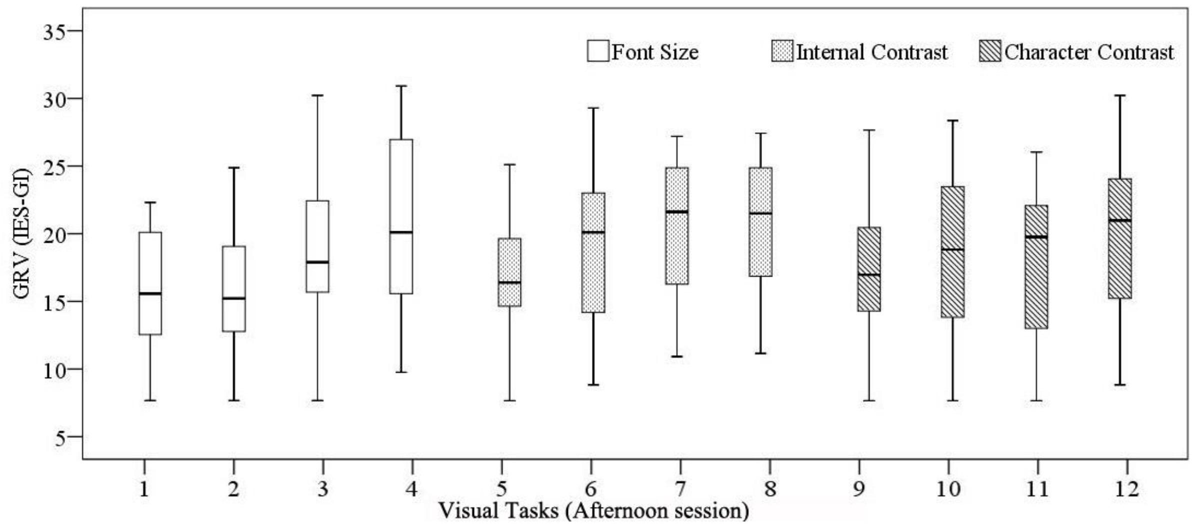


Figure 3-32 Boxplots of GRV(IES-GI) scores for each group of visual tasks in the Afternoon

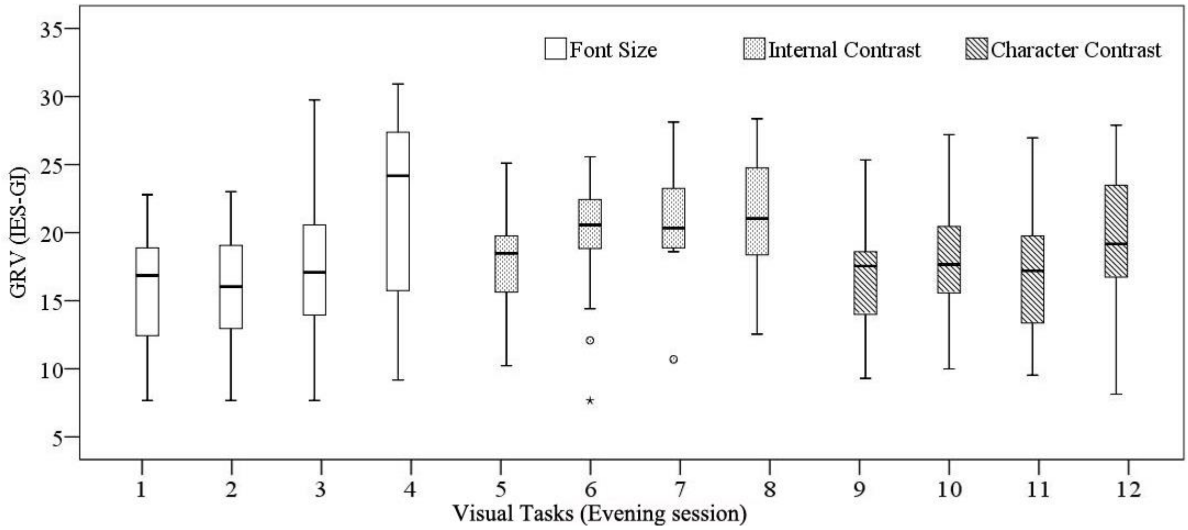


Figure 3-33 Boxplots of GRV(IES-GI) scores for each group of visual tasks in the Evening

Visual inspection of the data confirmed expected variations in the reported glare sensation experienced for different visual tasks. This inference can be made since the source luminance remained constant throughout the experimental procedure, therefore any changes in glare response can be related to the difficulty of the visual tasks presented to the participant (although, clearly, other uncontrolled confounding variables could also play a role).

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 test subjects featured progressively smaller 'font size' (Task 1 had a font size of 10, which was reduced with 2-point decrements for each subsequent task). 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 response. This trend of intensifying glare sensation further reinforces when the task information is distorted with vertical (Task 7) and horizontal lines (Task 8). Finally, in terms of the 'character contrast' group (Tasks 9-12), a tendency is recognisable where, as the contrast of the characters decreases (from 50% of Task 9 to 15% of Task 12), the perceived level of glare sensation increases. The tendencies for all three groups of visual tasks present, almost invariably, temporal consistency; that is, similar trends across tasks can be recognised at all test sessions.

To analyse the data, a Multivariate Analysis of Variance (MANOVA) was performed, whereby the dependent variables (i.e., the GRV(IES-GI) at each test session) were compared to the grouping variables (visual tasks). As the visual tasks were presented to participants in a randomised order, and subjects only performed one test session per day, it could be reasonably assumed for this analysis across visual tasks that the potential effects of temporal variables (e.g., fatigue) and other confounding factors (e.g., learning) were masked from the experiment. Therefore, covariates were considered as controlled in the analysis (Field, 2013).

The MANOVA for the visual tasks detected at least one statistically significant difference, suggesting that task difficulty and manipulation influences the variation of glare sensation (Sivak *et al.*, 1989): $F(44, 862.75) = 1.76$, $p = 0.00$**, Wilk's $\Lambda = 0.72$, $p\eta^2 = 0.08$. *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 of the sessions are as follows: Morning: $F(11) = 4.94$, $p = 0.00$**, $p\eta^2 = 0.19$; Midday: $F(11) = 3.73$, $p = 0\ .00$***, $p\eta^2 = 0.15$; Afternoon: $F(11) = 2.50$, $p = 0.01$**, $p\eta^2 = 0.11$; Evening: $F(11) = 3.54$, $p = 0.00$***, $p\eta^2 = 0.15$. These findings suggest that not only glare response may be dependent on task difficulty, but that this relationship appears to vary as the day progresses, as per the different effect sizes associated with each test 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 the risk of inflating the significance level – could have made the study susceptible to Type II errors (Nakagawa, 2004). Instead, the visual tasks were divided into 3 grouping variables – 'font size' (Size)= Tasks 1-4; 'internal contrast' (I.Contrast)= Tasks 5-8; and, 'character contrast' (C.Contrast)= Tasks 9-12 – that were compared in an additional ANOVA (Table 3-46).

Table 3-46 ANOVA and effect sizes 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
Midday	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	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 3-46 provides evidence that the differences between the GRV(IES-GI) scores calculated for each groups of visual tasks at all times of day are highly significant in 2 out of 12 cases, significant in 3 cases, weakly significant in 3 cases, and not significant in 4 cases. The differences detected have a generally substantive effect size, ranging between moderate ($0.25\leq p\eta^2<0.64$) in 1 case and small ($0.04\leq p\eta^2<0.25$) in 10 cases. From these findings, there is evidence to suggest that inclusive features of the task (font size, internal contrast, and character contrast) may influence the temporal variation in perceived levels of glare sensation.

For the 'font size' (Size) of the visual task, the data present larger statistically significant and practically relevant differences in the Morning (*p*= 0.00***, $p\eta^2$= 0.25) and Midday (*p*= 0.00***, $p\eta^2$= 0.21), compared to the Afternoon (*p*= 0.01**, $p\eta^2$= 0.15) and Evening (*p*=

0.00**, $p\eta^2$= 0.16) sessions. This suggests that the decrease in font size of the visual task may have a larger effect over variations of glare response at earlier times of the day.

Likewise, the effect of varying the ‘character contrast’ of the font (C.Contrast) denotes a similar trend, although this tendency appears to be of lower magnitude as indicated by consideration of the effect sizes across test sessions. In fact, the largest statistically and practically significant differences due to reduction in the contrast of the characters are recorded in the Morning (p= 0.00**, $p\eta^2$= 0.16) and in the Midday (p= 0.05*, $p\eta^2$= 0.10).

This seems to suggest that *direct* manipulation of the information of interest contained within the visual task (i.e., the size and contrast of the characters) may have its strongest influence on variation of glare response at earlier test sessions and decrease its effect as the day progresses. However, these results have to be treated with caution since the ANOVA did not detect statistically significant differences in the Midday and Evening sessions upon consideration of character contrast.

Lastly, variation of the ‘internal contrast’ of the task (I.Contrast) shows an opposite tendency, since the influence of increased task difficulty due to saturation of the background and distortion of the task information appears to amplify with time of day, as demonstrated by increasing levels of effect size ($p\eta^2$ increasing from 0.05 to 0.10). These findings seem to indicate that *indirect* manipulation of the information featured within the visual stimulus (e.g., the background of the task) may have a larger effect on glare sensation at later times of the day. However, also in this case, results have to be interpreted prudently as no statistically significant differences were detected in the Morning and Midday sessions.

3.4.3.3. Influence of Temporal Variables on Glare Response

Consistent with previous analysis (Section 3.4.3.1 and Section 3.4.3.2), the GRV(IES-GI) 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 difficulty. To allow permutation testing, the temporal variables were organised in ordered categories, dividing the size of the continuous visual analogue scales (10 cm) in four groups (Wienold, 2009): 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 (e.g., 'not fatigued' and 'very fatigued') (see Figure 3-26).

The influence of task difficulty was not directly included in this analysis due to the large number of permutations that would have resulted. Conversely, the 12 visual tasks were grouped into the respective times of day when the glare assessments were given. Therefore, for each test 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 (Heiman, 2013). For all variables, the non-parametric Levine's test of homogeneity of variance – a test suitable for non-Gaussian distributions (Nordstokke & Zumbo, 2010; Nordstokke, *et al.*, 2011) – returned high significance in 16 cases, significance in 1 case, weak significance in 2 cases, and non-significance in 5 cases. The standardised magnitude of the differences returned small and negligible effect sizes (Ferguson, 2009); hence, non-parametric tests were adopted for this analysis (Siegel, 1957).

When independent groups have different variances associated with them, the interpretation of the alternative hypothesis using solely the statistical significance (*p*-value) can be problematic. Therefore, to support the interpretation of the outcomes, the mean ranks of the independent groups were also considered (Hart, 2001). Directionality of the hypothesis was confirmed by inspection of central tendencies and graphical displays (Hauschke & Steinijans, 1996). However, if no convincing directionality of differences between groups could be determined, a non-directional two-tailed hypothesis was applied (Ruxton & Neuhauser, 2010).

For all temporal variables considered, two inferential tests were adopted in the analysis. The Jonckheere-Terpstra test evaluates the statistical significance of priori ordering effects with respect to ordered variables and can be utilised to determine the magnitude and directionality of trends within the data. The Kruskal-Wallis one-way analysis of variance can be used to detect the statistical significance of differences between independent groups; this is an omnibus inferential test based on ranks, which calculates the H test statistic (Kruskal & Wallis, 1952). These tests were initially run to compare the glare responses for each temporal variable at all test sessions. A *posteriori* analysis was then performed using the Mann-Whitney U test, whereby all permutations between temporal variables, for each time of the day, were compared against each other. The Mann-Whitney U test was used to isolate where the main effects detected in the Jonckheere-Terpstra and Kruskal-Wallis tests were.

To counterbalance the experiment-wise error rate caused by the significant level inflating across multiple pairwise comparisons, due to the same data being analysed repeatedly (Cabin & Mitchell, 2000), Bonferroni corrections were applied. Since Null Hypothesis Significance Testing (NHST) depends on both the size of the sample and the size of the effect under examination (Cohen, 1965; Schiavon & Altomonte, 2014), the emphasis of the inferential analysis was placed primarily on the effect size and the calculated difference between descriptive statistics (i.e., the median), and not only on its statistical significance (which, particularly for uneven sample groups, could confound effect size and sample size). The effect size was calculated by making use of equivalence with the Pearson's coefficient r, and was considered small, moderate, or large respectively for $r \geq 0.20$, 0.50, and 0.80 (Ferguson, 2009), according to the following formula (Fields, 2013; Rosenthal & DiMatteo, 2001) (Equation 3-12):

$$Effect\ Size\ (r) = \frac{Z_{Score}}{\sqrt{N}} \qquad \text{Equation 3-12}$$

Where N is the number of observations, and the test statistic (Z_{score}) is extracted from the inferential tests (respectively, Jonckheere-Terpstra and Mann-Whitney U).

For each temporal variable, the following sections provide the graphical displays of the data (boxplots), the outcomes of the Jonckheere-Terpstra and of the Kruskal-Wallis tests, and the results related to the statistically significant and practically relevant Mann-Whitney U pairwise comparisons between ordered categories of variables.

For each contrast, tables are used to illustrate the sample size (N) of independent groups, 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), the mean ranks of the groups, the test statistic (U), and the effect size (r).

3.4.3.4. Fatigue

Figure 3-34 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, on the y-axis, the GRV(IES-GI) calculated on the individual glare assessments provided by test subjects and, on the x-axis, the four test sessions related to the different times of the day.

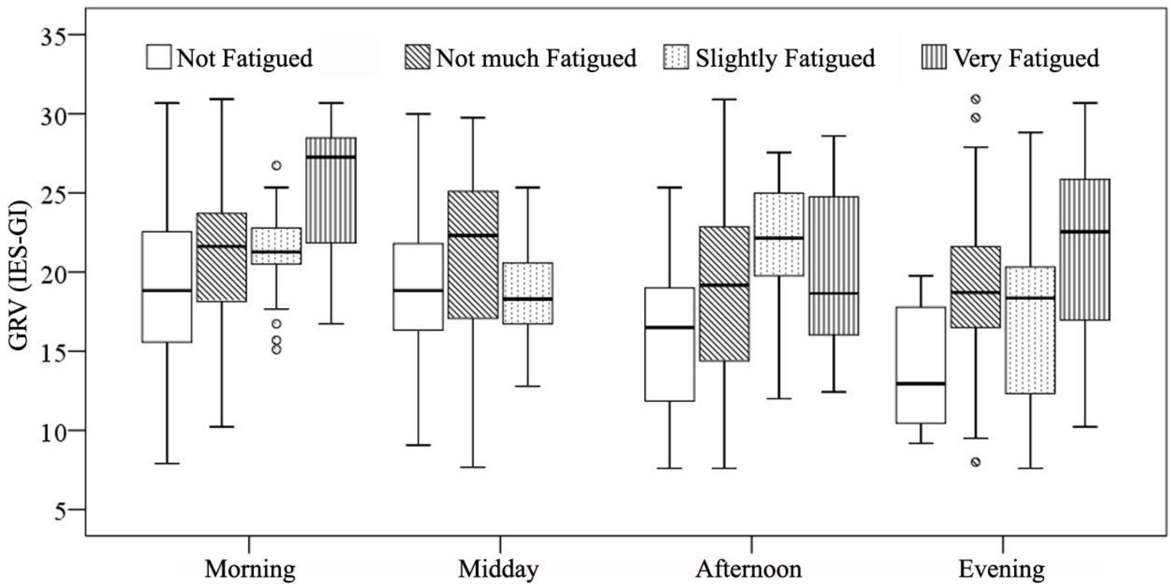


Figure 3-34 Boxplots of GRV(IES-GI) statistical parameters at each test session for the four ordered categories of the variable Fatigue

Graphical inspection of the data led to the alternative hypothesis that a prevailing uni-directional direct relationship existed between glare response votes and fatigue, as suggested by higher GRV(IES-GI) corresponding to increasing levels of this temporal variable. The one-tailed Jonckheere-Terpstra tests showed evidence of two highly significant and practically relevant differences at the Morning (J-T= 6,639.50, p= 0.00***, r= 0.27 (small)) and Afternoon sessions (J-T= 6,268.00, p= 0.00***, r= 0.30 (small)). The test did not return significant differences in the Midday and the Evening. Conversely, the Kruskal-Wallis tests resulted in highly significant differences for all the test sessions: Morning: $H(3)$= 20.10, p= 0.00***; Midday: $H(2)$= 12.62, p= 0.00***; Afternoon: $H(3)$= 24.82, p= 0.00***; Evening: $H(3)$= 29.21, p= 0.00***. Comparative *post-hoc* Mann-Whitney U tests were performed to isolate the main-effects detected. The results supported the postulated tendency of a direct relationship between fatigue and glare response, with higher glare sensation being reported by test subjects as their perceived level of fatigue increased. Table 3-47 provides the sample size (N) of each independent group (x_1 and x_2 corresponding to the test subjects featured within each ordered category of the temporal variable), the median (M_{dn}) and inter-quartile range (IQR) of the GRV(IES-GI) scores at each test sessions, the median difference (ΔM_{dn}) and its statistical significance (NHST, with Bonferroni-corrected *p*-value calculated with a one-tailed test), the mean ranks of the independent groups, the test statistic (U), and the effect size (r). The differences of GRV(IES-GI) detected at all sessions and for every category of fatigue are significant for all but 5 pairwise comparisons (to note that, only contrasts that are both statistically significant and practically relevant are reported in Table 3-47). The magnitude of the effects is generally substantive, with a practically relevant influence being detected in 13 out of 21 pairwise comparisons performed on the data.

Table 3-47 Pairwise comparisons between ordered categories of the variable Fatigue and effect sizes at each test session

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
Mid.	2 vs. 1	132, 84	22.31 (8.20)	18.83 (5.58)	3.48**	119.17	91.73	4,135.50	-0.21
Aft.	2 vs. 1	132, 60	19.18 (8.54)	16.50 (7.52)	2.68***	106.16	75.24	2,684.50	-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 at all 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

3.4.3.5. Food Intake

Figure 3-35 presents the boxplots related to the temporal variable ‘Food intake’. The graphical displays plot on the y-axis the calculated GRV(IES-GI) score based on the glare sensation votes provided by test subjects and, on the x-axis, the times of the day at which the glare assessments were given.

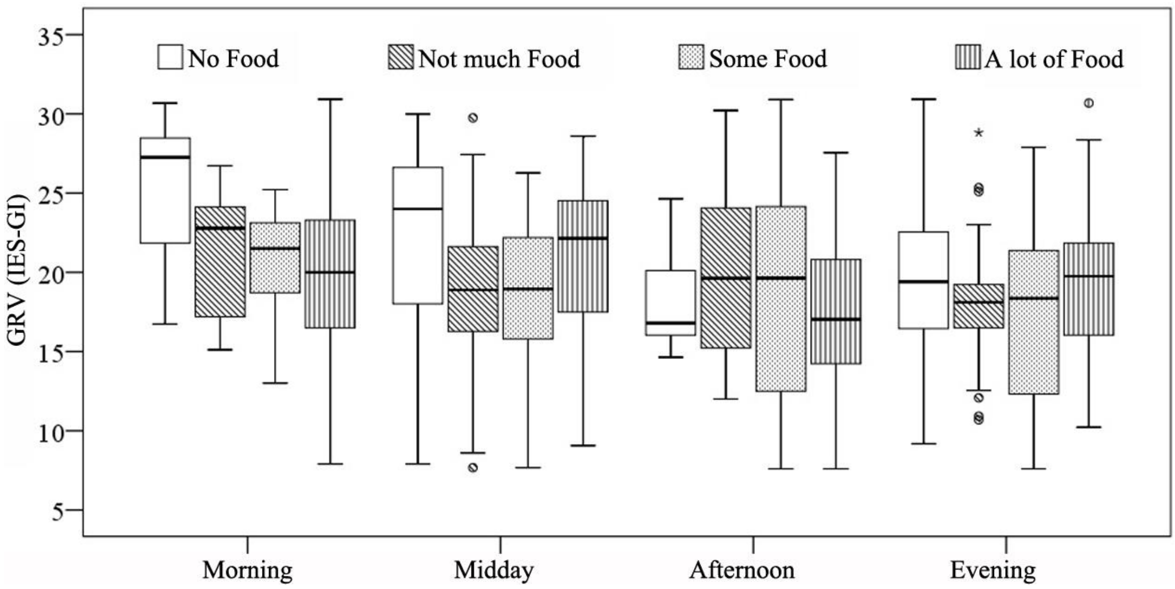


Figure 3-35 Boxplots of GRV(IES-GI) statistical parameters at each test session for the four ordered categories of the variable Food Intake

For this temporal variable, a non-directional two-tailed alternative hypothesis was adopted following preliminary 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= 3,676.00, p= 0.00***, r= -0.19) and a weakly significant influence in the Afternoon (J-T= 8,609.00, p= 0.03*, r= -0.14). This data led to the hypothesis that subjects who had had lower food intake reported higher levels of glare sensation, although the magnitude of the effects measured was, for both statistically significant influences, below the threshold of practical relevance. The results from the Kruskal-Wallis tests resulted in significant differences at all test sessions, with the exception of the Afternoon: Morning: $H(3)$= 12.49, p= 0.01**; Midday: $H(3)$= 22.54, p= 0.00***; Afternoon: $H(3)$= 6.10, p= 0.11 n.s.; Evening: $H(3)$= 9.96, p= 0.01**. The Mann-Whitney U tests suggested an inverse relationship between variables in the Morning session, as indicated by consistently negative values of the median differences (ΔM_{dn}). At this time of the day, three pairwise comparisons resulted in statistically significant and practically relevant differences (Table 3-48) coherent with the previous postulation that lower food intake resulted in higher votes of glare sensation. The tendency of an inverse relationship between glare response and food intake was apparent also in two statistically and practically significant comparisons at the Midday session, although this trend was not detected when analysing other permutations between ordered categories of this temporal variable.

Table 3-48 Pairwise comparisons between ordered categories of the variable Food Intake and effect sizes at each test session

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
Mid.	2 vs. 1	60, 60	18.88 (5.35)	24.00 (9.07)	-5.12***	49.12	71.88	2,947.00	-0.33
	3 vs. 1	72, 60	18.94 (6.45)	24.00 (9.07)	-5.06***	55.30	79.94	1,353.50	-0.32
	4 vs. 2	48, 60	22.41 (7.35)	18.88 (5.35)	3.53*	64.17	46.77	2,806.00	-0.28
	4 vs. 3	48, 72	22.41 (7.35)	18.94 (6.45)	3.47*	71.92	52.89	1,180.00	-0.27

Ordered Categories: No Food= 1, Not Much Food= 2, Some Full= 3, A lot of Food= 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

3.4.3.6. Caffeine Ingestion

Figure 3-36 presents the boxplots in relation to the ordered categories of ‘caffeine ingestion’.

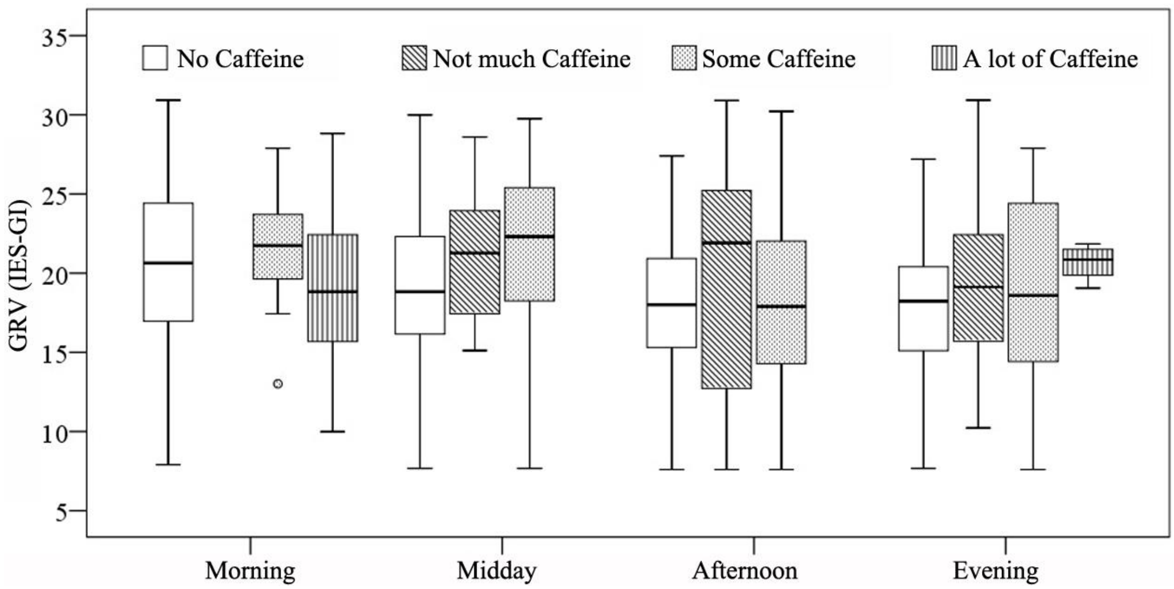


Figure 3-36 Boxplots of GRV(IES-GI) statistical parameters at each test session for the four ordered categories of the variable Caffeine Ingestion

Graphical inspection of the data revealed no prevailing directionality of central tendencies. Therefore, the alternative hypothesis supported either a direct or inverse relationship between this temporal variable and glare response. The Jonckheere-Terpstra tests showed evidence of two statistically significant direct trends at the Midday (J-T= 11,101.50, p= 0.001***, r= 0.22) and Evening (J-T= 10,607.00, p= 0.02*, r= 0.15) sessions, suggesting that higher caffeine ingestion led to an increase in glare response. For the latter session, however, the magnitude of the influence detected was of negligible size. The results from the Kruskal-Wallis test indicated significant differences between groups at all sessions except the Afternoon: Morning: $H(2)$= 9.14, p= 0.01**; Midday: $H(2)$= 11.68, p= 0.01**; Afternoon: $H(2)$= 3.63, p= 0.16 n.s.; Evening: $H(3)$= 8.87, p= 0.03*. Statistically significant and practically relevant differences were detected by the Mann-Whitney U tests in the Midday and Evening sessions (Table 3-49) that support the hypothesis of a direct relationship between caffeine ingestion and glare response. This trend was, however, reverted when considering the results of a significant and relevant comparison in the Morning session.

Table 3-49 Pairwise comparisons between ordered categories of the variable Caffeine Ingestion and effect sizes at each test session

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
Mid.	2 vs. 1	48, 108	21.27 (6.51)	18.83 (6.21)	2.44*	92.04	72.48	1,942.00	-0.20
	3 vs. 1	84, 108	22.31 (7.30)	18.83 (6.21)	3.48**	110.38	85.70	3,370.00	-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

3.4.3.7. Mood

Figure 3-37 presents the boxplots related to consideration of the temporal variable 'Mood'. Along the y-axis, the graph plots the GRV(IES-GI) calculated on the individual glare assessments provided by test subject and, on the x-axis, the four test sessions.

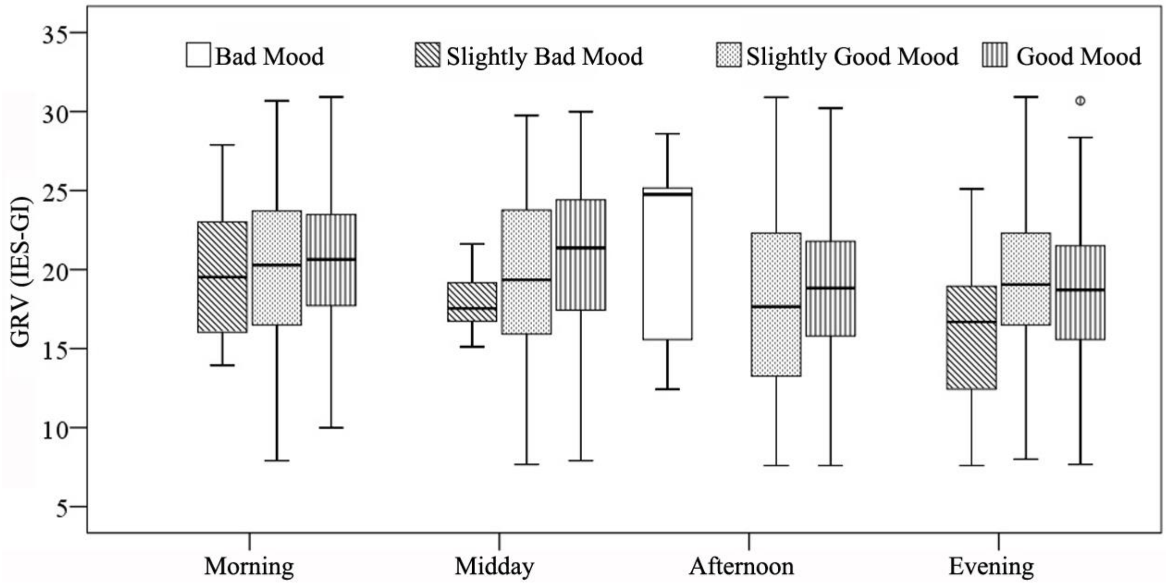


Figure 3-37 Boxplots of GRV(IES-GI) statistical parameters at each test session for the four ordered categories of the variable Mood

The displays reveal the lack of scores in some ordered categories of this variable. Graphical inspection of the data did not lead to any univocal interpretation of the distribution of statistical parameters, suggesting the adoption of a non-directional two-tailed alternative hypothesis. The Jonckheere-Terpstra tests signalled one statistically significant difference in the Midday session (J-T= 9,224.00, 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 Midday and in the Evening sessions: Morning: $H(2)$= 1.33, p= 0.51 n.s.; Midday; $H(2)$= 6.80, p= 0.03*; Afternoon, $H(2)$= 3.89, p= 0.14 n.s.; and Evening: $H(2)$= 11.47, p= 0.00**. Posteriori analysis was performed using the Mann-Whitney U tests, which detected 3 statistically significant and practically relevant pairwise comparisons (Table 3-50). The descriptive statistics for these significant and substantive contrasts suggest a direct relationship between variables (as shown by positive ΔM_{dn}), leading to infer that better mood might correspond to higher levels of glare sensation.

Table 3-50 Pairwise comparisons between ordered categories of the variable Mood and effect sizes at each test session

Session	Categories	N (x_1,x_2)	M_{dn} (IQR)	M_{dn} (IQR)	ΔM_{dn}^{NHST}	$MRank_{x1}$	$MRank_{x2}$	U	Effect Size (r)
Mid.	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	1,472.50	-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

3.4.3.8. Prior Daylight Exposure

Figure 3-38 presents the boxplots related to consideration of the 'prior daylight exposure' that subjects declared to have experienced before each test session.

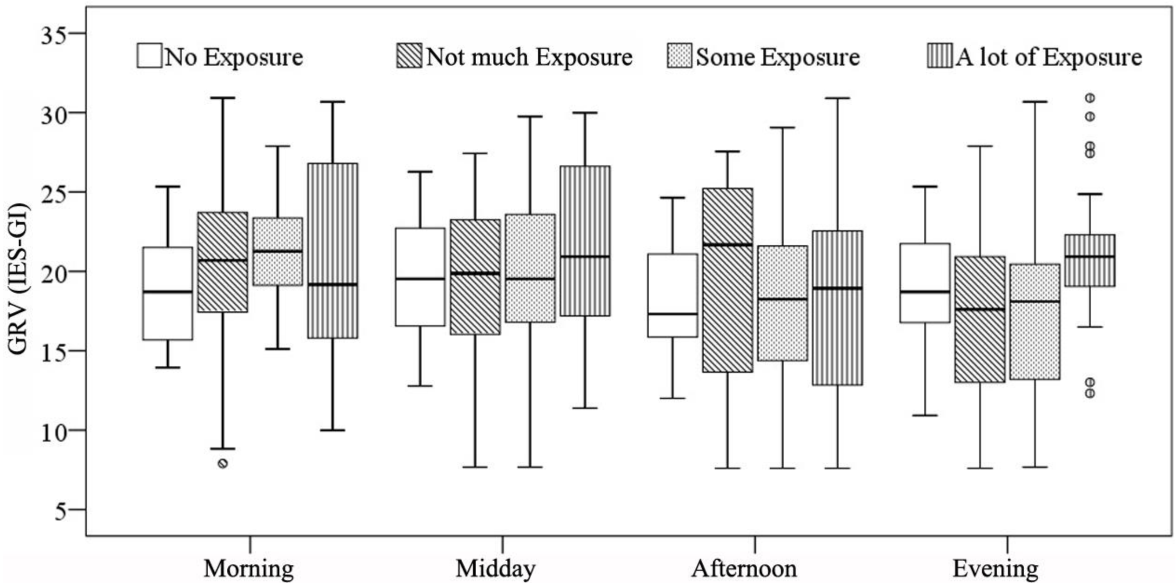


Figure 3-38 Boxplots of GRV(IES-GI) statistical parameters at each test session for the four ordered categories of the variable Prior Daylight Exposure

The graphical displays did not enable to identify any univocal trend of distributions of statistical parameters, this resulting in the adoption of an alternative (two-tailed) hypothesis supporting either a direct or inverse relationship between variables. The Jonckheere-Terpstra tests revealed evidence of one statistically significant effect at the Evening session (J-T=

11,902.50, p= 0.01**, r= 0.16), whose magnitude was however below the threshold of practical relevance. Also the Kruskal-Wallis tests showed significant differences between groups only in the Evening. For all other times of day, no significant difference across self-reported levels of 'prior daylight exposure' could be detected: Morning: $H(3)$= 5.50, p= 0.14 n.s.; Midday: $H(3)$= 4.55, p= 0.21 n.s.; Afternoon: $H(3)$= 1.38, p= 0.71 n.s.; Evening: $H(3)$= 22.11, p= 0.00***. Post-hoc Mann-Whitney U tests indicated that differences were statistically significant and practically relevant only in 3 cases at the Evening session (Table 3-51). For these comparisons, a significant and substantive direct trend was detected between glare response and levels of prior exposure to daylight. This suggests that, at this time of day, subjects who had been exposed to more daylight before the test session reported higher levels of glare sensation from the small diffusive screen under a constant source luminance. However, analysis of other comparisons did not lead to identify a clear univocal trend in the data emerging from central tendencies (medians) and inferential statistics (mean ranks).

Table 3-51 Pairwise comparisons between ordered categories of the variable Prior Daylight Exposure and effect sizes at each test session

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	1,002.50	-0.35
	4 vs. 3	48, 84	20.92 (3.25)	18.10 (7.17)	2.82***	85.43	55.68	1,107.50	-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

3.4.3.9. Sky Condition

Figure 3-39 presents the boxplots related to the prevailing 'sky condition' that participants reported before the start of each test session.

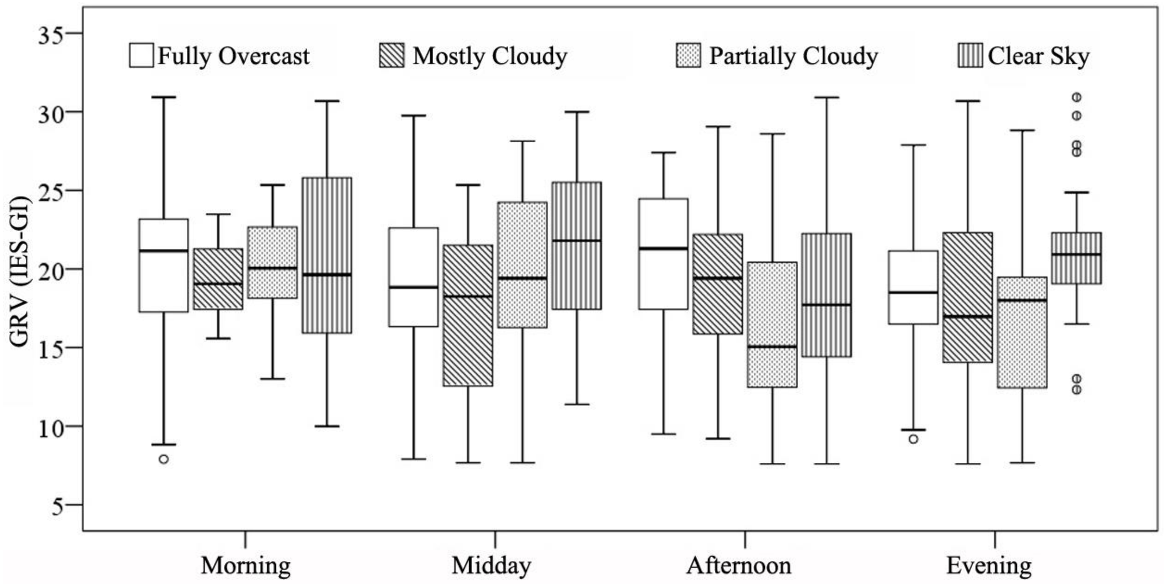


Figure 3-39 Boxplots of GRV(IES-GI) statistical parameters at each test session for the four ordered categories of the variable Sky Condition

Visual inspection of the data did not suggest any prevailing tendency between statistical parameters, hence supporting the adoption of two-tailed inferential analysis. The Jonckheere-Terpstra tests provided evidence of one statistically significant difference at the Midday session (J-T= 12,046.00, *p*= 0.01**, r= 0.19), although with an effect at the borderline of practical relevance. The Kruskal-Wallis tests led to the detection of significant differences at three test sessions: Morning: *H*(3)= 0.84, *p*= 0.83 n.s.; Midday: *H*(3)= 12.78, *p*= 0.01**; Afternoon: *H*(3)= 9.92, *p*= 0.02*; Evening: *H*(3)= 24.15, *p*= 0.00***. Post-hoc Mann-Whitney U tests were performed on the data to isolate the main effects. For the Evening session, three statistically and practically relevant differences were detected, suggesting a direct relationship between variables, with a higher glare response being given by subjects who experienced a clearer sky before the tests (Table 3-52). A similar trend was found for one significant pairwise comparison at the Midday session. However, two statistically significant and practically relevant inverse trends were detected in the Afternoon session, therefore hindering the possibility of inferring a consistent relationship between subjective glare response and self-reported prevailing sky condition.

Table 3-52 Pairwise comparisons between ordered categories of the variable Sky Condition and effect sizes at each test session

Session	Categories	N (x_1,x_2)	M_{dn} (IQR)	M_{dn} (IQR)	ΔM_{dn}^{NHST}	$MRank_{x1}$	$MRank_{x2}$	U	Effect Size (r)
Mid.	4 vs. 2	84, 24	21.79 (8.23)	18.24 (9.59)	3.55**	59.46	37.15	591.50	-0.30
Aft.	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	1,053.00	-0.33
	4 vs. 2	48, 72	20.92 (3.25)	16.97 (8.52)	3.95***	72.72	52.35	1,141.50	-0.29
	4 vs. 3	48, 48	20.92 (3.25)	18.00 (7.24)	2.92***	62.23	34.77	1,669.00	-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

3.4.4. Discussion

Previous laboratory experiments (Sections 3.1 and 3.2) identified statistically significant evidence of a substantive influence of time of day on glare response from a small source artificial lighting. This effect was found to be directly related to the independent variable itself (that is, the influence of time of day was found to be more substantive when considering a larger time gap between sessions), and not likely to be brought on by any influence of learning, although potentially confounded by other variables (e.g., fatigue). Further analysis (Section 3.3) detected significant and practically relevant influences of temporal variables and personal factors on the subjective evaluation of glare sensation as the day progresses.

The findings obtained from analysing the influence of visual task difficulty on glare response support the previous conclusion that the length of the interval between test sessions shows a direct effect on the increased tolerance to source luminance along the day. Moreover, inferential tests also 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 visual task difficulty.

To investigate more comprehensively the relationship between visual task difficulty, temporal variation in glare response, and time of day, Figure 3-40 plots – for each task reported in the contrasts of Table 3-45 (Tasks 3 to 12) – the effect sizes (d) detected from the pairwise comparisons between the Morning vs. Midday (left), Morning vs. Afternoon (centre), and Morning vs. Evening test sessions (right).

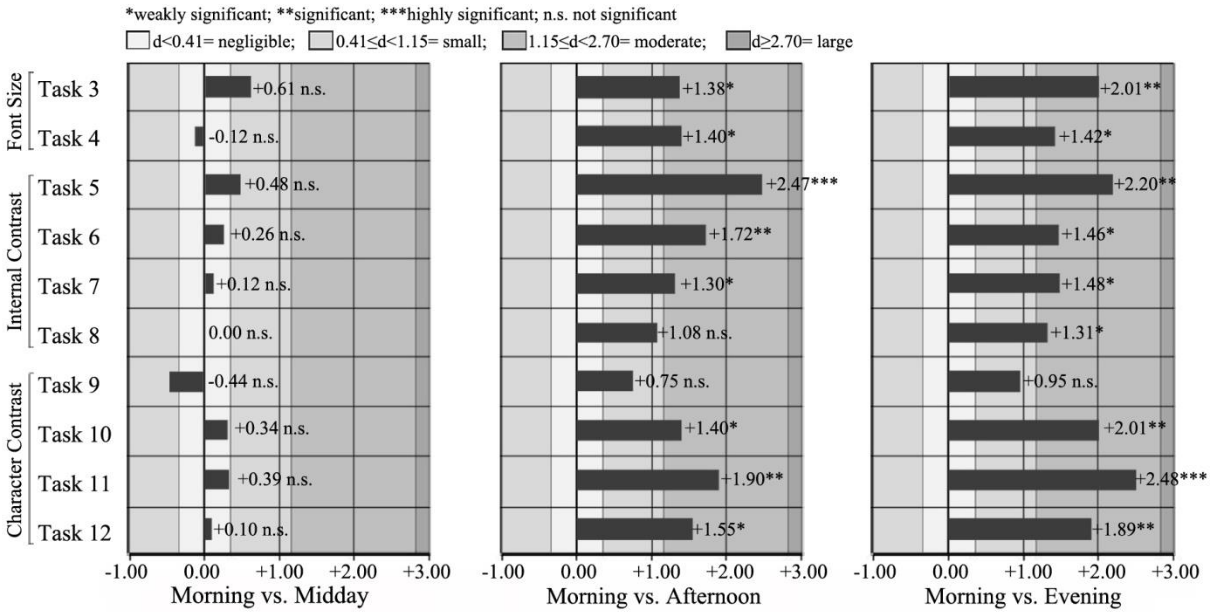


Figure 3-40 Effect sizes (d) detected in the pairwise comparisons (see Table 3-45) between test sessions (x-axis) for each group of visual tasks (y-axis)

Although all pairwise comparisons between the Morning and Midday sessions (corresponding to an interval of 3 hours between tests) were not statistically significant, the graphical displays reveal a tendency for the magnitudes of effect size to amplify with the increase in the temporal gap between test sessions. The rate of variation of the effect sizes is, however, not consistent between different visual tasks and across comparisons between sessions. This suggest a moderating effect of task manipulation and visual task difficulty influencing the effect that time of day has on glare response. The only task showing a consistent 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. The range of variation of effect sizes between test sessions is broadly in line with what had been measured in previous experiments (see Section 3.1.3) for the 'Just Uncomfortable' equivalent reported level of glare sensation. On the Hopkinson scale (Hopkinson, 1971), in fact, this GSV criterion corresponds to the constant luminance of the diffusive screen that was set for studying the influence of visual task difficulty on glare response (10,541.31 cd/m^2). As an example, when converting the previously measured effect sizes (Table 3-6) into an equivalent estimator (Cohen's d), the magnitude of the effect of time of the day on glare sensation between the Morning and the Afternoon session returns a similar value of $d= +2.27$.

With reference to individual task manipulation, the tasks that varied in 'internal contrast' (Tasks 5 to 8) show a tendency, for each comparison between sessions, for the effect size to decrease when increasing the difficulty of the visual task. This suggests that tasks that were rated by the independent sample group as more difficult – i.e., the information contained within the stimulus was harder to extract – correspond to lower effect sizes as the day progresses. Conversely, the tasks that varied in 'character contrast' (Tasks 9 to 12) show an opposing trend, with the effect size amplifying as the tasks become more visually difficult. In this group, an exception to this tendency is represented by Task 12, which was rated with the highest task difficulty of 4.00 (0.00). No univocal trend could be identified for the two tasks featured in the group varying 'font size' (Tasks 3 and 4), which were both rated with a high level of visual difficulty, respectively 3.00 (0.00) for Task 3 and 4.00 (0.00) for Task 4.

This supports 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 across the day. This is broadly in line with existing research, which has found that lighting preferences (i.e., luminance distribution, illuminance levels, etc.) are task-dependent (Veitch, 2001). In addition, the data suggest that, for highly demanding (difficult) visual tasks, the effect of time of day may be masked or confounded by task difficulty or by other temporal influences.

In consideration of temporal variables, Figure 3-41 provides a comprehensive visualisation of the variation of effect sizes (r), extracted from the Jonckheere-Terpstra tests, at each time of day. These tests were performed to detect ordered trends between categories of all temporal variables considered across the dependent variable (GRV(IES-GI)), and evaluate the magnitude and the sign of the trends as the day progresses (Field, 2013). For all tests, an alternative two-way 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 this temporal variable, in fact, initial graphical inspection of the data had suggested a prevailing univocal trend, hence leading to a one-way hypothesis (i.e., as fatigue intensifies so does the glare response reported by test subjects).

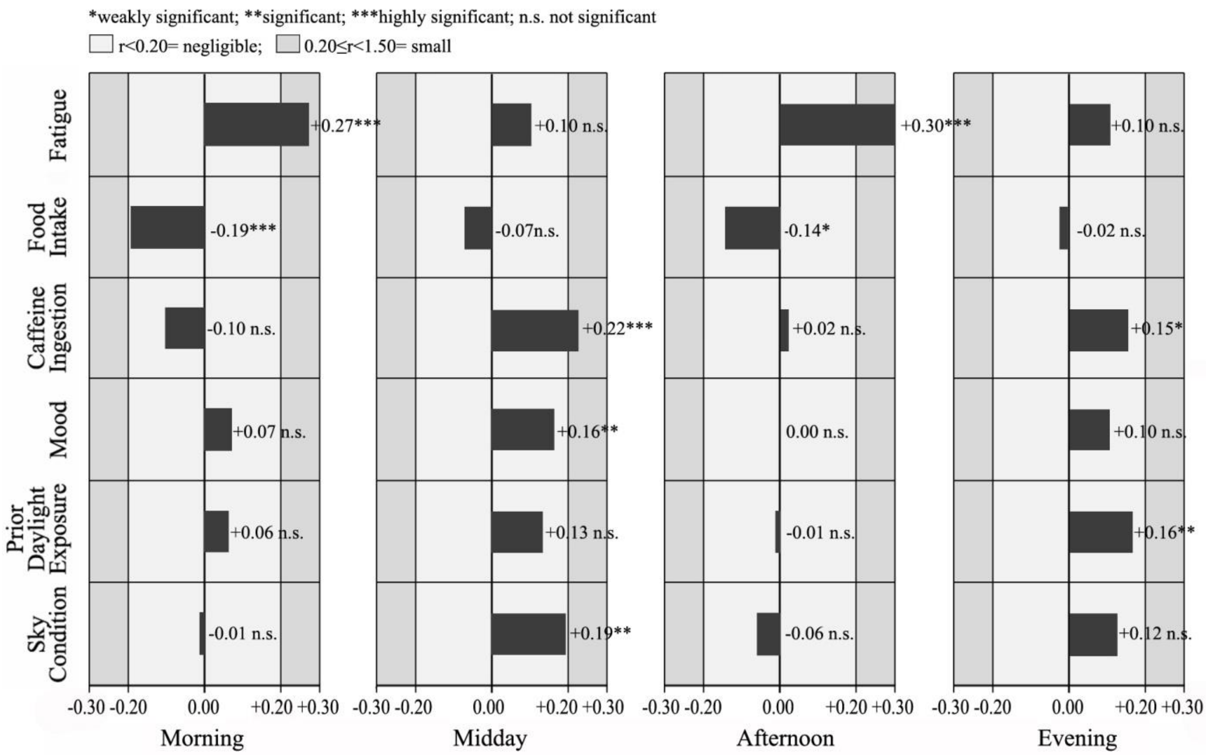


Figure 3-41 Effect sizes (r) extracted from the Jonckheere-Terpstra tests at each session (x-axis) and for all temporal variables (y-axis)

Consideration of self-assessed ‘fatigue’ shows the strongest effect on variation of glare response between all variables considered, which is identified by the largest magnitudes of effect sizes. The influence of fatigue is particularly evident in the Morning (r= +0.27***) and Afternoon (r= +0.30***) sessions, with a consistent direct relationship between reported levels of fatigue and glare sensation along the day. This finding is in line with earlier results (Section 3.1 and 3.3), although the absence of a visual task in previous experiments did not make it possible to identify a univocal statistically significant and practically relevant effect.

Analysis of ‘food intake’ suggests a consistently inverse influence on glare response. Findings demonstrate statistically significant relationships in the Morning (r= -0.19***) and Afternoon sessions (r= -0.14*), although trends were marginally not substantive in effect size.

A significant and practically relevant direct effect of ‘caffeine ingestion’ on glare sensation is evident in the Midday (r= +0.22***) and, to a lower extent, Evening (r= +0.15*) sessions. This is consistent with the results of previous experiments (Section 3.3.3.3) that had detected higher tolerance to source luminance for test subjects not having ingested caffeine, with this effect being less pronounced in the afternoon.

A similar direct relationship with reported levels of glare sensation is also noticeable for consideration of the temporal variable ‘mood’. This tendency is statistically significant in the Midday session (r= +0.16**), although at a magnitude that is marginally below the threshold of practical relevance (r≥0.20).

For the variables ‘prior daylight exposure’ and ‘sky condition’, the tests did not lead to the identification of 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 is detected in the Evening (r= +0.16**). Conversely, a significant direct relationship between the self-assessed prevailing sky condition and reported glare sensation is recognisable in the Afternoon session (r= +0.19**).In interpreting these outcomes, it is reasonable to hypothesise that some temporal variables may not be independent from one another, interacting and influencing each other along the day. In fact, the statistically significant and practically relevant direct effects of fatigue on glare response in the Morning and Afternoon sessions seem to complement the inverse relationship between food intake and visual response at these times of day. That is, subjects who declared to have ingested lower amounts of food might have also given a self-assessment of greater fatigue, this resulting in a higher glare response. Consideration of fatigue and food intake did not lead to the detection of any statistically significant effect on glare response in the Midday and Evening sessions. However, at these times, a significant and practically relevant direct influence of caffeine ingestion was detected on glare sensation.

In essence, it would be plausible that higher levels of caffeine ingestion or the lack of food intake may act as personal stressors that potentially influence (i.e., confound or amplify) the self-assessment of fatigue, this resulting in a lower tolerance to source luminance and higher reported levels of glare sensation. This is in line with research in the psycho-physiological and behavioural sciences that, among other aspects, has detected anxiety trait and long hours of work to be associated with heightened central arousal and sensitisation to luminous stimuli (Emdad *et al.*, 1998). Recent studies have also shown that caffeine ingestion and the effects of light exposure might have a dependent relationship, since caffeine has been found to influence human circadian timing and delay the melatonin rhythm (Burke *et al.*, 2015). Similarly matching patterns of variation of effect sizes could be identified when considering the temporal variables of mood and, particularly, prior daylight exposure and sky condition.

According to the results of this experiment, the direct effect of time of day on glare response appears to be related to visual task difficulty (Sections 3.4.3.1 and 3.4.3.1) and to the influence of several temporal variables (Section 3.4.3.4 to Section 3.4.3.9). Therefore, it can be assumed that, if these variables are not controlled when testing the research hypothesis of this investigation under daylit conditions (i.e., when recording votes of glare sensation from a window), they could contribute to the scatter commonly associated with subjective evaluation of discomfort glare and confound the detection of the effect of experimental interest.

As a consequence, the second stage of this investigation – described in the following sections – explored the research hypothesis in a naturally lit environment with direct access to daylight and a view to the outside, while experimentally controlling (where feasible) the variables isolated in previous laboratory experiments. To this purpose, the effects of visual task difficulty needed to be ‘masked’ from the experimental design, while – when it was impractical to control other variables (i.e., fatigue, prior daylight exposure, etc.) that intrinsically vary over the independent variable – these were measured separately and statistically controlled. This allowed to remove from the analysis the unknown variability related to these confounding temporal variables, reducing the distortion associated with these “nuisance” factors (MacKinnon *et al.*, 2000), and enabling a more accurate assessment of the effect of time of day on glare response within a less controlled environment (Field, 2013).

Chapter 4

Test Room Studies

4. Test Room Studies

4.1. Time of Day and Discomfort Glare from Daylight

The second stage of this investigation aimed at searching for an effect of time of day on reported levels of glare sensation from daylight. Previous findings from a series of controlled laboratory experiments detected greater tolerance to artificial light source luminance as the day progresses (see Section 3.1, Section 3.2, and Section 3.3). The results showed that this influence was directly linked to the independent variable itself (i.e., the effect of time of day on glare response amplified when considering a larger interval between test sessions). However, in interpreting these outcomes, it must be considered that the glare stimulus used (i.e., the luminance of a small diffusive screen mounted in front of a projector connected to a computer) was not dependent on many of the characteristics that are generally associated with a window and daylight (i.e., orientation, dynamics, view, etc.). Hence, the experiment described below – conducted in a test room with direct access to daylight – was designed primarily for replicating the original findings from the laboratory studies. This was to verify whether, once removing the influence of confounding variables from the analysis, the same effect could be detected under less controlled conditions (Tuaycharoen & Tregenza, 2007).

4.1.1. Introduction

A previous series of experiments conducted under a laboratory setting revealed a tendency towards greater tolerance to luminance increases in artificial lighting as the day progresses. This trend was found not to be statistically related to the confounding variable of learning, leading to postulate an effect of time of the day on glare response (discussed in Chapter 3). In the experiments, all known variables that could potentially influence glare sensation (i.e., source and background luminance, source size, position index, uniformity of the source, etc.), along with personal factors (e.g., age, education, etc.), were controlled. Although ethnicity and cultural differences were not strictly controlled in subject sampling, participants were requested to provide this information at the start of each test, so that these factors could be monitored and their potential influence on glare response be subsequently analysed.

Coherent with the literature (e.g., Tregenza and Wilson, 2011), when source luminance levels for each vote of glare sensation provided by test subjects were regressed against the times of sessions, a large scatter of data was observed. This suggested that there could be other factors varying with time of the day, not controlled through experimental manipulation over the independent variable, which could influence the subjective evaluation of glare sensation.

Among these variables, significant and substantive evidence was found of greater tolerance to artificial lighting increases for earlier chronotypes and for subjects not having ingested caffeine (Section 3.3). A further series of laboratory experiments, conducted while subjects completed visual tasks of various difficulties, confirmed the postulated increase tolerance to source luminance as the day progresses. The effect of time of day on glare response appeared to be dependent on the level of visual discomfort experienced and to be influenced by task difficulty and manipulation of the information of interest contained therein (Section 3.4).

The analysis also detected a significant and substantive direct influence of fatigue and caffeine ingestion on glare response, and an inverse effect of food intake on visual discomfort. Some significant and relevant evidence was found of a direct influence on glare response of mood, prior daylight exposure, and sky condition, although the scatter of the data did not allow inferring a convincing relationship with subjective visual perception. These results led to hypothesise that some temporal variables may interact with each other over times of the day, and significantly influence the variation of glare response as the day progresses.

4.1.2. Methods

4.1.2.1. Experimental Design and Procedure

4.1.2.1.1. The Test Room

To investigate the influence of time of day on glare sensation in the presence of daylight, an experiment was designed based on the use of a test room with access to direct sunlight and with a view to an external natural scene. The experiment was set up in order to allow as many potentially confounding variables as possible to be controlled (Figure 4-1).

Figure 4-1 The test room at the University of Nottingham (front and rear)

The test room was located at the Department of Architecture and Built Environment, The University of Nottingham, UK (latitude: 52° 56' 19'' N; longitude: 1° 11' 42'' W). The internal layout of the test room and the design of the experimental procedure were informed by the studies conducted by Christoffersen & Wienold (2008), Kuhn *et al.* (2013), and Wienold & Christoffersen (2006). The room had internal dimensions of 3.45 m x 2.55 m and an internal ceiling height of 2.35 m. It featured a South-East facing window (azimuth angle= 165°) of 0.87 m width and 1.47 m height, with a sill height of 0.7m (Figure 4-2).

Figure 4-2 Internal view of the test room

The internal walls of the room had the following reflectance properties: ρ_{wall}= 0.6, $\rho_{ceiling}$= 0.8, ρfloor= 0.2. The reflectance properties were estimated using the colour of the surfaces (walls, ceiling and floor) and the Munsell value (a value representative of a surface's hue, value (lightness), and chroma), as per Equation (4-1)) (Tregenza & Loe, 1998):

$$\rho \approx \frac{Munsell\,Value \cdot (Munsell\,Value - 1)}{100} \quad \text{(Equation 4-1)}$$

For this experiment, the window was equipped with an internal user-controlled white aluminium Venetian blind with estimated reflectances of ρ= 0.90 (upper surface) and ρ= 0.72 (lower surface). Each slat of the blind was convex in shape, with dimensions of 110 cm x 2.5 cm and a distance of 2.5 cm between each slat. The shading device was mounted on the internal wall above the window head at a distance of 11.5 cm from the glazing (Figure 4-3).

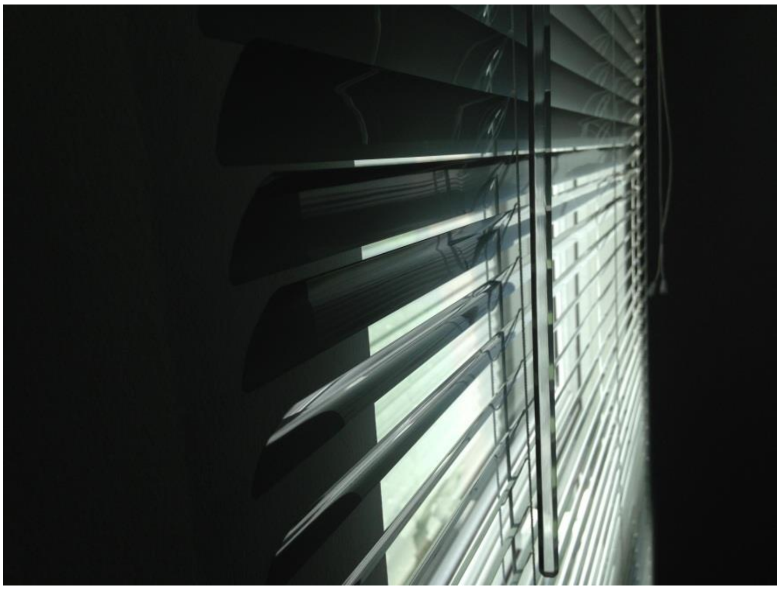


Figure 4-3 Internally-mounted Venetian blinds

The operation of the blind enabled to block, if and when necessary, direct sunlight penetration inside the room at different times of the day. A workstation (desk, office chair, and desktop computer) was placed inside the room at a 45° position from the peripheral walls, with the top corner of the desk located at a distance of 30 cm from the window (Figure 4-4). The surface of the desk had a reflectance of ρ= 0.42, and had dimensions of 120 cm x 60 cm and a height of 72 cm from the floor. A flat screen 19” iiyama ProLite B19065 liquid crystal display (LCD) (mean self-luminance= 201.64cd/m^2) was used as the Visual Display Unit (VDU) to present the visual tasks that were used as part of the experimental procedure.

A diagonal arrangement of the workstation was selected for this experiment instead of a desk positioned parallel or perpendicular to the window wall. In fact, previous glare studies conducted with parallel or perpendicular layouts found that, when subjects were requested to provide a glare assessment, they might often deviate their sight from the display and look directly at the window, while photometric measuring instruments would still capture the luminous condition of the VDU (Wienold & Christoffersen, 2006; Wienold, 2009). Conversely, an arrangement of the workstation positioned 45° clockwise from the window helped to mitigate the risk of unwanted head movements between the PC monitor display (VDU, task area) and the window when glare assessments were collected (Figure 4-4).

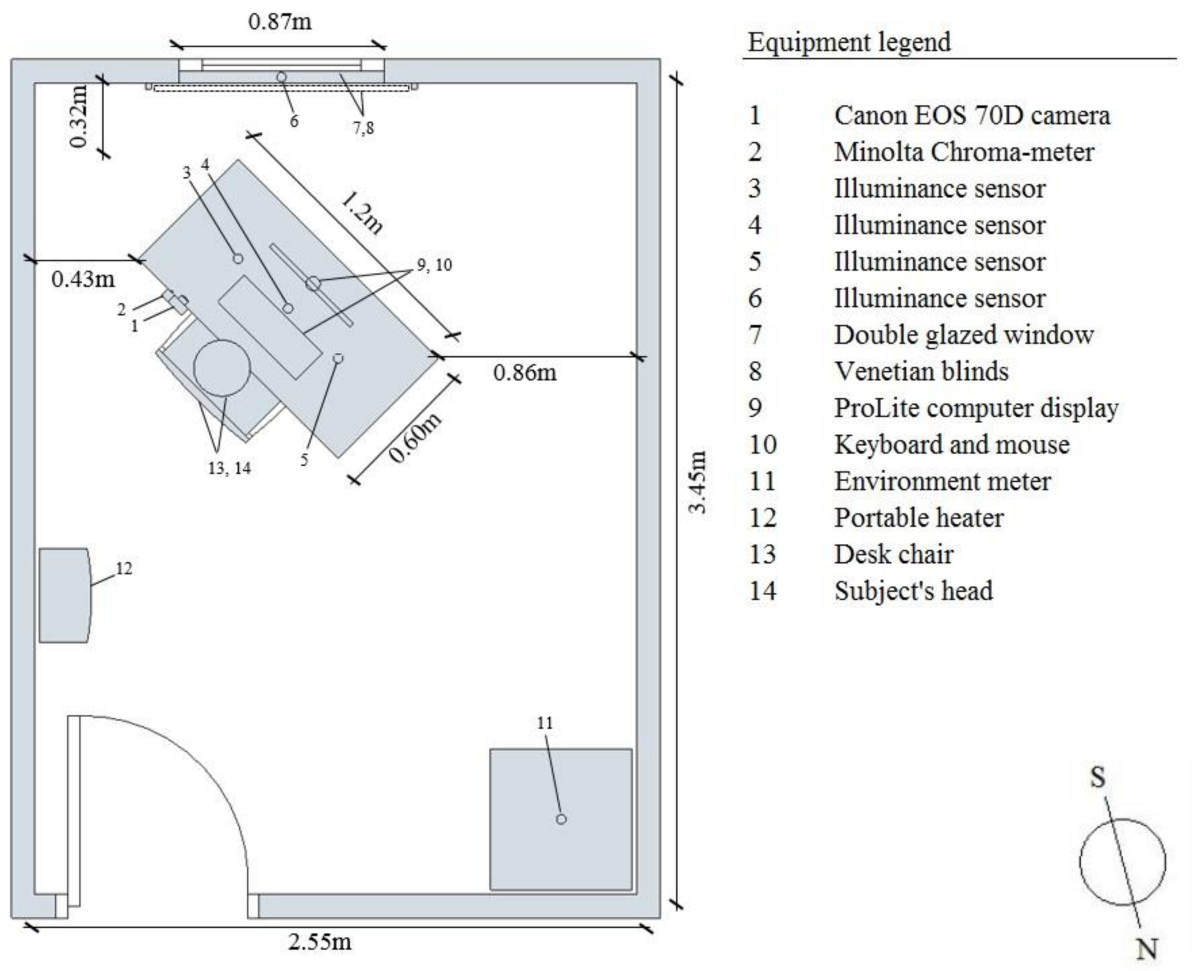


Figure 4-4 Layout of the test room and list of equipment

Figure 4-5 Position of the test subject during glare assessments

4.1.2.1.2. Photometric Measurements

For this experiment, three photometric instruments were used to ‘instantaneously’ capture the luminous environment of the observer: 1) a Charged-Couple Device (CCD) camera equipped with a fish-eye lens; 2) an illuminance chromameter; and, 3) a series of horizontal illuminance sensors connected to a datalogger. The camera and illuminance chromameter were mounted on the desk adjacent to the position of the subject’s head and were pointed towards the VDU, which was assumed to be the subject’s visual fixation area (Figure 4-6).

Figure 4-6 Setup of the CCD camera and illuminance Chromameter

The CCD camera was a Canon EOS 70D equipped with a 4.5 mm f/2.5 EX DC GSM 180^{o} Sigma fish-eye-lens mounted on a Monfrotto extendable arm. CCD cameras can utilise conventional photographic techniques to derive photometric values; in other words, photographs captured from CCD outputs can be used for luminance measurements (Inanici, 2006). The CCD camera allows a large range of luminance values to be quickly captured. These luminance values are contained within the image pixels corresponding to different points of measurement in the captured image (Bellia *et al.*, 2002). To obtain full range of luminance measurements, seven independent Low Dynamic Range Images (LDRI) were taken with the camera, varying exposure values (EVs) (Table 4-1) to capture the wide luminance variation within the field of view (FOV) (Inanici, 2006). The quality of the images taken with the CCD camera is dependent on both the aperture (*f*/N, whereby the f-number is the ratio between the lens's focal length and the entrance diameter) and the time during which the shutter is open (exposure time, (v)). At a constant focal length, the aperture becomes proportional to the square of its value ($1/f^2$). The combination of both settings is often referred to as image quality, or exposure value, and is proportional to v/f^2. Therefore, if the sensitivity (ISO) and gain (whereby the gain is the ratio between the number of photoelectrons received by the CCD and the number of pixels within the captured image) of the camera sensor are constant, the quality of at the image produced by the CCD is solely proportional to v/f^2 (i.e., the exposure time and aperture) and can be fully expressed using (Equation 4-2) (Coutelier & Dumortier, 2002):

$$EV = 3.32\ log_{10}\left(\frac{f^2}{v}\right) \qquad \text{(Equation 4-2)}$$

Table 4-1 presents the corresponding camera settings for each of the seven independent LDRI images used for this experiment. For each column, the table presents the aperture, exposure time, sensitivity, and exposure value calculated utilising Equation 4-2.

Table 4-1 CCD camera settings for the seven independent LDRI images

Image	Aperture (*f*/N)	Exposure Time (1/s)	Sensitivity (ISO)	Exposure Value (EV)
1	2.8	1/10	400	6.33
2	2.8	1/100	400	9.67
3	2.8	1/500	400	12.00
4	2.8	1/1000	400	13.00
5	2.8	1/2000	400	14.00
6	2.8	1/4000	400	15.00
7	2.8	1/8000	400	16.00

The LDRI images taken with the Canon EOS 70D camera were combined into a Radiance-formatted High Dynamic Range Image (HDRI) using the Photosphere software (Figure 4-7) from Anyhere Software (2005)[1]. Photosphere is a 'data fusion' software that combines

[1] The Photosphere software is available at: http://www.anyhere.com/

several independent LDRI into a single HDRI (Cai & Chung, 2011). The HDRI images could then be evaluated using the *Evalglare* tool (Wienold, 2009) available from the Fraunhofer ISE (2015)[2]. This software is a command line-based tool that uses HDRI images from predominantly daylit scenes to evaluate discomfort glare (*Evalglare*, 2004).

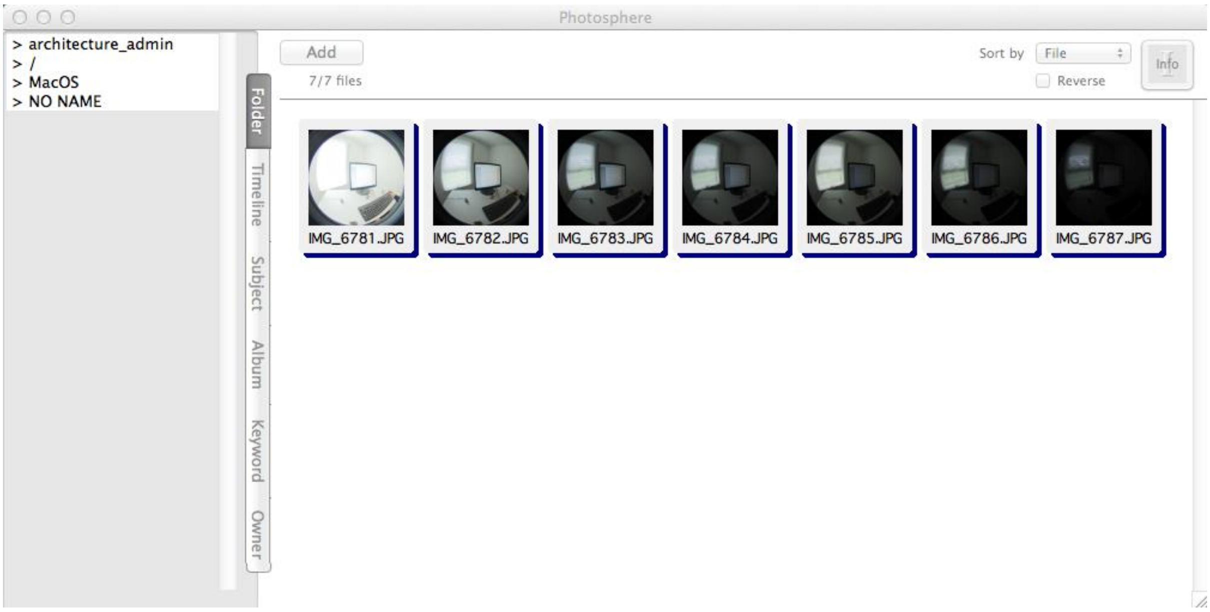


Figure 4-7 Combination of the seven independent LDRI images into a Radiance-formatted HDRI image using the Photosphere software

Once the LDRI images were combined, the camera response function – a regression curve demonstrating the relationship between a luminance value and a given pixel located within the scene image – was computationally derived through a self-calibration process via the use of a Minolta LS-100 luminance meter (Inanici, 2006). A single spot-point measurement under stable sky conditions was taken, which was directly applied to the HDRI previously obtained (Figure 4-8). Once the self-calibration process for a single HDRI was finalised, the response curve for the Canon EOS 70D camera with 4.5 mm f/2.5 EX DC GSM Sigma fish-eye-lens

[2] The *Evalglare* tool developed by Dr Jan Wienold is available at: http://www.ise.fraunhofer.de/de/geschaeftsfelder/energieeffiziente-gebaeude/themen/lichttechnik/fue-leistungen/lichtsimulation/radiance

was saved by Photosphere. Therefore, the calibrated response curve for the given camera and lens could be applied to any subsequent HDRI image produced by Photosphere.

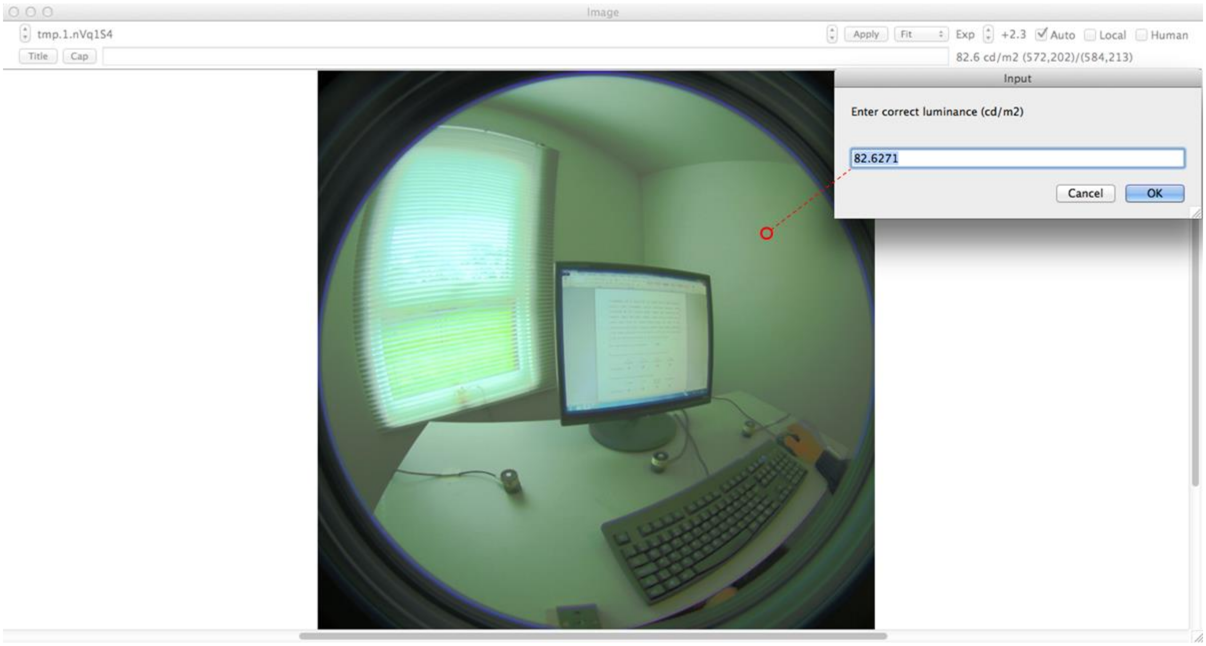


Figure 4-8 Self-calibration process of the HDRI image using a single-point luminance measurement

The second photometric instrument used in this experiment was a Minolta Chromameter CL-200a mounted vertically on the desk adjacent to the camera (Wienold & Christoffersen, 2006; Wienold, 2009). This was used to independently take vertical illuminance measurements and evaluate the photometric integrity of each luminance image. To check for integrity, the vertical illuminance of each luminance image taken by the CCD camera needed to be calculated. Therefore, the total amount of light entering the camera lens and the light reaching the illuminance meter's sensor had to be compared. The vertical illuminance corresponding to each luminance image was calculated using *Evalglare* (Equation 4-3) (Wienold, 2009).

$E_{vertical} = \sum_i L_i \cdot \omega_i \cdot \cos\theta_i$ (Equation 4-3)

In the equation, the vertical illuminance is weighted by the pixel luminance (L), the solid angle (ω), and the cosine of the angle (θ) between the view direction (the centre of the image) and the actual pixel considered. Both the camera and the illuminance meter were mounted on adjustable arms so that they could be positioned as close as possible to the observer's head without causing visual impairment or distraction to the subject (Kuhn *et al.*, 2013).

The third photometric instrumentation used in this experiment was represented by three calibrated horizontal illuminance sensors that were distributed evenly at a distance of 20 cm from each other on the desk surface (Figure 4-9), and one horizontal illuminance sensor that was placed centrally on the window sill, positioned between the internal venetian blind and the window glazing. These sensors were connected to a DataHog 2 data-logger and recorded horizontal illuminance every 10 seconds (Wienold & Christoffersen, 2006).

Figure 4-9 Position of the three horizontal illuminance sensors on the desk

In addition to these photometric instruments, a spot-point CEM DT-8820 Environment Meter was used to record dry bulb temperature, relative humidity, and sound levels before and after each test session. A portable electric heater was also placed in the test room to help maintain a constant temperature during the testing period (February to April).

4.1.2.1.3. Experimental Procedure

The experimental procedure required subjects to participate to three test sessions, whose order was randomised over three consecutive days, evenly distributed at 3-hour intervals:

- Morning: 09:00 or 09:30
- Midday: 12:00 or 12:30
- Afternoon: 15:00 or 15:30

During each test session, subjects were asked to perform two series of three visual tasks; each series was completed under a different shading setting (one ‘default’ and one ‘user-set’), for a total of six visual tasks (Wienold, 2009). The procedure was coherent with the laboratory tests described in Section 3.1.2.2, Section 3.2.2, and Section 3.3.2.1, although the evening session (18:00 or 18:30) was excluded from this experiment due to seasonal variations in day length and sunset times (these ranged between 16:43 and 19:08 throughout the testing period).

Considering that discomfort glare is a personal sensation that requires subjective evaluation methods (Velds, 2002), consistent with previous laboratory experiments, during the tests subjects were asked to make judgements of glare sensation utilising as benchmarks the adaptations of Glare Sensation Votes (GSVs) used by Iwata *et al.* (1992a; 1992b), Iwata & Tokura (1998), and Mochizuki *et al.* (2009). These glare criteria correspond to the sensation of visual discomfort experienced by subjects: ‘Just (Im)Perceptible’, ‘Just Noticeable’, ‘Just

Uncomfortable', and 'Just Intolerable'. Since it was considered that each criterion may be open to self-interpretation – this potentially influencing further scattering in responses – to aid the understanding of each criterion, GSVs were linked to time-span descriptors (Tuaycharoen & Tregenza, 2007; Velds, 2002). In addition, for this experiment it was also considered that natural light conditions continuously change over time (Wienold & Christoffersen, 2006), and therefore it may be plausible that the glare sensation perceived by test subjects could fall below the minimum criterion of 'Just (Im)Perceptible' (Fotios, 2015). As a consequence, the time-span descriptor for 'Just (Im)Perceptible' was modified slightly with respect to previous laboratory tests, and was described as the point where the window is bright but not necessarily giving a sensation of glare (MacGowan, 2010).

Whilst performing each glare assessment, subjects were also asked to make judgments of how satisfied they were with the amount of natural light that fell onto the desk and how important they considered a view to the outside. Both satisfaction with horizontal desk illuminance and view importance were evaluated using 4-point scales. For desk illuminance, the points corresponded to the perceived levels of satisfaction of: 'Too Low'; 'Somewhat Satisfied'; 'Very Satisfied'; and, 'Too High'. For view, the points corresponded to the following perceived levels of importance: 'Undesirable'; 'Indifferent'; 'Somewhat Important'; and, 'Very Important'. Both scales were adapted from Wienold (2009).

Before the subjects entered the test room, the venetian blinds were set at a default 'cut-off' position; that is, blinds were arranged with a slat tilt that ensured predominantly diffused daylit conditions inside the test room but still permitting some view to the outside from the position of the test subject (Figure 4-10). This was to ensure that no direct sunlight was present in the field of view of the observer during the first part of the test. The cut-off angle was set at an offset ranging from 5-10^{o} in response to external conditions. This was measured manually using a protractor at the start of every session (Wienold & Christoffersen, 2006).

Figure 4-10 Internal view of the test room with the blinds set at their default position

At the beginning of the test procedure, subjects were required to position themselves at the desk facing the computer screen (the subject's head being at a distance of around 1.5 m from the window). A detailed set of instructions was then given, including a definition of discomfort glare, the meaning of each GSV criterion, and a description of how the session would run. At this point, test subjects were asked to fill in a short questionnaire featuring demographic information (age, gender, ethnicity), and requesting participants to provide self-reports of personal factors (chronotype, photosensitivity), and temporal variables (fatigue, hunger, caffeine ingestion, mood, prior light exposure, sky condition).

A brief pre-test period under the default shading setting was then provided for the subject to adjust to the luminous conditions, comprehend the experimental setup, glare scale, etc., and ask questions to the experimenter. On their first test, subjects were also required to perform a series of 'trial' abbreviated visual tasks to familiarise themselves with the test procedure.

The first trial task consisted of a 'landolt ring' pre-test, whereby subjects were requested to look at a chart and count the number of rings that had a gap in a specified orientation (top, bottom, left, or right) (Boyce, 2014; Kuhn *et al.*, 2013). In the following 'letter searching' pre-test, subjects were asked to look at a short pseudo-text, featuring random letters and numbers, and locate specific characters within it (see examples in Appendix D). Finally, the 'typing' pre-test consisted of a short pseudo-text paragraph that had to be manually typed in a space on the computer screen (Kuhn *et al.*, 2013; Wienold & Christoffersen, 2006; Wienold, 2009). All trial tasks were presented on the Visual Display Unit (VDU).

Following the completion of each task, using a GSV scale displayed on the screen, subjects were asked to indicate the perceived glare sensation given by the daylight coming from the window (Wienold, 2009; Velds, 2002; Wienold & Christoffersen, 2006). At this point, the experimenter collected a series of seven LDRI images with varying exposure values, as well as a single vertical illuminance measurement taken by the Chromameter. All data from the pre-test were recorded but were not included in the analysis. The pre-test stage was followed by a brief relaxation period (~1-2 minutes), whereby any further questions regarding the test could be clarified. At the end of the pre-test, the full experimental procedure started.

Unlike the trials, the visual tasks presented to test subjects during the experimental stage were not abbreviated and sought to provide an evaluation of visual discomfort while subjects fully engaged with each task. The same visual tasks as the pre-test procedure – 'landolt ring', 'letter searching', and 'typing' – were used, and were presented to each subject under a randomised sequence, while the shading device (in the first part of the experiment) was still maintained at a default ‘cut-off’ position. In terms of the selection of visual tasks to be used for this study, it was considered that, since the experiment used a repeated-measures design, unwanted carry-over effects (learning) could occur if normal text (i.e., newspaper articles) were used (Field & Hole, 2013). Moreover, providing letter searching and typing tasks with content that was both independent from each other and homogenous enough for analysis in a within-subjects study was not considered feasible. Instead, pseudo-texts featuring random letters, with upper and lower case, and numbers were used (Wienold, 2009). This prevented repetition of identical stimuli, while retaining the experimental integrity needed for statistical analysis (Roufs & Boschman, 1997; Boschman & Roufs, 1997). Nevertheless, to counterbalance any unlikely possibility of carry-over effects, the content of the pseudo-texts were randomised for each visual task without directly changing the content of the stimuli.

In this context, the literature suggests that the perceived level of glare sensation may be influenced by a variation in task difficulty due to a change in size, contrast, or background of the text (Tregenza & Loe, 2014; Sivak *et al.*, 1989). This postulation was confirmed by previous laboratory experiments (see Section 3.4). Therefore, the effect of visual task difficulty was controlled by presenting font characters always set at Arial, with a size of 12 points, and at double line spacing.

The experimental procedure followed the same methodology of the pre-tests, with subjects performing the three visual tasks, expressing their vote of glare sensation, and photometric measurements being taken after each assessment. Once the first series of three tasks was concluded under the ‘default’ shading, subjects were asked to modify the shading setting from its cut-off position, and adjust the Venetian blind to their ‘user-set’ preferred configuration (Kuhn *et al.*, 2013; Wienold & Christoffersen, 2006). At this point, the experimental procedure was repeated with three more visual tasks presented in a randomised order. Glare assessments were provided by subjects after completion of each task under both shading settings. Each experimental session lasted around 25 minutes. Data were collected from a total of 40 volunteer test subjects (all postgraduate students).

4.1.2.1.4. Glare Detection Methods and Glare Indices

The experimental procedure adopted for this study required drawing a relationship between subjective glare assessments and objectively measured quantities (Velds, 1999; 2002). This involved requesting participants to make judgements of their visual conditions via subjective assessment methods and combining these with instantaneous photometric measurements of the observers' luminous environment. The glare search algorithm adopted by the *Evalglare* tool (utilised in this study to evaluate the HDRI images from the Photosphere software) uses a task definition criterion whereby a fixation area covering most of the Visual Display Unit is outlined within the image (this is represented by the blue circle in Figure 4-11).

In *Evalglare*, each pixel with a luminance value that is 5-times (sensitivity parameter) higher than the average luminance of the fixation area is treated as a glare source. Other studies in the literature have recommended that the threshold for the sensitivity parameter needs to be set between 2 to 7 times higher than the average luminance of the fixation area (Sarey Khanie *et al.*, 2015). In the experimental procedure adopted for the development of the Daylight Glare Probability index (DGP; Wienold, 2009), the task-area threshold in the test room located at the Fraunhofer Institute for Solar Energy Systems (ISE, Germany) was set at 4-times higher than the average luminance of the fixation area. Conversely, a task-area threshold of 7-times higher than the average luminance of the VDU was set for the room located at the Danish Building Research Institute (SBi, Denmark). This was due to the fact that the computer screen VDU in the Danish test room was switched off and, therefore, the average luminance of the task area was lower than that located in the German test room.

At a constant task area threshold, a decrease in the average luminance of the task area increases the sensitivity of the search algorithm used within *Evalglare*, and potentially risks over-detection of glare pixels (i.e., pixels that are not 'particularly' bright may be detected as potential glare sources). This detection method is considered to be very robust against different lighting scenarios, since the sensitivity parameter depends directly on the visual environment that is being investigated (Wienold, 2009).

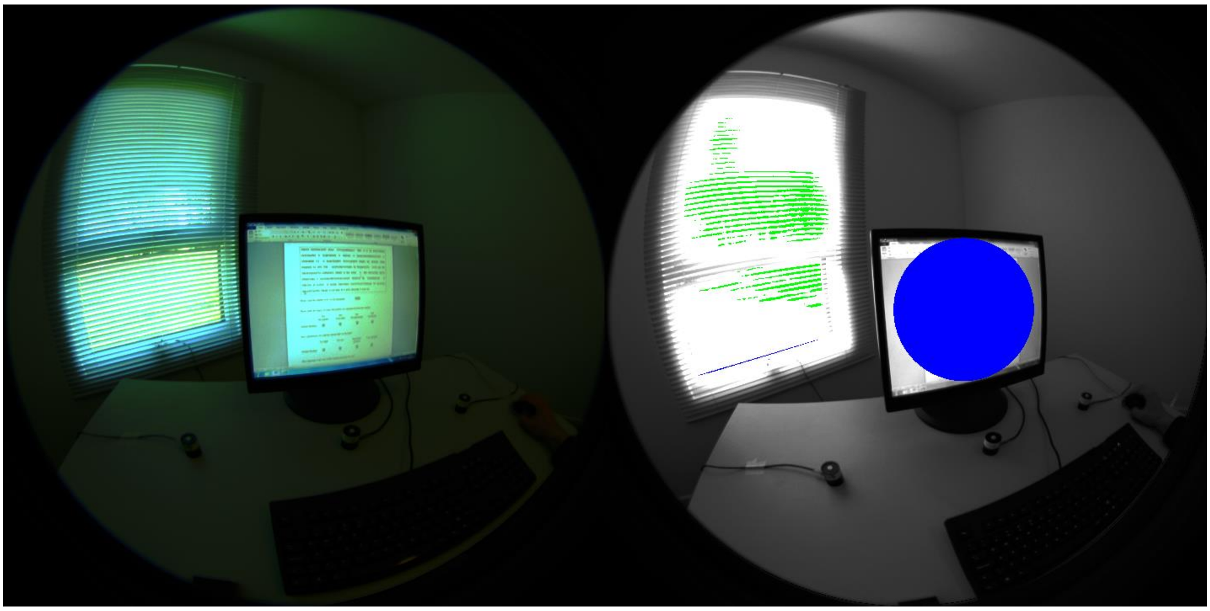


Figure 4-11 HDRI image with Radiance formatting (left); *Evalglare* image with task definition (blue circle) and detected glare sources (green areas) (right)

As described previously in Section 4.1.2.1.2, a comparison of independent measurements was made between the vertical illuminance recorded by the Chromameter and the illuminance values calculated by *Evalglare* from the CCD luminance images collected throughout the experiment. Figure 4-12 shows the comparison between measured and calculated vertical illuminance levels near the head position of the test subjects. Two different shading settings are presented: one (default shading) where blinds were arranged with a 'cut-off' slat tilt that ensured predominantly diffused daylight conditions; and, one (user-set shading) where blinds were adjusted to the test subject's own preferences. The graphs indicate that the correlation between values is high, while any minor differences could be accounted for by the slightly different position of the camera lens and the illuminance meter (Wienold & Christoffersen, 2006). Larger differences could be due to direct sunlight transmission through gaps in the blind slats that, in some cases, might have hit the camera lens but not the illuminance sensor. According to Wienold (2009), these differences are not to be regarded as problematic. Therefore, the luminance images collected during the experimental procedure with test subjects can be considered reliable to inform further analysis of data (Wienold, 2009).

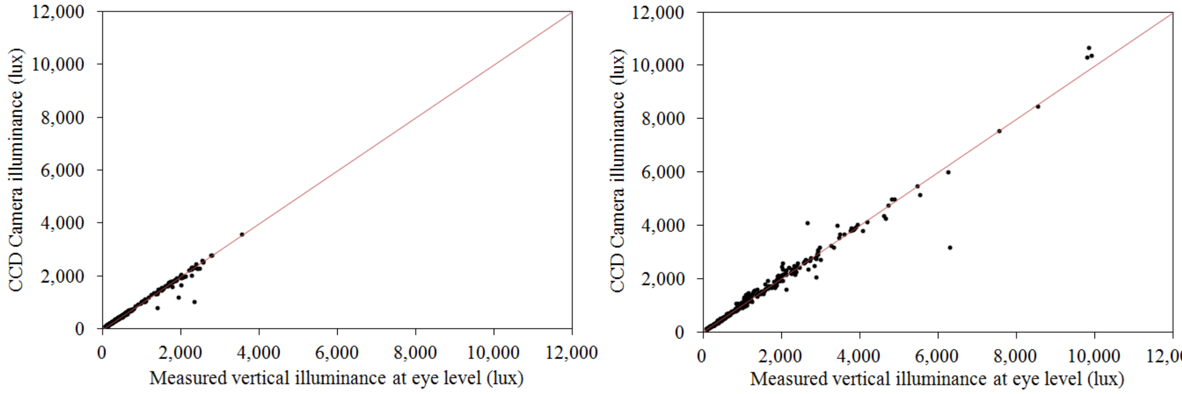


Figure 4-12 Comparison between calculated (from the CCD images, y-axis) and measured (from the Chromameter, x-axis) vertical illuminance: Default shading setting (left); User-set shading setting (right)

To provide a more rigorous analysis (and also support graphical inferences), both the Mean Absolute Deviation (MAD) and the Root-Mean-Square-Error (RMSE) were calculated for the illuminance values obtained from the CCD camera luminance images and those measured by the Chromameter, according to the following formulas (Equations 4-4 and 4-5) (Chai & Draxler, 2014; Willmott & Matsuura, 2005):

$$MAD = \frac{1}{N} \sum_{i=1}^{n} |E_{CCD} - E_{CM}| \qquad \text{(Equation 4-4)}$$

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^{N} E_{CCD} - {E_{CM}}^2} \qquad \text{(Equation 4-5)}$$

Where E_{CCD} is the illuminance obtained from the CCD images and E_{CM} is the illuminance recorded by the Chromameter.

The MAD and RMSE are metrics used to estimate the average error expressed in the units of the variable of interest (i.e., lux) (Chai & Draxler, 2014; Willmott & Matsuura, 2005). In the case of this investigation, the average error refers to the difference in illuminance values obtained from the Chromameter and from the CCD camera. The MAD measures how two data sets are likely to differ from their mean by taking the average absolute value (E_{CCD} – E_{CM}). This aims to avoid that differences with opposing signs might cancel each other out (Willmott & Matsuura, 2005). The RMSE is a measure of the deviation, on average, of a data point from the null hypothesis line. According to the literature, both the MAD and RMSE are widely used metrics to assess model performance (Chai & Draxler, 2014). Table 4-2 displays, for both shading settings, the mean and standard deviation (SD) for the illuminance values obtained from the images captured by the CCD camera ($Mean_{CCD}$) and those recorded by the Chromameter ($Mean_{CM}$), the MAD, and the RMSE.

Table 4-2 Descriptive analysis of the vertical illuminance values calculated from the CCD camera images and measured by the Chromameter

Shading Setting	$Mean_{CCD}$ (SD)	$Mean_{CM}$ (SD)	MAD	RMSE
Default Shading	634.41 (667.41)	654.54 (691.13)	26.70	95.89
User-set Shading	1,369.40 (1,485.11)	1,369.01 (1,479.54)	78.94	220.46

The results indicate that, for both the MAD and the RMSE, the average errors between illuminance values are lower under the default shading setting. This suggests that, when blinds were set at their 'cut-off' position, the differences between the illuminance values obtained from the CCD camera and those measured by the Chromameter are smaller, thereby confirming graphical observations.

To select the appropriate metric to be adopted for the analysis sought in this investigation, a range of photometric measures (with logarithmic transformation) and glare indices were collected and/or calculated. Consideration of logarithmic transformation was due to the fact that, according to the literature, the magnitude of perception of a glare stimulus is logarithmic (Wienold & Christoffersen, 2006). The photometric measures included: illuminance at the eye obtained from *Evalglare* (E_{eye}); illuminance at the window sill (E_{sill}); luminance of the source (L_{source}); and, average luminance (L_{avg}) (this being the average value of luminance corresponding to every pixel within a given HDRI scene evaluated by *Evalglare*). The glare indices featured: Daylight Glare Probability (DGP); Daylight Glare Index (DGI); Unified Glare Rating (UGR); Visual Comfort Probability (VCP); and, CIE Glare Index (CGI).

Rather than independently measuring the statistical and practical significance of the experimental effect of interest (temporal variation of glare response) using all photometric values and glare indices considered – which, potentially could be strongly correlated between one another – all outcomes corresponding to the time of day in which they were obtained (morning, midday, and afternoon) were plotted onto a Pearson's correlation matrix. The matrix plots all measured outcomes against each other, and uses the Pearson's correlation coefficient r as a measure of the strength of the relationship that exists between variables (Field, 2013). The test can demonstrate whether a single outcome variable (i.e., a metric) may be sufficient to evaluate the experimental effect of interest. In fact, the presence of large correlations between photometric values and glare indices would show evidence of very similar trends and differences when evaluating the effect of time of day on glare sensation.

To test the statistical significance of the correlations, a one-tailed hypothesis was adopted, postulating a direct relationship (i.e., as one value increases, so does the other) between variables. However, it was considered that the Visual Comfort Probability (VCP) index provides a rating of comfort whereby higher values correspond to the probability that more people would be comfortable under the given luminous conditions. This outcome is opposite to all other photometric measures and glare indices here considered, for which higher values correspond to a glare source that is perceived by observers as being less comfortable. As a consequence, the Pearson's correlation coefficient was calculated for all measures and indexes against the VCP adopting the hypothesis of an inverse relationship (i.e., as one value (DGP, DGI, etc.) increases, the other (VCP) decreases). The interpretation of the outcome for each correlation was derived from the benchmarks given by Ferguson (2009) for measuring small, moderate, and large effect sizes by the Pearson's r (respectively, $r \geq 0.20, 0.50, 0.80$).

Table 4-3 to Table 4-5 display, for the DGP, DGI, UGR, VCP, CGI, E_{eye}, L_{avg}, L_{source}, and E_{sill}, the Pearson's correlation coefficients r (effect size), measuring the strength of the association that exists between variables along with their statistical significance.

Table 4-3 Pearson product-moment correlation matrix for the Morning test session

	DGP	DGI	UGR	VCP	CGI	E_{eye}	L_{avg}	L_{source}	E_{sill}
DGP	1.00	0.91***	0.93***	-0.86***	0.93***	0.90***	0.87***	0.75***	0.40***
DGI		1.00	0.99***	-0.95***	0.98***	0.81***	0.78***	0.68***	0.42***
UGR			1.00	-0.95***	0.99***	0.81***	0.79***	0.73***	0.37***
VCP				1.00	-0.94***	-0.78***	-0.75***	-0.62***	-0.45***
CGI					1.00	0.80***	0.78***	0.70***	0.40***
E_{eye}						1.00	0.98***	0.78***	0.38***
L_{avg}							1.00	0.78***	0.35***
L_{source}								1.00	-0.02 n.s.
E_{sill}									1.00

*** highly significant; ** significant; * weakly 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

Table 4-4 Pearson product-moment correlation matrix for the Midday test session

	DGP	DGI	UGR	VCP	CGI	E_{eye}	L_{avg}	L_{source}	E_{sill}
DGP	1.00	0.89***	0.90***	-0.76***	0.90***	0.94***	0.94***	0.78***	0.57***
DGI		1.00	0.90***	-0.92***	0.98***	0.75***	0.75***	0.68***	0.62***
UGR			1.00	-0.93***	0.99***	0.75***	0.76***	0.71***	0.60***
VCP				1.00	-0.92***	-0.60***	-0.60***	-0.53***	-0.65***
CGI					1.00	0.75***	0.76***	0.70***	0.62***
E_{eye}						1.00	1.00***	0.79***	0.50***
L_{avg}							1.00	0.80***	0.49***
L_{source}								1.00	0.17**
E_{sill}									1.00

*** highly significant; ** significant; * weakly 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

Table 4-5 Pearson product-moment correlation matrix for the Afternoon test session

	DGP	DGI	UGR	VCP	CGI	E_{eye}	L_{avg}	L_{source}	E_{sill}
DGP	1.00	0.89***	0.91***	-0.87***	0.89***	0.79***	0.76***	0.78***	0.52***
DGI		1.00	0.99***	-0.94***	0.97***	0.71***	0.68***	0.71***	0.33***
UGR			1.00	-0.94***	0.97***	0.73***	0.71***	0.76***	0.30***
VCP				1.00	-0.89***	-0.75***	-0.73***	-0.73***	-0.35***
CGI					1.00	0.72***	0.70***	0.74***	0.32***
E_{eye}						1.00	0.98***	0.92***	0.49***
L_{avg}							1.00	0.91***	0.47***
L_{source}								1.00	0.33***
E_{sill}									1.00

*** highly significant; ** significant; * weakly 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

On the diagonal, the matrix evidently indicates that when comparisons are made between the same parameters (i.e., DGP vs. DGP, DGI vs. DGI, etc.), a perfect correlation exists. The positive sign of the coefficients provides evidence of the direct correlation that exists between variables, with the exception of the VCP due to the inverse nature of the relationships previously highlighted. The Pearson's matrix shows that the correlations between variables have a generally substantive effect size, ranging from large ($r \geq 0.80$) in 41 cases out of 108, to moderate ($0.50 \leq r < 0.80$) in 50 cases, and small ($0.20 \leq r < 0.50$) in 15 cases. The effect sizes are negligible ($r < 0.20$) only in 2 cases. The correlations are highly significant in 106 cases out of 108, significant in 1 case, and not significant in 1 case. Therefore, it can be inferred from this analysis that, across the time of the day, there is a strong correlation between the photometric values measured, those calculated by *Evalglare*, and the glare indexes considered. The only exception is represented by the E_{sill}, which presents a weak correlation to all other variables. This is likely due to the fact that, when taking this measurement, the illuminance sensor was placed on the window sill at a position between the glazing and the Venetian blind. Therefore, the sensor was effectively unaffected by the changes caused by the adjustment of the shading setting that characterised the observer's luminous field. As a result, the sill illuminance can be considered as an unsuitable variable to evaluate the effect of time of day on glare sensation within this investigation.

To identify the most appropriate metric to test the research hypothesis, a review of the literature was performed. From logistical regression analysis, Wienold (2009) found that a range of photometric values (illuminance at eye and average luminance – i.e., the luminance of all the pixels weighed by their solid angle ω in a given image evaluated by *Evalglare*) and glare indices (DGP, DGI, UGR, CGI, and VCP) demonstrated statistical significance when predicting the possibility that an observer would be disturbed by the glare coming from a window. Out of these parameters, the DGP showed the strongest correlation with the probability of occurrence of glare. In this context, it must be considered that most glare indices – with the exception of DGP and DGI (Wymelenberg & Inanici, 2014) – have been developed (and are still used) for predicting discomfort glare from artificial lighting sources. Conversely, these indices have not been designed to deal with non-uniform sources such as, for example, venetian blinds, luminance variations generated by different elements within a view (i.e., ground, buildings, sky with variation in cloud cover) (Tuaycharoen & Tregenza, 2007; Wienold & Christoffersen, 2006), or small glare sources subtending an angle at the eye of the observer below 0.01 sr (Osterhaus, 2005; Velds, 1999). On the other hand, although the DGI has been used to predict discomfort glare from windows, it must be considered that it was originally developed in the absence of sunlight using large-area electric light (Hopkinson, 1972; Chauvel *et al.*, 1982). Based on these considerations, the Daylight Glare Probability (DGP) was selected as the main evaluation parameter for this study (Equation 4-6).

$$DGP = 5.87 \cdot E_v + 9.18 \cdot 10^{-2} \cdot log \sum_i 1 + \left(\frac{L_{s,i}^2 \cdot \omega_{s,i}}{E_v^{1.87} \cdot P_i^2}\right) + 0.16 \qquad \text{(Equation 4-6)}$$

In the equation, E_v is the vertical illuminance at the eye (lux), L_s is the luminance of the glare source detected by *Evalglare* (cd/m^2), ω_s is the subtended size of the source (sr), and P_i is the position index. The DGP provides an indication of the percentage of observers who would be 'disturbed' by the daylight glare that is present within the field of view (Wienold, 2009). Unlike other indices, the DGP is mainly dependent on the vertical illuminance at the eye. Conversely, it is less affected by the sensitivity parameter (i.e., the task-area threshold), since the remaining factors within the DGP formula – L_s, ω_s, and P_i, which are heavily influenced by the sensitivity parameter – have smaller weighted terms (Sarey Khanie *et al.*, 2015).

4.1.2.1.5. Size and Position of the Glare Source

Preliminary analysis of the data – collected from the 40 test subjects that participated to the experiment – revealed that both the size and the position of the glare sources detected by *Evalglare* varied with time of day under either the default (Figure 4-13) and user-set shading settings (Figure 4-14). In particular, when test subjects were permitted to adjust the Venetian blinds to their own preference ('user-set' shading), *Evalglare* detected glare sources that were inside the room, but not within the window area (Figure 4-14). This effectively invalidated the use of the *Evalglare* masking option for the user-set shading – which extracts glare source information in consideration of a specified target area (i.e., the window) – since luminous information associated with detected glare sources (pixels) located outside of the masked area would be lost if only the window area was to be analysed within each scene image.

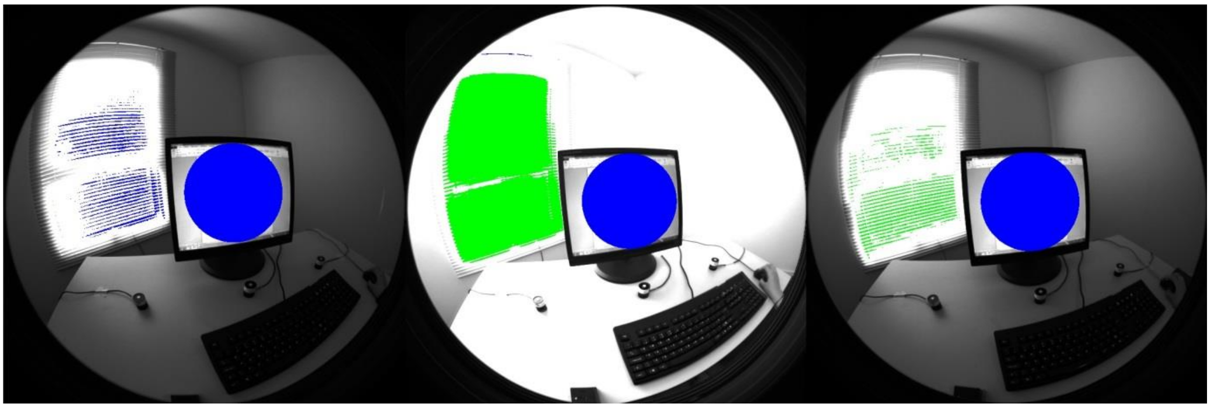


Figure 4-13 Examples of task zone (blue circle) and glare sources detected by *Evalglare* at the Morning (left), Midday (middle) and Afternoon (right) test sessions under the default shading (colours are arbitrarily set by the tool, without being linked to glare magnitude)

Figure 4-14 Examples of task zone (blue circle) and glare sources detected by *Evalglare* at the Morning (left), Midday (middle) and Afternoon (right) test sessions under the user-set shading (colours are arbitrarily set by the tool, without being linked to glare magnitude)

The glare detection algorithm implemented within *Evalglare* uses the average task luminance as the threshold for detecting the glare source(s) (Wienold, 2009). In cases where no clear visual task is present, *Evalglare* reverts to its default detection method by either: calculating the average luminance of the entire scene image and treating as a glare source every pixel with a luminance value that is x-times (sensitivity parameter) higher than the average scene luminance; or, taking a fixed value of luminance and treating every pixel that is higher than this as a glare source (Wienold, 2009). Again, the sensitivity parameter for the search algorithm is recommended to be within the range of 2 to 7 times the average task or scene luminance, although this is part of ongoing research (Sarey Khanie *et al.*, 2015). As noted, it must be reminded here that, in the evaluations made by *Evalglare*, the glare source may not necessarily correspond to the area of the window. Instead, a glare source detected within *Evalglare* can be defined as the summation of the 'glare pixels' whose luminance exceeds by more than x-times (x being a value between 2 and 7) the task/scene average luminance.

As previously mentioned, when evaluating the HDRI images for each time of day in this study, both the size and the position of the glare sources detected by *Evalglare* varied. This presented a problem to the analysis, since the literature indicates that the magnitude of glare sensation can be influenced by both the size (Tregenza & Wilson, 2011) and the position of the source relative to the line of sight (Iwata & Tokura, 1997). To address this issue, rather than measuring and controlling for the size and position indices of the individual glare sources independently – which are distinct parameters in the DGP formula – the solid angle (ω) subtended by the glare sources, modified by the position index (P), was used since this is a combination of both parameters (Equations 4-7 and 4-8). Therefore, by only controlling for the effect of one variable (Equations 4-10 and 4-11→4-12) – i.e., the solid angle subtended by the glare sources modified by the position index (Ω) – high statistical power could be retained, while controlling for the effects of both glare source size and position.

$$\omega = \frac{A_s \cos\theta \cdot \cos\varphi}{d^2} \qquad \text{(Equation 4-7)}$$

$$\Omega = \sum [\, d\omega_s \cdot P_s] \qquad \text{(Equation 4-8)}$$

Where ω is the solid angle subtended by the glare source (sr), A is the subtended area of the glare source (m^2), d is distance from the camera to the glare source (m), θ is the horizontal angle between the line of sight and the glare source, φ is the vertical angle between the line of sight and the glare source, and P is the weighted position index of the glare source. The subscript s is used to denote the glare sources detected by *Evalglare*.

A review of the literature shows that a similar adjustment was also adopted by Tuaycharoen & Tregenza (2007). In fact, from a series of laboratory studies, these authors found that, when comparing images with the same average luminance but with different luminance ratios (i.e., maximum luminance of the glare source against its instantaneous mean luminance), the perceived level of glare sensation was higher for the image that contained higher contrast (Tuaycharoen, 2006). Therefore, when evaluating the effect of interest of view from different windows on reported glare response, an ANCOVA was utilised to control for the influence of luminance contrast on the calculated DGI values.

The position index calculated by *Evalglare* takes into consideration differences in visual discomfort caused by the position of the glare source relative to the line of sight. Within *Evalglare*, the line of sight refers to the task-area. Any glare source located above the task-area (e.g., upper window, ceiling) is calculated using the conventional formula developed by Luckiesh & Guth (1949) (Equation 4-9). Any glare source located below the task-area (e.g., lower window, desk) is calculated using the work of Iwata & Tokura (1997) and the formulas proposed by Einhorn (1997) (Equation 4-10 and 4-11). The expression of the position index below the task-area is dependent on the distance between the glare source and the fixation point (i.e., the VDU containing the visual task). When this distance is small ($R < 0.6D$), Equation 4-10 is applied. When this distance is large ($R \geq 0.6D$), Equation 4.11 is applied.

$$\ln P = \left[35.2 - 0.31889\tau - 1.22e^{-\frac{2\tau}{9}}\right]10^{-3}\sigma + [21 + 0.26667\tau - 0.002963\tau^2]10^{-5}\sigma^2$$

(Equation 4-9)

Where τ is the angle between the vertical plane containing the glare source and the line of sight (degrees), and σ is the angle between the line of sight and the line from the observer to the glare source (degrees).

$P = 1 + 0.8 \cdot \frac{R}{D}$ (Equation 4-10) $\{R < 0.6D\}$

$P = 1 + 1.2 \cdot \frac{R}{D}$ (Equation 4-11) $\{R >= 0.6D\}$

$R = \sqrt{H^2 + Y^2}$ (Equation 4-12)

Where R is the distance between the glare source and the fixation point (m), D is the distance between the observer's eye and the glare source located on the vertical plane (m), H is the vertical distance between the glare source and the view direction (m), and Y is the horizontal distance between the glare source and the view direction (m).

The second sensitivity parameter within *Evalglare* describes the search radius for the glare source pixels. Once glare pixels are detected (i.e., pixels that – in this study – have luminance that is 5-times greater than the task-area luminance), they are merged into larger areas – these areas then become defined as glare sources by *Evalglare* – if the distance between the pixels is small (Wienold, 2009). This distance is defined as the search radius within *Evalglare* and, according to the literature, this value is recommended to be set at 0.20 sr (Sarey Khanie *et al.*, 2015). Therefore, if the distance between glare pixels is lower than 0.20 sr, these pixels are merged into one glare source. Once a detected glare pixel is added to an existing glare source, the average luminance, source luminance, background luminance, solid angle, and position index of the glare source are modified according to the added pixel (Wienold, 2009). Alternatively, if a detected glare pixel is located at a distance that is greater than 0.20 sr from an existing glare source, then this glare pixel is treated as an independent glare source.

4.1.3. Statistical Analysis

4.1.3.1. Fixed-Effects

To analyse the data collected, a multi-level model (MLM) analysis with 'fixed-effects' (i.e., variations in the dependent variable caused by the experimental manipulation of an independent variable, e.g., time of day) was performed to compare the DGP values for all variables that were experimentally manipulated against each other, while controlling for the effects of glare source size and position. Within the model, the fixed-effects specified were:

- Time of day;
- Shading setting;
- Task type;
- GSV reported by test subjects.

For the purpose of this investigation, the GSV reported by test subjects was treated as a fixed-effect although, based on the literature (Seltman, 2014; Snijders, 2005), this could have also been classified as a 'random-effect'. However, in the latter case, any variation on reported GSV would have not been considered as due to experimental manipulation, but rather to intrinsic variations in the environmental conditions or in variables personal to the test subject – e.g., fatigue – influencing their responses to the measureable outcome. In previous laboratory experiments, the measurable outcome (GSV) was separated in discrete categories, which were then analysed by inferential tests to evaluate the effect of experimental interest (i.e., the variation of glare response as the day progresses) at each point a subject reported their glare sensation (Kuhn *et al.*, 2013; Tuaycharoen & Tregenza, 2005, Wienold, 2009).

4.1.3.2. Multilevel Model

Figure 4-15 presents each level of the MLM and the number of conditions associated with the independent variable. The figure also shows the location of each condition at various levels within the model.

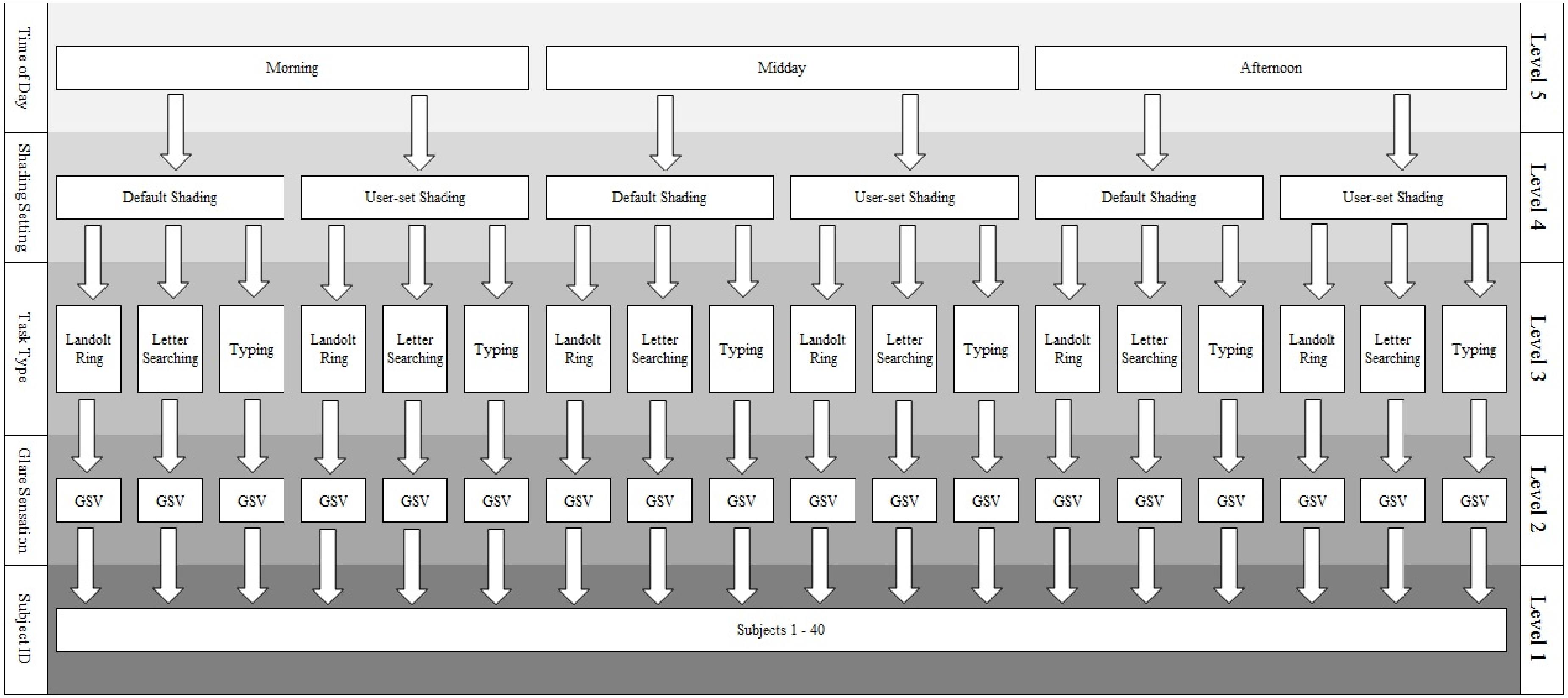


Figure 4-15 Structure of the Multilevel model for the glare assessment data

MLM analysis was selected for this study since it is a method suitable to test observed outcomes using independent variables that are measured at multiple levels (Goldstein, 2010). MLMs are similar to many uni-level models (e.g., MANOVA, ANOVA, ANCOVA, etc.), whereby the dependent variable is used to determine the variation caused by a single (uni-level) effect by making comparisons between two or more independent variables. The primary difference between MLM and uni-level models is that, in MLM analysis, the independent variables are nested in a model with multiple effects (or levels) (Hayes, 2006).

Figure 4-15 displays the multiple levels on which the fixed-effects are nested upon (left), and the level in which each fixed-effect has been specified within the MLM (right). The arrows indicate the direction in which the glare assessments (i.e., the GSVs reported by test subjects) are distributed among independent variables. Table 4-6 shows the independent variables under examination (i.e., the fixed-effects within the MLM), the corresponding level in the model within which the independent variable is located, the number of conditions (independent groups) within each independent variable, and the number of glare assessments collected for each conditional group related to the independent variable.

As an example, for the independent variable ‘Time of Day’ (level 5), the number of glare assessments collected for each conditional group (i.e., test sessions) is the highest, since the total number of glare assessments (N= 720, that is 40 subjects providing glare assessments at 3 times of the day after performing 3 visual tasks under 2 shading conditions) is divided by three independent conditions (Morning, Midday, and Afternoon). For level 2 ‘Glare Sensation Vote (GSV)’, the number of glare assessments for each outcome variable (‘Just (Im)Perceptible’, ‘Just Noticeable’, ‘Just Uncomfortable’, and ‘Just Intolerable’) is the lowest, since the number of glare assessments collected is divided by the number of conditional groups at that level, in addition to all the conditional groups that exist within the independent variables located at higher levels of the model. Hence, the measureable outcome for each condition of GSV cannot be specified accurately in Table 4-6 (<40).

Table 4-6 Distribution of conditions and glare assessments across the independent variables

Independent Variables (Fixed-Effects)	Level	No. Conditions for the Independent Variables	No. Glare Assessments
Time of Day	5	3	240
Shading Setting	4	2	120
Task Type	3	3	40
GSV	2	3	<40
Subject ID	1	-	-

As mentioned, in this study the GSV reported by test subjects was treated as a fixed-effect within the MLM, and was measured using multiple independent uni-level statistical tests. Despite having features that could characterise GSV as a random-effect, the literature suggests that, when estimating average values of a variable under examination within the sample population, the investigator will often specify a variable as a fixed-effect (rather than a random-effect) if it is of primary interest (Kreft & de Leeuw, 1998; Snijders, 2005). Level 1 ‘Subject ID’ is included at the bottom level of the model on the basis of the repeated-measures procedure used within this experiment (Field, 2013; Seltman, 2014). This postulates that the glare assessments recorded in each of the upper level measurements (levels 2, 3, 4, and 5) should be both correlated within each uni-level (i.e., time of day) and across each of the multiple levels (i.e., time of day and shading setting). In other words, since test subjects were requested to provide votes of glare sensation on multiple occasions, at each level within the model, and for each conditional group, there will be some form of relationship that exists between reported levels of GSV. However, this relationship causes a lack of independence between observations, which are clustered on multiple levels. According to the literature, this might determine serious problems, if not properly addressed (Julian, 2009; Romano, 2007).

4.1.3.3. Independence of Observations

In performing the analysis, it must be considered that the glare assessments collected were not independent from each other (Huang & Cornell, 2015). A review of the literature suggests that, when interpreting the outcome of a MLM with fixed-effects, the approach adopted should be the same as that used in many parametric uni-level tests examining mean differences between independent variables (Field, 2013; Seltman, 2014). In parametric inferential testing, in fact, the main assumptions associated with the sample population are:

- Observations in each group are normally distributed around the mean parameter;
- Variances associated with each independent group are homogeneous (Heiman, 2013);
- Observations between and within each group are independent from each other (Romano, 2007).

A review of the literature shows that, when the first two assumptions (i.e., normality and homogeneity of variances) have not been met, their influence on the inflation of the alpha-level (i.e., the risk of incurring in Type I errors) is not as severe as when the assumption of independence has been violated (Scariano & Davenport, 1987; Hox, 1998).

A study by Landman & Dawes (1982) identified different scenarios that might lead to a violation of the assumption of independence, most notably when:

- Multiple measures of the same outcome are obtained from the same subject;
- Assessments of the same outcome are obtained by the same subject at more than one point in time.

In this investigation, to examine whether the assumption of independence was satisfied, the intra-class correlation (ICC) was estimated according to the following formula (Equation 4-13) (Hayes, 2006; Peugh, 2010):

$$ICC = \frac{\tau_{00}}{(\tau_{00} + \sigma^2)} \qquad \text{(Equation 4-13)}$$

Where τ_{00} is the estimated variance (i.e., the variation in calculated DGP values at levels 2, 3, 4 and 5 (within group variance)), and σ^2 is the residual variance (i.e., the variation in calculated DGP values at level 1 (between group variance)) (Hayes, 2006; Peugh, 2010).

According to the literature, if ICC is sufficiently close to zero, the assumption of independence has not been violated (Julian, 2001). However, this does not necessarily mean that the data should be analysed using conventional uni-level inferential tests (Hayes, 2006; Peugh, 2010). Utilising Equation 4-13, the ICC was calculated for the data collected within this study: ICC= 0.002577 / (0.002577+0.000340) = 0.883442 (large). The interpretation of the outcome was derived from the benchmark values provided by Julian (2001) for low, moderate and high ICCs (respectively, ICC$\geq$ 0.05, 0.15 and 0.45). Therefore, the result (ICC= 0.88) indicates that the assumption of independence for this investigation was not satisfied.

Several studies from the literature suggest that, when the assumption of independence has been violated, the MLM represents a more appropriate method of analysis (Beretevas & Pastor, 2003; Field, 2013; Hox 1998; Romano, 2007; Peugh, 2010). In fact, MLM are tests that account for dependencies among observations at multiple levels (Huang & Cornell, 2015) by estimating a single variance structure, which represents how spread-out the random intercepts are around the common intercept of each group (Seltman, 2014). In other words, the covariance structure estimates how the variance parameters for each participant are related across the fixed-effects within the MLM (Muthen, 1994) and compares this with a specified covariance structure using a goodness-of-fit index (Field, 2013). This ‘relaxes’ the assumption of independence within the inferential tests, hence reinforcing the reasons behind the selection of MLM analysis in the context of this investigation (Muthen, 1994; 2014).

4.1.3.4. Covariance Structure and Goodness-of-Fit

MLMs offer a flexible approach to estimate variance parameters, since direct assumptions regarding covariance structure (i.e., how the variances associated with each independent group are related to each other) can be specified (Goldstein, 1995). These assumptions are dependent upon experimental design; for example, in time-based studies, the variance associated with independent groups may systematically change due to experience and practice effects (Field & Hole, 2013). In this investigation, an autoregressive (AR(1)) covariance structure was assumed within the variances associated with each independent group. For instance, measurements taken at closer time intervals are postulated to be more highly correlated than when intervals are distant from each other (Field, 2013; Seltman, 2014). Thus, it is assumed that the variances systematically change over time (Field, 2013).

To assess the suitability of the covariance structure applied to the MLM, a goodness-of-fit index was evaluated. This index assesses the overall fit – using a χ^2 likelihood ratio test – between the estimated variance parameters and the selected covariance structure (Field, 2013). The literature recommends use of the Bayesian Information Criterion (BIC) (Seltman, 2014), which adjusts the statistical outcome based on the number of fixed-effects and the sample size used within the MLM (Field, 2013; Seltman, 2014). In interpreting the outcome of the BIC, the smaller the value (χ^2), the better the model fit. However, no absolute interpretation can be made of the statistical outcome – i.e., no meaningful information can be inferred from a single statistical value (Field, 2013). Instead, the BIC can be compared with equivalent values from other models that contain either additional random-effects (retaining the original MLM with fixed-effects) or use a different covariance structure (Seltman, 2014). Therefore, by increasing the complexity of the model (e.g., by including additional random-effects), or fitting a more suitable covariance structure, the BIC provides statistical information that can be used to estimate the presence of unknown parameters within the MLM.

4.1.3.5. Null Hypothesis Significance Testing

For each fixed-effect, the main interaction was studied. This implied analysing the difference between two or more means at a single level – for example, level 5 ‘Time of day’: Morning vs. Midday – without grouping the dependent variable (adjusted mean value of the DGP, Daylight Glare Probability) by shading setting, task type, or GSV (i.e., the other fixed effects). The null hypothesis (Equation 4-14) and alternative hypothesis (Equation 4-15) adopted for testing the effect of time of day on glare sensation within the MLM were:

$H_0: \mu_{morning} = \mu_{midday} = \mu_{afternoon}$ (Equation 4-14)

$H_a: \mu_{morning} \neq \mu_{midday} \neq \mu_{afternoon}$ (Equation 4-15)

Where μ is the adjusted mean value of the DGP calculated at each time of day.

In addition, interactive effects between the fixed-effects and the covariate (i.e., the solid angle subtended by the glare source modified by the position index (Ω)) were specified for inclusion in the inferential analysis. This required testing the differences between two or more means across independent variables, as an example level 5 ‘Time of day’ and level 4 ‘Shading’. For instance, if comparing the differences in mean values of DGP for the default and user-set shading settings between the morning, midday, and afternoon sessions, the null hypothesis (Equation 4-16) and alternative hypothesis (Equation 4-17) for consideration of the interactive effects, in this case, between time of day and shading were as follows:

$H_0: \Delta\mu_{morning} = \Delta\mu_{midday} = \Delta\mu_{afternoon}$ (Equation 4-16)

$H_a: \Delta\mu_{morning} \neq \Delta\mu_{midday} \neq \Delta\mu_{afternoon}$ (Equation 4-17)

Where Δμ is the difference between the mean DGP values calculated under the default and the user-set shading settings at each time of day.

In the MLM, all possible outcomes on the dependent variable (DGP) were specified using the variables that were known to vary within the experiment (i.e., GSV, time of day, shading setting, and task type). This can be assumed to 'consume' as much as possible of the scatter that is commonly associated with subjective evaluations of glare sensation (Tregenza & Wilson, 2011). In respect to the primary aim of this investigation (see Section 1.2), the MLM was used to determine whether there was sufficient evidence in the data collected from the test room experiment to reject the null hypothesis (Equation 4-16) and accept the alternative hypothesis (Equation 4-17), therefore implying that the mean DGP values at each time of the day were statistically and practically different from each other. This would allow postulating an effect of time of day on glare sensation in the presence of daylight from a window.

4.1.3.6. Estimates of Covariance Parameters

As mentioned, all outcomes on the DGP were specified in the MLM using the fixed-effects available (i.e., main interaction and interactive effects). The MLM analysis provides estimates of mean parameters and their associated statistical difference. When a statistically significant difference is detected, there is a reduction in the total amount of variance present within the model (Seltman, 2014). Once all mean parameters have been estimated, the MLM then calculates whether the remaining unexplained variance is significantly different from a model that has a variance equal to zero; this test is called the Wald Z statistic (Peugh, 2010). The statistical power of the Wald Z statistic is dependent on the size and evenness of the sample and, more importantly, on the number of interactive effects specified within the model (Hox, 1998). This shows that, as more interactive effects are specified, the smaller the estimated variance parameter becomes. The null hypothesis is that the unexplained variance within the model is equal to zero (Seltman, 2014). As an alternative hypothesis, the Wald Z test seeks to demonstrate that there is sufficient evidence to suggest that the unexplained variance within the model is not equal to zero. Rejection of the null hypothesis would imply that there are other important unmeasured variables that are unaccounted for within the MLM.

4.1.3.7. Parameter Estimation

In MLM analysis, there are two methods that can be used to estimate parameters: Maximum Likelihood (ML) and Restricted Maximum Likelihood (REML). The main difference between these centres upon how the two methods estimate the variance parameters for the fixed-effects considered (Peugh, 2010). The ML provides more robust estimates of fixed regression parameters (Field, 2013), although it is dependent on large sample sizes (Hox, 2004; Peugh, 2010). This limitation is not problematic for main-interactions or interactive effects at group level (i.e., when post-hoc analysis is performed and the sample size is distributed to isolate the main-effects). However, when the sample size becomes low due to the number of levels within the MLM, parameter estimates may not be robust (Mass & Hox, 2005). This is one of the main disadvantages associated with the use of MLM analysis (Beretvas & Pastor, 2003), since the sample distribution will always be the lowest at group level. In contrast, the REML provides more robust estimates of variance parameters and does not rely on large sample sizes (Kenwood & Rogers, 1997; Peugh, 2010). Nevertheless, the major caveat of the MLM is that different models can only become comparable if the ML estimation is used (Field, 2013). For this reason, the ML was adopted for this study.

4.1.4. Results and Discussion

40 test subjects volunteered to take part to the experiment, which was carried out between the months of February and April. Participants were all postgraduate students: 12 male and 28 female, the mean age was 25.00 (standard deviation, SD= 2.59), 3 were left-handed, 37 right-

handed, and 15 wore corrective lenses. The criterion adopted for the selection of participants was purposive sampling, and subjects were recruited via an online advertisement addressed to all postgraduate students at the Department of Architecture and Built Environment, University of Nottingham. There were no criteria used for the exclusion of volunteers. Figure 4-16 plots, on the y-axis, the adjusted DGP marginal means (with 95% confidence intervals) controlled for the effects of both glare source size and position, as calculated by *Evalglare*. The x-axis presents the time of day when glare assessments were provided, organised in terms of the reported levels of GSV. Since relatively few votes of 'Just Uncomfortable' and 'Just Intolerable' were given by test subjects, these GSV criteria were merged so as to enable performing meaningful statistical analysis. A similar modification was also undertaken in the study by Kuhn *et al.* (2013). For each criterion of GSV, the plots are distributed according to the setting of the shading device, corresponding respectively to Default and User-set shading.

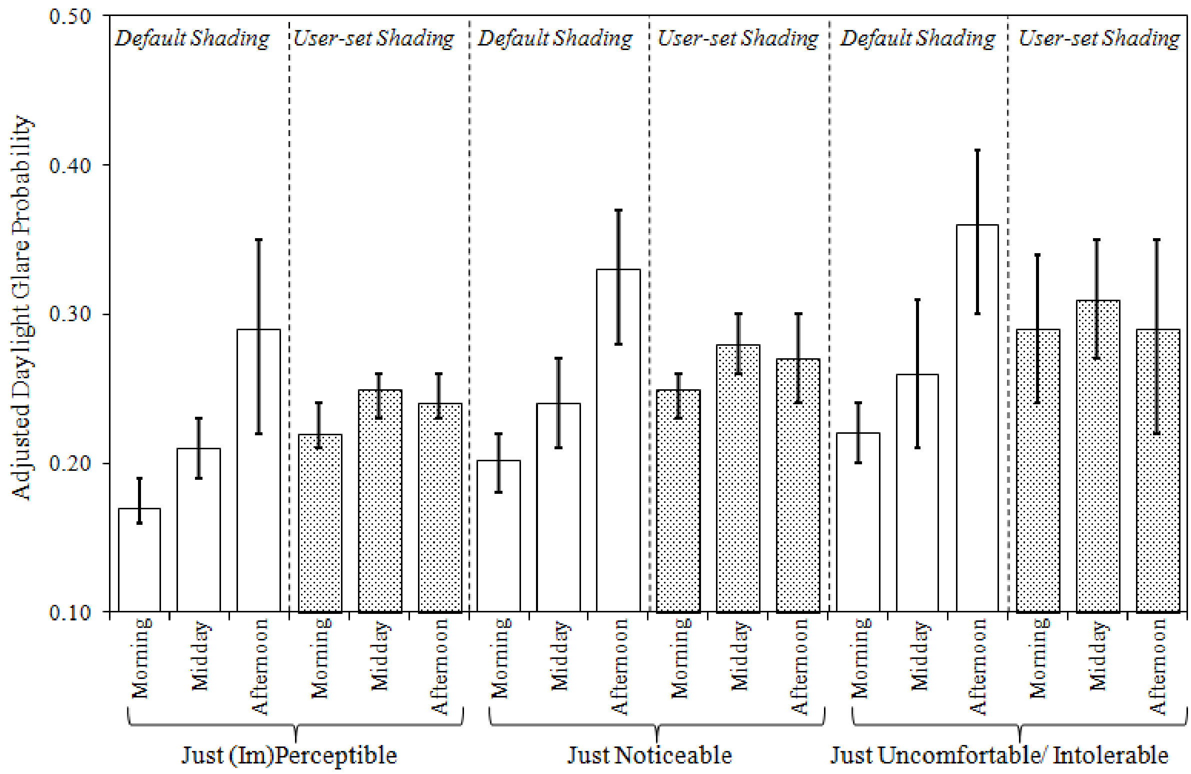


Figure 4-16 Mean plots with 95% confidence intervals of adjusted DGP values (y-axis) for each GSV criterion, test session, and shading setting (x-axis); fixed-effects MLM

Graphical inspection of the boxplots suggests that, under the default shading setting, there is a consistent trend for central tendencies to correspond to higher levels of DGP for each criterion of glare sensation as the day progresses. Also, as expected, the adjusted mean DGP values show a tendency to amplify with increasing levels of glare sensation at each test session (i.e., this pattern appears when comparing the plots for individual sessions – e.g., Morning – across different benchmarks of glare sensation votes).

It is important to remember that a higher level of mean DGP signals a greater probability that an observer may be 'disturbed' by the combination of photometric and physical parameters associated with the glare source(s). Since the same level of glare sensation provided by subjects (for example, Just (Im)Perceptible) corresponds to increasing levels of mean DGP as the day progresses (from Morning to Midday to Afternoon), the boxplots suggest that – when the Venetian blinds were set at the default 'cut-off' shading setting – subjects became more tolerant to the combination of photometric and physical parameters associated with the glare source(s) along the day. Therefore, from initial graphical analysis, there are reasons to suspect that, under the default shading setting, test subjects showed higher tolerance to the glare source(s) at later sessions for all levels of glare sensation. If substantiated by inferential testing, this would support the directionality of the effect of time of day on glare sensation previously detected under laboratory conditions (see Sections 3.1, 3.2, 3.3, and 3.4).

For the user-set shading setting, the effect of time of day on glare response is not apparent. In fact, in this case, graphical inspection of statistical parameters shows no univocal interpretation of a prevailing tendency for any of the GSVs as the day progresses. Therefore, from the boxplots of Figure 4-16, there is no consistent evidence to suggest that an effect of time of day was present on the subjective assessment of glare sensation provided by test subjects when they were allowed to control the setting of the Venetian blind.

When grouping the data by time of day, shading setting, task type, and GSV, visual inspection of the plots did not reveal any noticeable difference in central tendencies (mean) for the effect of task type. Inferential analysis explored the overall effect of task type on glare sensation with the following null (Equation 4-18) and alternative hypotheses (Equation 4-19):

$$Ho = \mu_{Landolt\ ring} = \mu_{Letter\ searching} = \mu_{Typing} \qquad \text{(Equation 4-18)}$$

$$Ha = \mu_{Landolt\ ring} \neq \mu_{Letter\ searching} \neq \mu_{Typing} \qquad \text{(Equation 4-19)}$$

Where μ is the mean DGP value calculated for each visual task used in the experiment.

No statistically significant differences were detected upon consideration of the type of task, $F(2, 670.76) = 0.19$, $p = 0.83$, therefore task type was excluded from further post-hoc analysis.

For every inferential test performed, at all times of the day and at all reported levels of glare sensation, the DGP was calculated at a constant value of the solid angle subtended by the glare source modified by the position index (Ω) to control for its temporal influence on the dependent variable. This value (Ω) was defined by the statistical package (SPSS) and was an adjustment derived from the MLM through multiple regression, utilising the alpha-level (statistical significance), the coefficient of the outcome (i.e., the mean difference in DGP calculated from the HDRI images evaluated by *Evalglare*) weighted for the effect of the covariate (i.e., the solid angle subtended by the glare sources modified by the position index calculated from the images evaluated by *Evalglare*) regressed onto the fixed-effects (i.e., time of the day), and the unexplained residual variance remaining within the model (Field, 2013).

For this investigation, all HDRI images evaluated by *Evalglare* – calculating DGP values and the covariate (i.e., the solid angle subtended by the glare sources modified by the position index) – were analysed using the same glare source detection parameter. The detection parameter searched for glare pixels with a luminance that was 5-times larger than the average luminance of the defined task-area. By 'masking' the window area for the default shading setting – whereby detected glare sources can only be contained within the window area – and decreasing the glare detection parameter of the masked area, it would become feasible to examine DGP values from a glare source corresponding to the entire window area for all times of the day. However, in so doing, glare pixels detected within the images might correspond to low luminance values (for example, 73.1 cd/m^2 from ground elements, as tested on some scene images) that, if merged with adjacent detected glare pixels (<0.20 sr), could produce glare sources of small luminance level, potentially causing poor association with reported glare sensation. This is a finding validated by the work of Wienold (2009), which showed an improved predictive estimation between subjective glare responses and source luminance corresponding to pixels 5-times higher than the average luminance of the defined task-area, compared to when the source luminance was calculated from the entire window area. These considerations reinforce the selection of covariate analysis for this study.

In fact, based on the literature, when statistically controlling for the effect of an uncontrolled variable, the variance associated with the covariate – that is attributed to the dependent variable – is subtracted, holding it constant by assigning it a defined value (Vogt, 1999). Moreover, the removal of a covariate – known to also influence the dependent variable – not only eliminates the bias associate with average estimates, but it also reduces the unexplained variance within the model (Field, 2013). Consequently, this lessens the total amount of error, and allows isolating the effect of experimental interest with greater accuracy (Snijders, 2005).

For the DGP, upon consideration of the effect of time of day (i.e., Morning vs. Midday vs. Afternoon) while controlling for glare source and size, the MLM detected a highly significant difference, $F(2, 571.46)= 9.93$, $p \leq 0.001$. This result provides statistically significant evidence that the adjusted mean DGP values calculated at the three sessions were different from each other, thereby signalling an effect of time of day on the glare sensation reported by subjects.

Follow-up univariate tests were then performed by grouping the effect of time of the day by the GSV reported by subjects and by shading setting. These tests compared the DGP under all times of day, similar to an ANOVA. Table 4-7 presents the inferential data from the univariate tests, providing the shading setting, the GSV reported by test subjects, the degrees of freedom (df), the test statistic (*F*), and the statistical significance (*p*-value). As the type of the task showed no statistically significant differences as a main effect, nor did it demonstrate any statistically significant interaction with any other fixed-effect, consideration of task type was removed from the analysis. In fact, in order to prevent the occurrence of Type II errors, any non-significant fixed-effect should be excluded from post-hoc analysis so as to reduce the number of levels within the model (Seltman, 2014). This is the main disadvantage of MLM, since at the lowest level (i.e., Level 2 'GSV') the statistical power will be at its minimum (Beretvas & Pastor, 2003) and, as previously mentioned, the parameter estimation used (ML) is dependent on asymptotic calculations (i.e., large sample sizes) (Hox, 2004; Peugh, 2010). Therefore, since inclusion of more independent variables and conditional groups requires more parameters (mean and variance) to be estimated, larger samples are necessary to provide more robust estimates (Field, 2013).

Table 4-7 Univariate tests comparing DGP values under all times of day for each GSV criterion and shading setting (fixed-effects MLM)

Shading	GSV	Numerator df	Denominator df	***F***	***p*-value**
Default Shading	Just (Im)Perceptible	2	685.86	6.87	0.00***
	Just Noticeable	2	712.68	16.17	0.00***
	Just Uncomfortable/Intolerable	2	694.87	8.72	0.00***
User-set Shading	Just (Im)Perceptible	2	622.08	2.97	0.06 n.s.
	Just Noticeable	2	692.99	2.45	0.09 n.s.
	Just Uncomfortable/Intolerable	2	701.14	2.54	0.08 n.s.

* weakly significant; ** significant; *** highly significant; n.s. not significant

Under the default shading setting, Table 4-7 shows 3 highly significant differences; conversely, under the user-set shading, the tests detected 3 non-significant differences. Therefore, the univariate tests provided evidence that the variations of DGP values along the day are statistically significant when the Venetian blinds are set at their default ‘cut-off’ position. However, when the subjects adjusted the blinds based on their preferences, the tests did not offer supportive evidence of an effect of time of day on glare response for any GSV.

To isolate the main effects between variables, contrasts were made using pairwise comparisons (Bird & Hadzi-Pavlovic, 2014), whereby all permutations between times of day were compared against each other. The directionality of the hypothesis was informed by examination of descriptive statistics and visual inspection of central tendencies from graphical displays of data (Hauschke & Steinijans, 1996). Since no consistent directionality between the observed differences could be detected upon simultaneous consideration of both shading settings, two-tailed hypothesis testing was applied (Ruxton & Neuhauser, 2010).

In consideration of the experiment-wise error rate caused by the significance level inflating across multiple tests carried out on the same data (Cabin & Mitchell, 2000) – which was calculated as $1-(0.95)^n= 0.14$ (thus risking a 14% probability of making at least one Type I error), where n= 3, i.e. the number of pairwise comparisons performed – Bonferroni corrections were applied. As null hypothesis significance testing (NHST) depends both on the size of the sample and on the magnitude of the postulated effect under examination (Ellis, 2010), emphasis of the inferential tests was placed on the effect size (i.e., a standardised measure of the observed difference between sample groups) and not only on their statistical significance (which, particularly for small or uneven sample size, could confound effect size and sample size). For this analysis, the effect size was calculated by the Cohen’s d coefficient. This was obtained from the estimated marginal means and pooled standard deviation scores adjusted for the effect of glare source size and position by the MLM analysis, according to the following formula (Equation 4-20) (Fields, 2013; Rosenthal & DiMatteo, 2001):

$$Cohen's\ d = \frac{\Delta M}{\sigma pooled} \qquad \text{(Equation 4-20)}$$

Where ΔM is the difference between the estimated marginal means, and σ_{pooled} is the pooled standard deviation.

The interpretation of the outcome was derived from the benchmarks provided by Ferguson (2009), for small, moderate, and large effect sizes (d≥ 0.41, 1.15, and 2.70, respectively). Values below 0.41 were considered to be not substantive (i.e., not practically relevant).

For each GSV, Table 4-8 to Table 4-10 report the shading setting, the sample size (N) of the pairwise comparisons (x_0 and x_1 corresponding to the test sessions considered), the difference between the estimated marginal means adjusted for the effect of glare size and position (ΔM) and its associated two-tailed statistical significance (NHST, *p*-value with Bonferroni correction), the standard error, the degrees of freedom (df), the upper and lower confidence intervals for the difference between estimated marginal means, and the effect size (d).

Table 4-8 Pairwise comparisons and effect sizes for 'Just (Im)Perceptible'

Shading	Time of Day	N(x_0, x_1)	ΔM^{NHST}	Std. Error	df	CI_L	CI_U	Effect Size (d)
Default Shading	Morning vs. Midday	49, 37	-0.04*	0.01	678.78	-0.06	-0.01	-0.72
	Midday vs. Afternoon	37, 45	-0.08 n.s.	0.03	640.49	-0.16	0.00	-0.50
	Morning vs. Afternoon	49,45	-0.13**	0.03	669.21	-0.19	-0.04	-0.74
User-set Shading	Morning vs. Midday	72, 72	-0.03 n.s.	0.01	675.85	-0.05	0.00	-0.40
	Midday vs. Afternoon	72, 84	0.01 n.s.	0.01	679.91	-0.02	0.03	0.09
	Morning vs. Afternoon	72, 84	-0.02 n.s.	0.01	514.32	-0.05	0.01	-0.18

* weakly significant; ** significant; *** highly significant; n.s. not significant
d < 0.41= negligible; 0.41 ≤ d < 1.15= small; 1.15 ≤ d < 2.70= moderate; d ≥ 2.70= large

Table 4-9 Pairwise comparisons and effect sizes for 'Just Noticeable'

Shading	Time of Day	$N(x_0, x_1)$	ΔM^{NHST}	Std. Error	df	CI_L	CI_U	Effect Size (d)
Default Shading	Morning vs. Midday	47, 61	-0.04**	0.01	692.97	-0.06	-0.02	-0.61
	Midday vs. Afternoon	62, 44	-0.10***	0.02	675.00	-0.15	-0.04	-0.73
	Morning vs. Afternoon	47, 44	-0.13***	0.02	706.00	-0.19	-0.07	-1.08
User-set Shading	Morning vs. Midday	43, 38	-0.03 n.s.	0.01	719.56	-0.06	0.00	-0.46
	Midday vs. Afternoon	38, 26	0.01 n.s.	0.02	693.69	-0.04	0.05	0.14
	Morning vs. Afternoon	43, 26	-0.02 n.s.	0.02	654.47	-0.07	0.02	-0.27

* 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

Table 4-10 Pairwise comparisons and effect sizes for 'Just Uncomfortable'

Shading	Time of Day	$N(x_0, x_1)$	ΔM^{NHST}	Std. Error	df	CI_L	CI_U	Effect Size (d)
Default Shading	Morning vs. Midday	24, 22	-0.04 n.s.	0.02	714.30	-0.06	0.00	-0.80
	Midday vs. Afternoon	22, 31	-0.11***	0.03	661.40	-0.19	-0.04	-0.80
	Morning vs. Afternoon	24, 31	-0.14***	0.03	679.73	-0.21	-0.06	-1.12
User-set Shading	Morning vs. Midday	5, 10	-0.02 n.s.	0.03	683.06	-0.08	0.05	-0.44
	Midday vs. Afternoon	10, 10	0.02 n.s.	0.02	704.70	-0.01	0.04	0.44
	Morning vs. Afternoon	5, 10	0.00 n.s.	0.03	711.68	-0.03	0.03	0.00

* 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

Consistent with previous findings (see Section 3.1.3, Section 3.2.4, and Section 3.4.3), analysis of the descriptive statistics shows that, under the default shading setting, the mean differences (ΔM) are consistently negative. This suggests an increased value of DGP as the day progresses for each threshold of glare sensation. The pairwise comparisons provide evidence that, under the default shading setting, the ΔMs are highly significant in 4 cases, significant in 2 cases, weakly significant in 1 case, and not significant in 2 cases out of 9. The differences detected have a substantive effect size in all 9 comparisons (in this case, as a measure of effect size, the Cohen's d has to be considered in its absolute value). The inferential results, therefore, confirm the postulated hypothesis of a tendency for the DGP to increase at later times of the day under the default shading setting for each GSV criterion.

Conversely, observation of descriptive statistics (ΔM), confidence intervals, and effect sizes under the user-set shading setting shows no prevailing directionality, this providing no reason to hypothesise that the glare sensation votes reported by test subjects consistently correspond to higher or lower levels of DGP as the day progresses. The pairwise comparisons show that none of the ΔMs under the user-set shading setting are statistically significant. Effect sizes are of small practical relevance only in 3 cases (all not statistically significant), the remaining influences being of negligible magnitude. Therefore, under the user-set shading, the results do not offer any evidence to postulate an effect of time of day on reported glare sensation.

Figure 4-17 plots, for each threshold of GSV and for the two shading settings, the effect sizes (d) from the pairwise comparisons presented in Table 4-8 to Table 4-10. For each contrast, the figure also provides indication of mean differences (ΔM) and their statistical significance.

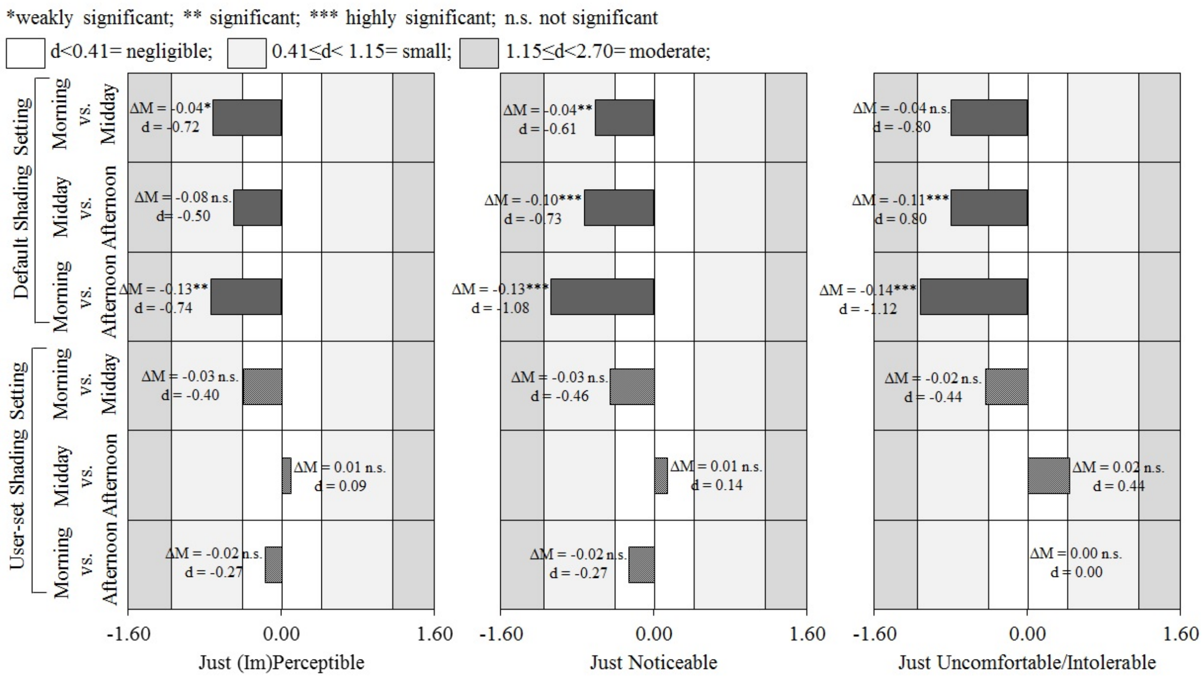


Figure 4-17 Effect sizes (d) detected in the pairwise comparisons (see Tables 4-8 to 4-10) between test sessions (y-axis), for each GSV criterion (x-axis) under the default (top) and user-set (bottom) shading settings; fixed-effects MLM

Under the default shading setting, Figure 4-16 reveals a tendency for the magnitudes of effect size to amplify with the increase in the time interval between test sessions. This is evident when comparing the 'Morning vs. Midday' (3 hours between tests) against the 'Morning vs Afternoon' (an interval of 6 hours) sessions, across all criteria of glare sensation. The increases in effect size signal that subjects reported the same GSV criterion for photometric and physical measures corresponding to higher DGP values at later times of day. This confirms previous results obtained under laboratory conditions (see Section 3.1.3, Section 3.2.4 and Section 3.4.3), supporting the postulated greater tolerance to discomfort glare as the day progresses under the default shading setting. Under the user-set shading conditions, no clear directionality can be identified in the variation of effect sizes along the day (and no statistically significant differences could be detected in the pairwise comparisons). Therefore, when comparing the findings under the two shading settings, it can be inferred that the effect of time of day on glare response between the two conditions is effectively independent.

In essence, the data presented in Table 4-8 to Table 4-10, and graphical inspection of Figure 4-16, provide statistically significant and practically relevant evidence of an effect of time of day on the glare sensation reported by test subjects in the presence of a window, when Venetian blinds avoided the penetration of direct sunlight over the desk area. Under this shading setting, the variation of tolerance to glare sensation appeared to amplify when considering a larger time gap between test sessions. Conversely, no clear indication of the same temporal influence was detected when users were given the chance to control the blinds based on their own preferences. This leads to hypothesise that, once the blinds have been adjusted, the presence of other uncontrolled variable(s) or condition(s) may have potentially masked or confounded the effect of time of day on glare response. In this context, for example, studies in the literature have suggested an influence of the view on the glare sensation reported by test subjects (Tuaycharoen & Tregenza, 2005; 2007) (see Section 4.3).

In interpreting the outcomes of this experiment, some limitations should be acknowledged. Among these, it must be considered that, when evaluating glare sensation at high levels of perceived visual discomfort (i.e., 'Just Uncomfortable' and 'Just Intolerable' GSVs), the experimental setup effectively protected the subjects from direct sunlight exposure under the default shading setting, while it allowed the subjects to adjust the Venetian blind to their own visual preference after the first round of assessments. This procedure may have prevented the subjects from experiencing high levels of glare sensation. A similar observation was also reported by other glare studies (Kuhn *et al.*, 2013; Van Den Wymelenberg & Inanici, 2014). The distribution of the population sample for the glare assessments reporting high levels of visual discomfort also demonstrates this limitation (Table 4-6). As a result, when evaluating data related to the highest GSVs, it was difficult to maintain high statistical power at group level since the sample size was distributed across multiple variables. Clearly, this limitation challenges the formulation of any univocal and consistent inference on the results obtained, although the outcomes of this experiment do support earlier findings from laboratory studies.

The results from this experiment are supportive of the research hypothesis that time of day directly influences the perceived level of glare sensation from daylight. However, within the MLM, the estimates of covariance parameters showed a highly significant difference (Wald Z= 12.28, $p \leq 0.001$), providing evidence to reject the null hypothesis – i.e., that the variances associated with each effect are equal to zero. This suggests that the fixed-effects included in the original MLM model are not sufficient to explain the variance present within the data and to justify the spread in glare responses reported by test subjects. Rather, there may have been other variables influencing glare sensation beyond the fixed-effects. In order to systematically identify these further effects (Rodriquez & Pattini, 2014; Tregenza & Wilson, 2011), the following stage of this study investigated the potential influence of temporal variables and personal factors on the level of glare sensation from daylight reported by test subjects.

4.2. Temporal Variables, Personal Factors, and Glare Sensation from Daylight

4.2.1. Introduction

This stage of the investigation aimed at analysing the effect of several temporal variables and personal factors on the remaining between-subject and within-subject variance within the original MLM. Respectively, the variance was calculated based on the Level 2, 3, 4, and 5 variance in DGP scores (τ_{00}, analogue to the between-groups sums of squares (SS_B)), and the Level 1 variance in DGP scores (σ^2, analogue to the within-groups sums of squares (SS_W)).

Based on previous experiments (see Section 3.1.3, Section 3.2.4 and Section 3.4.3) – which systematically identified potential causes of between-subject variance within each conditional group (i.e., Morning, Afternoon, and Evening test sessions) related to the independent variable (i.e., time of the day) – several temporal variables and personal factors were postulated to also cause random variation within the glare responses provided by subjects at various times of day, under the two shading settings (default and user-set), and while performing three visual tasks (landolt ring, letter searching, and typing). When evaluating the nature of the research question (i.e., the effect of time of day on glare response), it is essential that the central tendencies corresponding to each independent variable are isolated by minimising the deviation (i.e., variance) associated with them (i.e., $X_i - \bar{X}$, where X_i is the calculated DGP value given by individual test subjects and $\bar{X}$ is the mean DGP value) (Seltman, 2010). Moreover, it is important to highlight that both sources of variability (i.e., within-subject and between-subject variance) need to be considered when making inferences related to descriptive and inferential statistics. However, fixed-effects approaches only take into account the within-subject variance (Penny & Holmes, 2003). Conversely, mixed-effects approaches consider both sources, and therefore can account for the total variation (i.e., $\sigma^2 + \tau_{00}$) within the data, providing much more reliable parameter (mean) estimates.

4.2.2. Methods

4.2.2.1. Temporal Variables and Personal Factors

At the beginning of the test procedure (Section 4.1.2.1.3), subjects were requested to fill in a short questionnaire featuring demographic information (age, gender, and ethnicity) and to provide self-assessments of personal factors (photosensitivity and chronotype) and temporal variables (fatigue, hunger, caffeine ingestion, mood, prior light exposure – distinguished in direct, diffuse, and artificial – and sky condition). Coherent with previous laboratory studies, to assess photosensitivity, test subjects were asked questions regarding their self-evaluated sensitivity to natural and artificial light, their use of solar protection devices (e.g., sunglasses), their luminous preference within the indoor environment (e.g., bright or dim conditions), and the frequency of their interactions with environmental controls (e.g., blinds) (Aries, 2005). Likewise, the Samn-Perelli scale (Samn & Perelli, 1982) was utilised to measure subjective levels of fatigue, and the Munich Chronotype Questionnaire (MCTQ) (Roenneberg *et al.*, 2003) was used to collect self-assessments of chronotype. Since the measurement of both these latter variables is based on well-established methodologies structured on 7-point scales, when feasible all other variables were also measured on 7-point ratings of assessment. However, when it was not practical to measure a variable under a 7-point scale (e.g., in the case of caffeine ingestion), a dichotomous criterion of evaluation was used, thereby forcing test subjects to provide a discrete outcome (e.g., yes/no). The main benefit of using dichotomous scales – rather than a continuous, ordinal, interval or ratio measurement – is that, by forcing the dependent variable into discrete categories, differences can be isolated more straightforwardly when evaluating the effect under examination using a statistical inferential test (Field & Hole, 2013). For this experiment, a short synthetic descriptor was provided for each point of evaluation on the scales so as to designate a given level of magnitude for the effect under examination, thereby reducing the self-interpretation of the outcome.

4.2.3. Statistical Analysis

4.2.3.1. Mixed-Model

To evaluate the data, a mixed-effects MLM analysis with both 'fixed-effects' and 'random-effects' was performed, comparing the DGP values for the variables time of day, shading setting, and GSV against each other, while controlling for the effect of glare source size and position. Based on the literature (Hayes, 2006; Seltman, 2014), when the Wald Z statistic test is significant (in this study, Wald Z= 12.28, $p \leq 0.001$), there is evidence that the remaining variance within the MLM cannot be explained only by the specified fixed-effects (i.e., time of day, shading setting, task type, and GSV), thus suggesting that further random-effects need to be considered (Field, 2013; Seltman, 2014). However, the conditions upon which a variable should be specified as either fixed- or random- effect are not always clear in the literature (Snijders, 2005).

In the context of MLM analysis and time-based studies, a fixed-effect is generally a factor that does not change over time (e.g., gender and other personal factors such as photosensitivity, chronotype, etc.); these effects are often referred to as fixed variables. In contrast, a random-effect is likely to fluctuate over time (e.g., the temporal variables provided by test subjects in this study, such as fatigue, hunger, sky condition, etc.); hence, within the data collected, these variables can be assumed to randomly change from individual to individual (Field, 2013). In essence, the variables within a model can be either observed since they occur by chance (a random-effect), or they may vary by experimental control (a fixed-effect) (Snijders & Bosker, 1994). One of the main differences between specifying a variable as a fixed- or random- effect is represented by the calculation of the variance parameters (Beretevas & Pastor, 2003), since additional sources of variability are accounted for by the inclusion of random-effects into the model. The standard errors in the fixed-effect model tend to be underestimated due to the fact that additional causes of variance (caused by random-effects), contributing to the reliability estimates, are not included. As a result, fixed-effects models tend to have higher 'perceived' statistical power, causing an inflation of the test statistics and an elevation of the Type I error rate (Kreft & de Leeuw, 1998).

Within the previous MLM analysis (that only considered fixed-effects), the ML (Maximum Likelihood) method was adopted to estimate the variance parameters. In order to compare the mixed-effects MLM with the fixed-effects model by using a likelihood ratio test (testing whether the explained variances in both models are statistically different from each other), it is important that the sample sizes do not differ, that the same fixed-effects are used, and that the ML method is specified in both models (Field, 2013; Peugh, 2010).

As mentioned, the likelihood ratio test determines whether the explained variances in both models are statistically different from each other, thereby evaluating the inclusion of random-effects within the MLM in comparison to the fixed-effects only model. For this study, this was calculated by making use of the difference – defined as the 'deviance' – in the Schwarz's Bayesian Criterion (BIC) extrapolated from both the fixed-effects only MLM and the mixed-effects MLM models (Equation 4-21), and the degrees of freedom (*df*) in each MLM (Equation 4-22) (Enders, 2010; Field, 2013):

$$\chi^2_{Change} = (BIC_{(fixed-effects)}) - (BIC_{(mixed-effects)}) \qquad \text{(Equation 4-21)}$$

$$df_{Change} = k_{fixed-effects} - k_{mixed-effects} \qquad \text{(Equation 4-22)}$$

Where k is the number of parameters in each model.

The difference between the deviance (ΔBIC) is an approximation of the chi-squared (χ^2) distribution with degrees of freedom equal to the number of random-effects included in the mixed-effects MLM. In interpreting the outcome, it is important to mention that the likelihood ratio test is effectively a null hypothesis significance test. Therefore, it inherits the same limitations of the *p*-value, being dependent on both the size of the effect under examination and the size of the sample (Ellis, 2010). A study by Peugh (2010) on student grades based on a large dataset (N= 12,144) detected statistical significance in the likelihood ratio test ($\chi^2(1)= 5.51$, $p<0.05$) when examining the differences in variances associated with the random-effect of 'gender' across two MLM models. However, since a large sample size produced such as small chi-squared value (χ^2), the author concluded that gender had a relatively minor effect. Coherent with this finding, the emphasis of the inferential analysis for the likelihood ratio test was placed on the effect size, measured in this case by the pseudo squared partial correlation (r^2). This was calculated by making use of the quantified proportion of variance remaining in the model (residual variance (σ^2)) after accounting for the variability caused by the random-effects (within the mixed-effects model MLM (σ^2_{new})) and the variability explained only by the fixed-effects (within the fixed-effects only MLM (σ^2_{old})), according to the following formula (Equation 4-23) (Hayes, 2006; Peugh, 2010)):

$$Pseudo\ r^2 = 1 - \frac{\sigma^2_{new}}{\sigma^2_{old}} \qquad \text{(Equation 4-23)}$$

Where the pseudo r^2 benchmarks the variance explained relative to the total variance.

Also for this analysis, the tables by Ferguson (2009) provided values for small, moderate, and large effect sizes ($r^2 \geq 0.04$, 0.25, and 0.64, respectively) in order to support the interpretation of the outcome.

4.2.4. Results and Discussion

Utilising Equations 4-21 and 4-22, the likelihood ratio test detected a highly significant difference: $\chi^2(8)= 180.19$, $p\leq 0.001$, $r^2= 0.51$ (moderate), suggesting that the variances associated with the random-effects are significantly different from zero. This result provides statistically and practically relevant evidence to postulate that the inclusion of the random-effects represented by the temporal variables self-reported by test subjects – i.e., fatigue', 'hunger', 'caffeine intake', 'mood', 'direct exposure', 'diffused exposure', 'artificial exposure', and 'sky condition' – into the mixed-effects model offers a significantly better fit to the data than the fixed-effects only model. The inclusion of random-effects on the variance showed a 'moderate' ($0.25\leq r^2<0.64$) effect size. Therefore, by including the random-effects into the MLM, consideration of temporal variables can explain 51% (i.e., $r^2= 0.51$) of the unexplained variance that was not accounted for within the original fixed-effects only MLM.

For this analysis, the DGP was again used as the primary evaluation parameter to assess the variances associated with each temporal variable (random-effects). However, unlike other photometric metrics (i.e., illuminance at the eye, average luminance) and glare indices (i.e., DGI, UGR), the DGP is characterised by small numerical values that have a narrow range of variation (the DGP is always comprised between 0 and 1). Hence, the variances associated with the DGP are relatively little. Estimates of the covariance parameters were calculated by taking the square root of the estimated variance and making use of the standard deviation – which provides a statistical value that is easier to interpret, i.e. the amount of spread that the random-effect can explain within the model – for each temporal variable. Table 4-11 presents the standard deviation for each temporal variable, the standard error, the Wald Z test statistic, and the lower (CI_L) and upper (CI_U) 95% confidence intervals for the standard deviation associated with each random-effect calculated from the multivariate Wald test.

Table 4-11 Estimates of covariance parameters for each temporal variable

Variable	Standard Deviation	Std. Error	Wald Z	CI_L	CI_U
Fatigue	0.01	0.01	1.58	0.00	0.02
Hunger	0.01	0.01	1.61	0.00	0.02
Caffeine Intake	0.03	0.02	1.85	0.02	0.06
Mood[a]	0.00	0.00	-	-	-
Direct Exposure[a]	0.00	0.00	-	-	-
Diffused Exposure	0.01	0.01	1.62	0.00	0.02
Artificial Exposure	0.00	0.00	0.39	0.00	0.03
Sky Condition	0.01	0.01	3.15	0.01	0.02

* weakly significant; ** significant; *** highly significant; n.s. not significant

[a.] Covariance parameter is redundant. The test statistic and confidence intervals cannot be computed

The standard deviation (descriptive) and the Wald Z (inferential) statistics both provide a measure of the variance that each temporal variable causes on the DGP. With reference to the standard deviation, the results indicate that ‘caffeine intake’ causes the highest amount of variance in calculated DGP values (SD= 0.03). Conversely, with reference to the Wald Z test statistic, the variable ‘sky condition’ is associated with the highest amount of variance in the DGP (Wald Z= 3.15). This discrepancy is likely due to differences in the scaling of the self-reported levels of the temporal variables, whereby ‘caffeine intake’ was measured on a discrete dichotomous scale, while all other variables were measured on a 7-point Likert scale. However, since the Wald Z is a standardised value – comparable across temporal variables – the findings suggest that ‘sky condition’ is effectively the variable that can explain the highest amount of the residual variance remaining in the MLM, which cannot be explained by the fixed-effects. That is, the variance in DGP at a between-subject level (expressed by the personal regression slopes for each individual test subject – i.e., Level 1 of the fixed-effects MLM) is the largest when individual test subjects were exposed to different sky conditions while reporting their glare sensation for each fixed-effect specified in the MLM.

In the interpretation of the outcomes, the statistical significance associated with the Wald Z test has not been included in the inferential analysis. This is due to the fact that the Wald Z test assumes that sampling distributions of the parameters (e.g., variances, covariances, correlations) are normal; however, this is an assumption that, based on the literature, may lead to invalid conclusions (Raudenbush & Bryk, 2002). In addition, in the case of small sample sizes, the Wald Z test may be susceptible to Type I errors (Buse, 1982; Engle, 1984). To fully address these concerns, various studies recommend that the likelihood ratio test should also be used, other than the Wald Z test, to make more robust inferences from the inclusion of random-effects, by deriving statistical significance from the deviance of two models (χ^2_{change}) (Field, 2013; Hayes, 2006; Peugh, 2010). In fact, the likelihood ratio test makes no assumptions about the shape of the sampling distribution (Enders, 2010).

Figure 4-18 plots, on the y-axis, the adjusted DGP marginal means (with 95% confidence intervals) controlled for the effect of glare source size and position, and for the variances associated with the temporal variables included as random-effects within the MLM. On the x-axis, the figure presents the time of day when glare assessments were provided, distributed according to reported thresholds of glare response. Coherent with previous analysis (i.e., the MLM with fixed-effects), due to relatively small sample of votes at the highest levels of glare sensation, the GSV criteria of ‘Just Uncomfortable’ and ‘Just Intolerable’ were merged so that meaningful statistical analysis could be performed (Kuhn *et al.*, 2013).

For each criterion of GSV, the mean plots are distributed according to the setting of the shading device, corresponding respectively to the Default shading and the User-set shading. The dependent variable (DGP) was again evaluated at an adjusted solid angle subtended by the glare sources modified by the position index for all test sessions and at all reported levels of glare sensation. This value was calculated by the statistical package (SPSS) and was an adjustment derived by the MLM and adopted for all inferential tests (see Section 4.1.2.1.5 and Section 4.1.4).

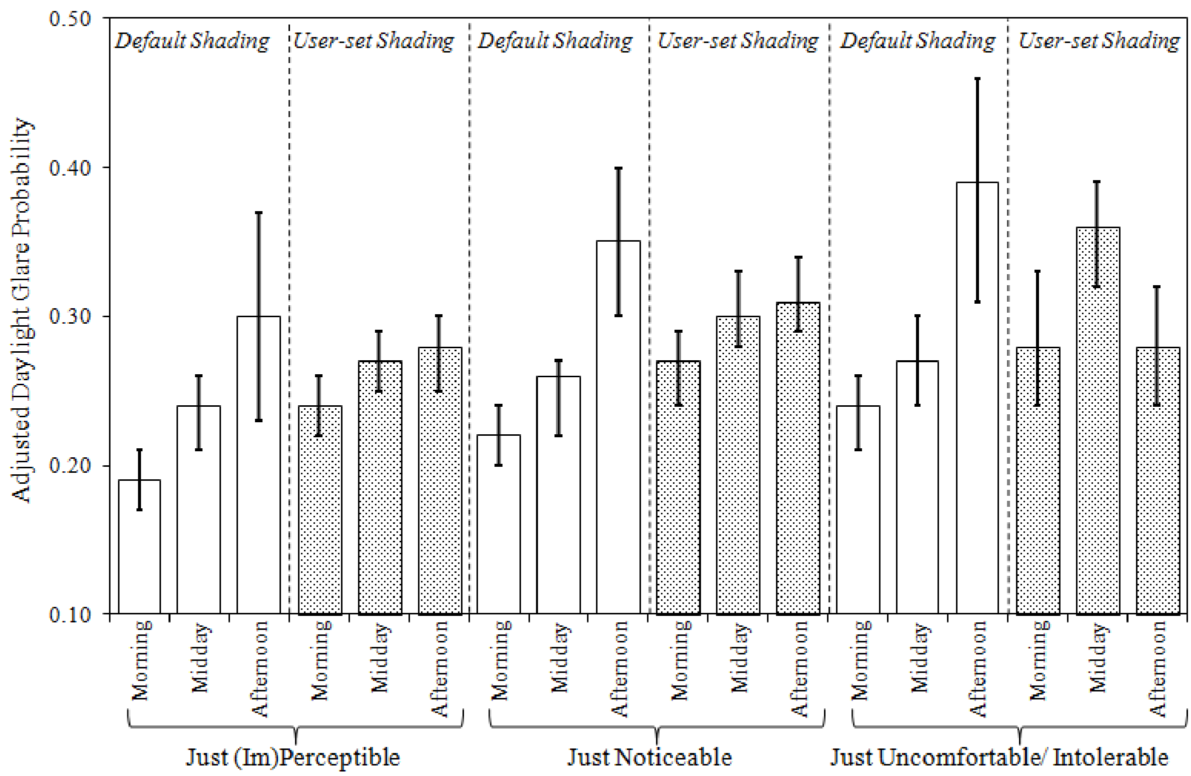


Figure 4-18 Mean plots with 95% confidence intervals of adjusted DGP values (y-axis) for each GSV criterion, test session, and shading setting (x-axis); mixed-effects MLM

Consistent with Figure 4-16 (fixed-effects only MLM), graphical inspection of the boxplots of Figure 4-18 suggests that, under the default shading setting, there is a consistent trend for central tendencies to correspond to higher levels of DGP for each criterion of glare sensation as the day progresses. Since a higher DGP signals a larger probability that an observer may be 'distributed' by the glare source, the boxplots lead to the hypothesis that, when the Venetian blinds were set at their 'cut-off' position, subjects became more tolerant to the combination of photometric and physical parameters associated with the glare source(s) along the day. In fact, the same glare sensation vote given by subjects (e.g., Just (Im)Perceptible) corresponds to increasing levels of mean DGP at later test sessions (from Morning to Midday to Afternoon). This confirms the outcomes of the fixed-effects MLM analysis. Contrary to the fixed-effects MLM, in Figure 4-18 an equal trend of central tendencies corresponding to higher levels of DGP as the day progresses can also be observed under the user-set shading setting for the GSV criteria of Just (Im)Perceptible and Just Noticeable. However, from graphical inspection of the data, this trend for the user-set shading does not seem to be as strong as for the default shading. Conversely, for the criterion of Just Uncomfortable/Intolerable, the visual displays from the mixed-effects MLM do not allow a univocal interpretation of a prevailing tendency along test sessions, hence providing no evidence of an effect of time of the day when subjects reported the highest levels of glare sensation after having controlled the Venetian blinds to their own preferences.

In the mixed-effects MLM, consideration of the effect of time of the day (i.e., Morning vs. Midday vs. Afternoon) – while controlling for the glare source size and position, and for the variances associated with the temporal variables – detected no statistically significant difference for the DGP between test sessions: $F(2, 326.10)= 2.51$, $p= 0.08$ n.s. This result offers no evidence that the adjusted mean DGP values calculated at the three test sessions were different from each other, hence potentially challenging any effect of time of the day on glare response. However, it must be considered that this test does not take into account the evaluations of glare sensation provided by subjects (GSVs). Previous analysis (fixed-effects MLM) had detected a highly significant difference for the DGP between test sessions ($F(2, 571.46)= 9.93$, $p\leq 0.001$) using the same sample, fixed-effects, covariate (subtended angle of the glare source modified by its position index), and estimation method (ML). Essentially, this indicates that inclusion of random-effects and their associated variance (not accounted in the fixed-effects model) leads to more robust inferential analysis in the mixed-effects MLM.

Follow-up univariate tests were then performed by grouping the effect of time of day by the GSV reported by test subjects and by shading setting. Table 4-12 shows the inferential data from the univariate tests, presenting the shading setting (Default and User-set shading), the GSV criteria (‘Just (Im)Perceptible’, ‘Just Noticeable’, and ‘Just Uncomfortable/Intolerable’), the degrees of freedom (df), the test statistic (F), and the statistical significance (p-value). As in the fixed-effects MLM, the influence of task type resulted in no statistically significant difference as a main effect, nor did it demonstrate any significant interaction with any other fixed-effect. Hence, task type was excluded from any subsequent analysis.

Table 4-12 Univariate tests comparing DGP values under all times of day for each GSV criterion and shading setting (mixed-effects MLM)

Shading Setting	GSV	Numerator df	Denominator df	***F***	*p*-value
Default Shading	Just (Im)Perceptible	2	483.19	6.90	0.00***
	Just Noticeable	2	505.79	21.21	0.00***
	Just Uncomfortable/Intolerable	2	575.36	10.90	0.00***
User-set Shading	Just (Im)Perceptible	2	358.37	9.61	0.00***
	Just Noticeable	2	581.18	4.43	0.01**
	Just Uncomfortable/Intolerable	2	551.30	7.89	0.00***

* weakly significant; ** significant; *** highly significant; n.s. not significant

Under the default shading setting, the results of Table 4-12 show 3 highly significant differences. Under the user-set shading setting, the tests detected 2 highly significant and 1 significant differences. The univariate tests, therefore, provide evidence of statistically significant differences in the DGP values at different times of day under both shading settings.

To isolate the main effects between variables, contrast were made using pairwise comparisons (Bird & Hadzi-Pavlovic, 2014), whereby all permutations between times of the day were compared against each other. In the fixed-effects MLM, examination of the descriptive statistics and visual inspection of central tendencies from graphical displays (Hauschke & Steinijans, 1996) had detected no univocal directionality between observed differences, hence suggesting the adoption of a two-tailed hypothesis (Ruxton & Neuhauser, 2010). Coherent with previous analysis, a two-tailed alternative hypothesis was also applied under a mixed-effects MLM to the pairwise comparisons performed for all times of day, GSV criteria, and shading settings. In consideration of the experiment-wise error rate caused by the significance level inflating across multiple tests carried out on the same data (Cabin &

Mitchell, 2000), Bonferroni corrections were applied. Emphasis of the analysis was placed on the effect size, estimated by the Cohen's d coefficient and interpreted on the benchmarks by Ferguson (2009) for small, moderate, and large effects (d≥ 0.41, 1.15 and 2.70, respectively). The effect size was calculated from the estimated marginal means and pooled standard deviation scores adjusted for the effect of glare source size and position, and for the variances associated with each temporal variable in the mixed-effects MLM, according to the following formula (Equation 4-24) (Fields, 2013; Rosenthal & DiMatteo, 2001):

$$Cohen's\ d = \frac{\Delta M}{\sigma pooled} \qquad \text{(Equation 4-24)}$$

Where ΔM is the difference between the estimated marginal means, and σpooled is the pooled standard deviation.

For each GSV, Table 4-13 to Table 4-15 report the shading setting, the sample size (N) of the pairwise comparisons (x_0 and x_1, corresponding to the test sessions), the difference between the estimated adjusted marginal means (ΔM) and its two-tailed statistical significance (*p*-value), the standard error, the degrees of freedom (df), the lower (CI_L) and upper (CI_U) 95% confidence intervals for the difference between marginal means, and the effect size (d).

Table 4-13 Pairwise comparisons and effect sizes for 'Just (Im)Perceptible'

Shading	Time of Day	$N(x_0, x_1)$	ΔM^{NHST}	Std. Error	df	CI_L	CI_U	Effect Size (d)
Default Shading	Morning vs. Midday	49, 37	-0.05**	0.01	448.67	-0.07	-0.02	-0.77
	Midday vs. Afternoon	37, 45	-0.06 n.s.	0.04	517.88	-0.14	0.03	-0.31
	Morning vs. Afternoon	49, 45	-0.11*	0.04	511.83	-0.17	-0.01	-0.56
User-set Shading	Morning vs. Midday	72, 72	-0.03***	0.01	393.96	-0.05	-0.01	-0.38
	Midday vs. Afternoon	72, 84	-0.01 n.s.	0.01	331.90	-0.03	0.02	-0.12
	Morning vs. Afternoon	72, 84	-0.04***	0.01	423.84	-0.06	-0.01	-0.47

* weakly significant; ** significant; *** highly significant; n.s. not significant
d < 0.41= negligible; 0.41 ≤ d < 1.15= small; 1.15 ≤ d < 2.70= moderate; d ≥ 2.70= large

Table 4-14 Pairwise comparisons and effect sizes for 'Just Noticeable'

Shading	Time of Day	$N_{(x_0, x_1)}$	ΔM^{NHST}	Std. Error	df	CI_L	CI_U	Effect Size (d)
Default Shading	Morning vs. Midday	47, 61	-0.04***	0.01	439.42	-0.06	-0.02	-0.53
	Midday vs. Afternoon	62, 44	-0.09***	0.03	582.54	-0.16	-0.04	-0.59
	Morning vs. Afternoon	47, 44	-0.13***	0.03	559.85	-0.20	-0.08	-0.87
User-set Shading	Morning vs. Midday	43, 38	-0.03*	0.01	491.42	-0.06	-0.01	-0.46
	Midday vs. Afternoon	38, 26	-0.01 n.s	0.02	626.91	-0.04	0.03	-0.12
	Morning vs. Afternoon	43, 26	-0.04 n.s.	0.02	613.83	-0.07	0.00	-0.47

* weakly significant; ** significant; *** highly significant; n.s. not significant
d < 0.41= negligible; 0.41 ≤ d < 1.15= small; 1.15 ≤ d < 2.70= moderate; d ≥ 2.70= large

Table 4-15 Pairwise comparisons and effect sizes for 'Just Uncomfortable/Intolerable'

Shading	Time of Day	$N_{(x_0, x_1)}$	ΔM^{NHST}	Std. Error	df	CI_L	CI_U	Effect Size (d)
Default Shading	Morning vs. Midday	24, 22	-0.03*	0.02	584.25	-0.07	-0.01	-0.41
	Midday vs. Afternoon	22, 31	-0.12**	0.04	576.32	-0.20	-0.03	-0.71
	Morning vs. Afternoon	24, 31	-0.15***	0.04	558.94	-0.24	-0.07	-0.94
User-set Shading	Morning vs. Midday	5, 10	-0.08**	0.02	557.35	-0.13	-0.02	-1.57
	Midday vs. Afternoon	10, 10	0.08***	0.02	570.44	-0.03	0.13	1.33
	Morning vs. Afternoon	5, 10	0.00 n.s.	0.02	573.90	-0.06	0.06	0.00

* weakly significant; ** significant; *** highly significant; n.s. not significant
d < 0.41= negligible; 0.41 ≤ d < 1.15= small; 1.15 ≤ d < 2.70= moderate; d ≥ 2.70= large

Consistent with the findings from the fixed-effects MLM (see Section 4.1.4) and from previous laboratory experiments (Section 3.1.3, Section 3.2.4, and Section 3.4.3), analysis of the descriptive and inferential statistics shows that, under the default shading setting, the mean differences (ΔM) and the effect sizes (d) are consistently negative. This signals an increased value of DGP as the day progresses at each reported level of glare sensation. Under the default shading setting, the pairwise comparisons provide evidence that the ΔMs are highly significant in 4 cases, significant in 2 cases, weakly significant in 2 cases, and not significant in 1 case. The differences detected have all a practically relevant magnitude (0.41≤d<1.15), except for 1 case with negligible effect size. The inferential tests, therefore, confirm the postulated hypothesis of a tendency for the DGP to increase as the day progresses under the default shading setting for all criteria of reported glare response. Inspection of descriptive statistics (ΔM) and effect sizes under the user-set shading setting shows evidence of consistent negative signs for both values, with the exception of the 'Just Uncomfortable/Intolerable' GSV, for which the results do not allow the definition of a prevailing tendency. For the 'Just (Im)Perceptible' and 'Just Noticeable' benchmarks of glare sensation, the data signal an increased value of DGP as the day progresses. The pairwise comparisons provide evidence that the ΔMs under the user-set shading setting are highly significant in 3 cases, significant in 1 cases, weakly significant in 1 case, and not significant in 4 cases. The differences detected have moderate effect size (1.15≤d<2.70) in 2 cases and small magnitudes (0.41≤d<1.15) in 3 cases. Therefore, also under the user-set shading – once the influences of temporal variables varying across the fixed-effects (time of day, shading setting, and report GSV) are controlled – the results support the postulated hypothesis of an effect of time of the day on the evaluation of glare sensation reported by test subjects. Figure 4-19 plots, for each GSV and shading setting, the effect sizes (d) extracted from Table 4-13 to Table 4-15, the associated mean differences and their statistical significance.

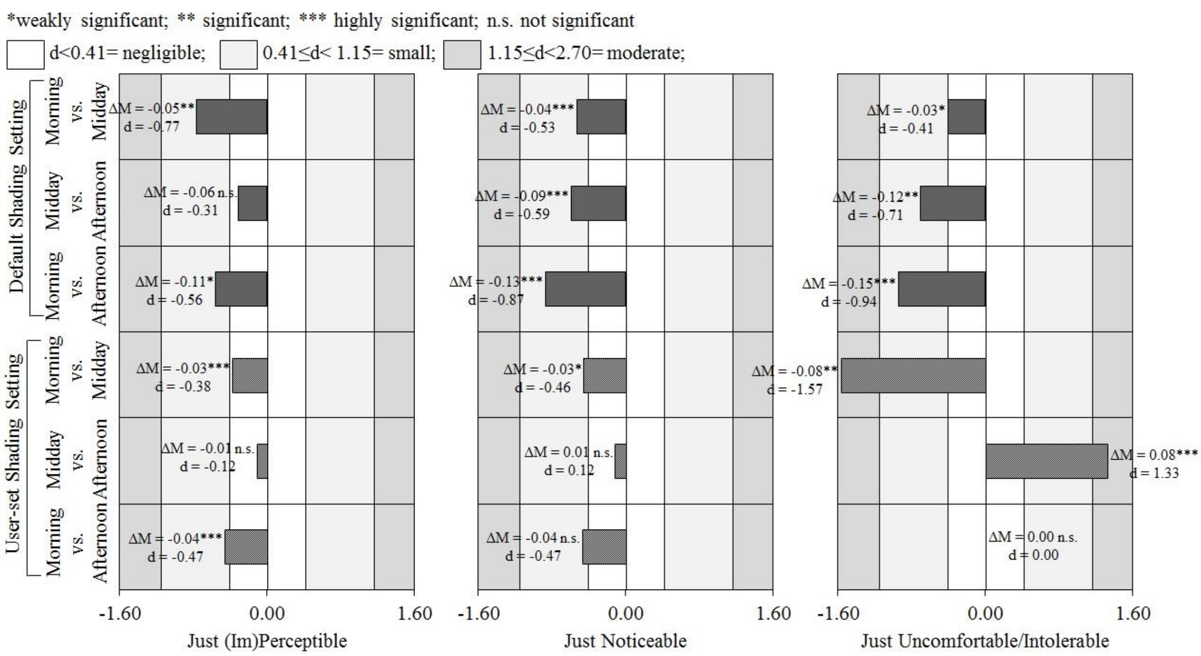


Figure 4-19 Effect sizes (d) detected in the pairwise comparisons (see Tables 4-13 to 4-15) between test sessions (y-axis), for each GSV criterion (x-axis) under the default (top) and user-set (bottom) shading settings; mixed-effects MLM

Visual inspection of the boxplots presented in Figure 4-19 reveals, under the default shading setting, consistently negative effect sizes, these signalling increased values of DGP as the day progresses. Under this setting of the Venetian blinds, the boxplots show a tendency for the magnitudes of effect size to amplify with the increase in the temporal gap between test sessions, with the exception of the 'Just (Im)Perceptible' threshold of glare sensation (left).

Under the user-set shading setting, the boxplots for the 'Just (Im)Perceptible' (left) and 'Just Noticeable' (middle) GSVs also show negative effect sizes, whose magnitudes increase for comparisons between sessions separated by a larger time interval. This supports the postulated hypothesis of an increased tolerance to discomfort glare as the day progresses. However, the results for the 'Just Uncomfortable/Just Intolerable' GSV under the user-set shading do not lead to any consistent interpretation of a prevailing trend within the data.

In essence, the results of the mixed-effects MLM analysis provide statistically significant and practically relevant evidence to suggest an effect of time of the day on the level of glare sensation reported by subjects in a test room with direct access to daylight. With respect to the previous findings from the fixed-effects MLM (Section 4.1.4), the outcomes of the mixed model analysis demonstrate that, when controlling for the effects of temporal variables, indications of a direct influence of time of day on glare response can be detected for both, the default and the user-set shading settings. In addition, the findings support the hypothesis that the effect of time of the day amplifies along the test sessions, thereby suggesting that subjects became more tolerant to discomfort glare as the day progresses.

It is worth to remind here that the fixed-effects MLM had provided no evidence of an effect of time of the day on glare response when users were given the opportunity to control the Venetian blinds based on their preferences. Conversely, the influences of the random-effects within the mixed-effects MLM suggest that the variances associated with the considered temporal variables partially confound the effect of time of day on glare sensation. That is, once the variances of temporal variables were controlled – hence, increasing the sensitivity of the inferential tests (Field, 2013) – the effect of time of day on glare response could be detected also under the user-set shading setting. However, since the magnitudes of the observed temporal influences (effect sizes) are in general smaller under the user-set shading (see Table 4-13 to Table 4-15) than under the default shading, it is plausible that – as previously noted (see Section 4.1.4) – when the blinds were adjusted by test subjects, the presence of other uncontrolled variable(s) or condition(s) may have potentially masked or confounded the magnitude of the temporal influence on the perception of glare.

Interestingly, when considering the default shading setting, the effect sizes calculated from the pairwise comparisons in the fixed-effects and in the mixed-effects MLM models do not appear to differ substantially (see Figure 4-16 and Figure 4-16). This seems to suggest that the temporal variables may exert greater influence on glare sensation once the test subjects adjusted the Venetian blinds to their own visual preference.

4.2.4.1. Chronotype

Unlike the temporal variables considered in the mixed-effects MLM analysis (e.g., fatigue, mood, sky condition, etc.), chronotype does not vary over the fixed-effects under examination, and therefore – in accordance with the literature – it has been classified as a fixed-effect in this study (Snijders, 2005; Snijders & Bosker, 1994). The aim of the analysis presented in this section is to determine whether the chronotype of subjects had any moderating impact on the detected effect of time of day on glare response in the presence of daylight from a window (refer to Section 3.3.3.2). For each chronotype, this analysis sought to explore whether a similar variation of tolerance to discomfort glare as the day progresses can be detected, and investigate if differences in variation of tolerance between chronotypes can be observed.

Figure 4-20 and Figure 4-21 display selected mean plots in consideration of chronotype for the reported level of glare sensation of 'Just (Im)Perceptible' under the Default Shading (Figure 4-20) and the User-set Shading settings (Figure 4-21). These visual displays were selected among others since they suggest, in a clearer way than for the other thresholds of glare sensation, a moderating effect between the variables considered (time of day and chronotype) under the default shading setting. Conversely, under the user-set shading, this effect appears to be confounded (possibly due to the presence of other variables).

To remind here that the chronotype of test subjects was identified via the MCTQ (Roenneberg, 2003) self-assessment, which categorises chronotype using a 7-point ordinal scale ranging from 'Extremely Early'= 1 to 'Extremely Late'= 7. In the test-room experiment, the distribution of test subjects (N= 40) according to their self-assessed chronotype was as follows: Type 1= 3, Type 2= 8, Type 3= 3, Type 4= 7, Type 5= 8, Type 6= 10, and Type 7= 1.

The figures present, on the y-axis, the DGP values adjusted for glare source size and position calculated at the three test sessions when subjects reported the glare sensation of 'Just (Im)Perceptible'. On the x-axis, the figures present the chronotypes of the test subjects. Since the distribution of votes was low under certain categories of chronotype, some of them were combined together so that meaningful statistical analysis could be performed. Specifically: Type 1, Type 2, and Type 3 were merged into a grouping corresponding to 'Early Chronotype' (Type 1, 2, 3); Type 5, Type 6, and Type 7 were combined into a classification of 'Late Chronotype' (Type 5, 6, 7). Lastly, Type 4 was retained in its original discrete category referring to a chronotype that is 'Neither Early nor Late'. For each grouping of chronotypes, the plots are distributed according to the time of the day in which the glare sensation vote was reported, corresponding respectively to the Morning, Midday, and Afternoon test sessions.

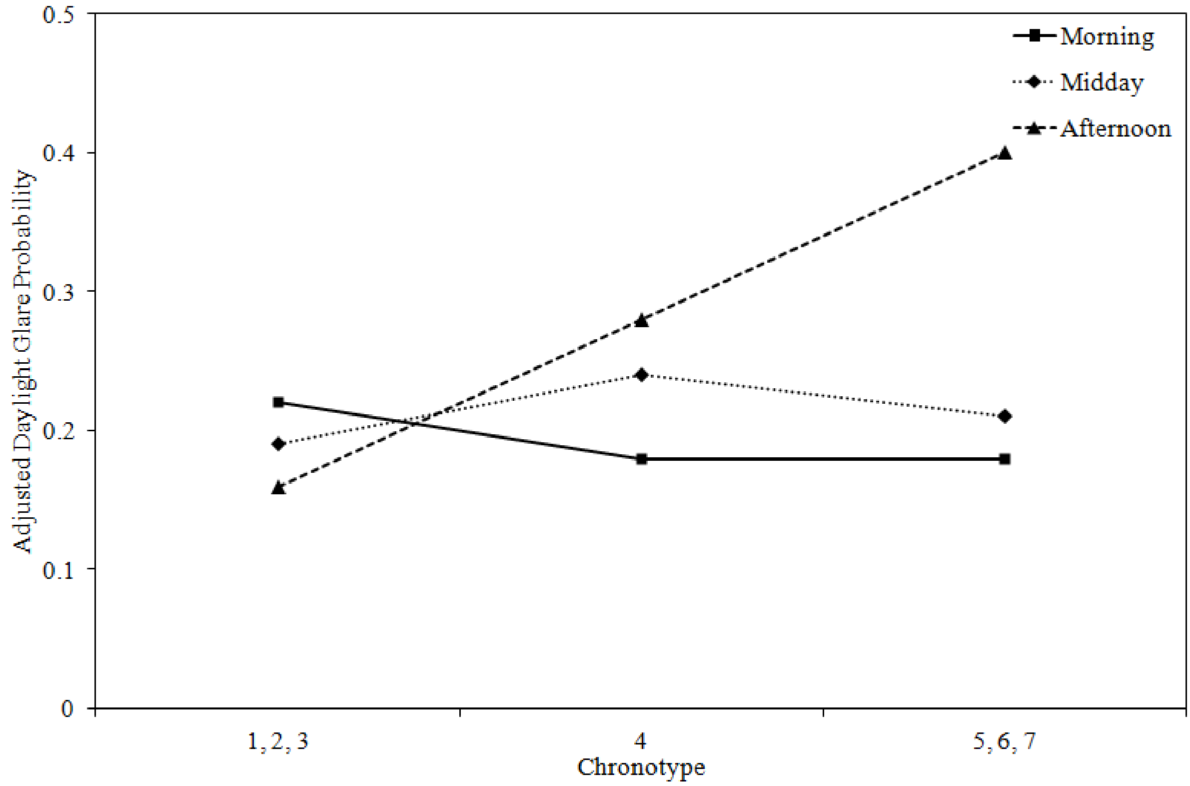


Figure 4-20 Mean plots of adjusted DGP values (y-axis) for groupings of chronotype (x-axis), for the 'Just (Im)Perceptible' GSV criterion under the default shading setting

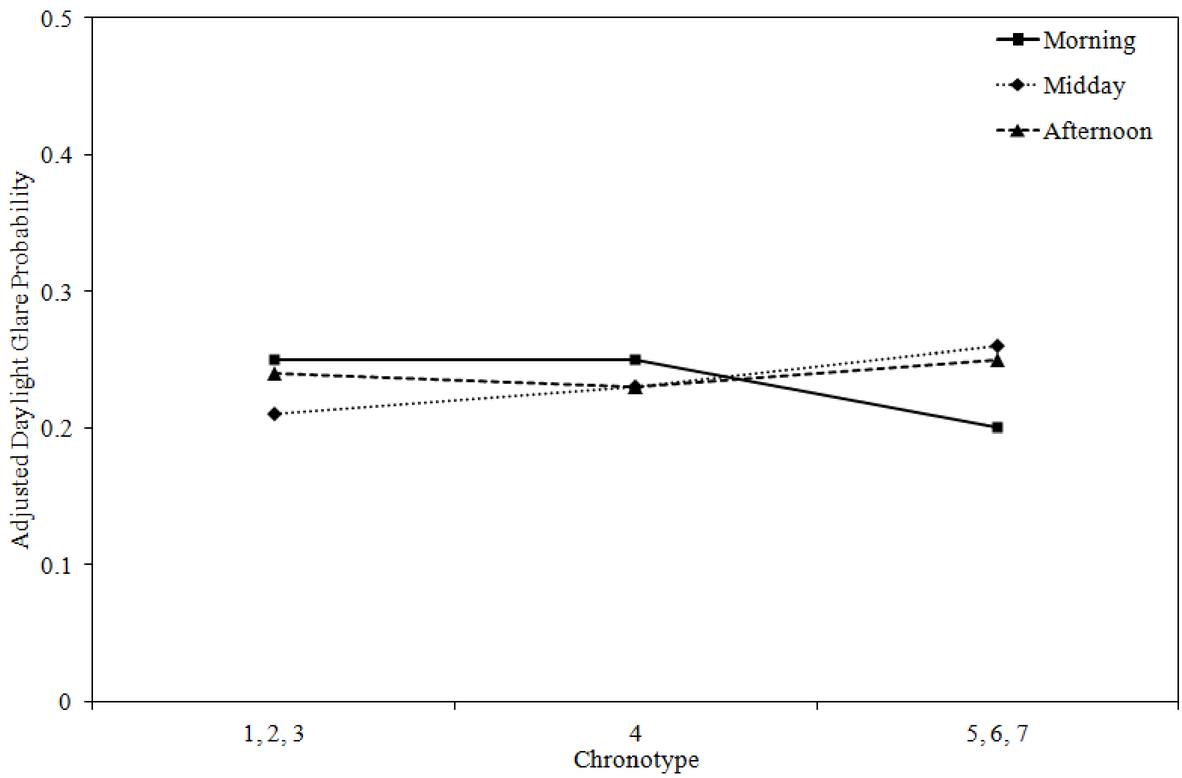


Figure 4-21 Mean plots of adjusted DGP values (y-axis) for groupings of chronotype (x-axis), for the 'Just (Im)Perceptible' GSV criterion under the user-set shading setting

For the GSV criterion of ‘Just (Im)Perceptible’ under the default shading setting (Figure 4-20), graphical inspection of Figure 4-20 suggests that, when considering the ‘Early Chronotype’ group (Type 1, 2, 3), there appears to be an inverse interaction between variables (time of day and chronotype). That is, as the day progresses, the plots suggest a decreased tolerance to discomfort glare (corresponding to decreasing values of DGP at later sessions). Conversely, when considering Type 4, ‘Neither Early nor Late’, and the Type 5, 6, 7 group ‘Late Chronotype’, the displays show an increasing tolerance to discomfort glare as the day progresses. Under the user-set shading (Figure 4-21), the relative ‘flatness’ and close proximity of the interaction lines associated with the mean points do not enable a consistent and univocal interpretation of the data for any of the groups of chronotype.

To examine the data, a fixed-effect MLM was again selected featuring the same fixed effects of previous analyses with the addition of the level of chronotype, while controlling for the effects of glare source size and position. The fixed-effects specified in the MLM were:

- Time of day;
- Shading setting;
- Task type;
- GSV reported by the test subjects;
- Chronotype;

The addition of chronotype at the lowest level of the MLM does not make this inferential model comparable to that presented in Section 4.1.3 (Field, 2013; Peugh, 2010).

Table 4-16 presents, for all shading settings, criteria of glare sensation, and groups of chronotype, the degrees of freedom (df), the *F* test statistic, and the statistical significance (*p*-value) of the univariate analysis.

Table 4-16 Univariate tests comparing DGP values under all times of day for each GSV criterion and shading setting for each grouping of chronotype

Shading	GSV	Chronotype	Numerator df	Denominator df	***F***	*p*-value
Default Shading	Just (Im)Perceptible	1, 2, 3	2	597.98	0.51	0.60 n.s.
		4	2	629.29	4.02	0.02*
		5, 6, 7	2	635.20	12.23	0.00***
	Just Noticeable	1, 2, 3	2	647.76	8.09	0.00***
		4	2	641.78	6.94	0.00***
		5, 6, 7	2	642.54	7.37	0.00***
	Just Uncomfortable/ Intolerable	1, 2, 3	2	611.25	6.91	0.00***
		4	2	652.84	2.74	0.07 n.s.
		5, 6, 7	2	649.16	4.74	0.01**
User-set Shading	Just (Im)Perceptible	1, 2, 3	2	417.11	2.81	0.06 n.s.
		4	2	456.12	0.17	0.84 n.s.
		5, 6, 7	2	471.29	11.32	0.00***
	Just Noticeable	1, 2, 3	2	571.63	1.13	0.33 n.s.
		4	2	625.97	0.56	0.57 n.s.
		5, 6, 7	2	612..68	13.66	0.00***
	Just Uncomfortable/ Intolerable	1, 2, 3	1	697.53	0.25	0.62 n.s.
		4	0	-	-	-
		5, 6, 7	1	669.82	7.72	0.01**

* weakly significant; ** significant; *** highly significant; n.s. not significant

The results of the univariate analysis considering all test sessions (times of the day) show 7 highly significant differences, 2 significant differences, 1 weakly significant difference, and 7 non-significant differences. To isolate the main effects detected between variables (i.e., test sessions corresponding to different times of the day), contrasts were made using pairwise comparisons (Bird & Hadzi-Pavlovic, 2014). In the inferential analysis, the effect size was again calculated by the Cohen's d coefficient and considered small, moderate and large, respectively for d≥ 0.41, 1.15 and 2.70 (Ferguson, 2009). The effect size was obtained from the estimated marginal means and pooled standard deviation values adjusted for the effect of glare source size and position by the MLM analysis, according to the following formula (Equation 4-25) (Fields, 2013; Rosenthal & DiMatteo, 2001):

$$Cohen's\ d = \frac{\Delta M}{\sigma pooled} \qquad \text{(Equation 4-25)}$$

Where ΔM is the difference between the estimated marginal means, and σ_{pooled} is the pooled standard deviation.

Table 4-17 presents the results of the pairwise comparisons, displaying the groups of the independent variable under examination (chronotype), the sample distribution (N) across the independent variable (x_0 and x_1 respectively corresponding to the test sessions considered in the comparison), the difference between the estimated marginal means adjusted for the effect of glare size and position (ΔM) and its statistical significance (*p*-value, calculated using a two-tailed hypothesis with Bonferroni correction), the standard error, the lower (CI_L) and upper (CI_U) 95% confidence intervals for the mean difference, and the effect size (d).

Since no votes for the GSV criteria of Just Uncomfortable/Intolerable, under the user-set shading setting, were given by test subjects who had self-assessed their chronotype as Type 4 and Type 5, 6, 7 in the morning, Type 4 at midday, and Type 1, 2, 3 in the afternoon, the pairwise comparisons for these test sessions have not been reported.

Table 4-17 Pairwise comparisons and effect sizes for groupings of chronotype

Shading	GSV	Chronotype	Test Sessions	$N_{(x0,\ x1)}$	ΔM^{NHST}	Std. Error	CI_L	CI_U	Effect Size
Default Shading	Just (Im)Perceptible	1, 2, 3	Morn. vs. Mid.	14, 16	0.03 n.s.	0.04	-0.06	0.12	0.74
			Morn. vs. Aft.	14, 9	0.06 n.s.	0.06	-0.06	0.18	1.22
			Mid. vs. Aft.	16, 9	0.03 n.s.	0.05	-0.06	0.12	1.00
		4	Morn. vs. Mid.	10, 9	-0.06**	0.02	-0.10	-0.02	-2.45
			Morn. vs. Aft.	10, 11	-0.11 n.s.	0.09	-0.29	0.08	-1.50
			Mid. vs. Aft.	9, 11	-0.04 n.s.	0.09	-0.22	0.14	-0.60
		5, 6, 7	Morn. vs. Mid.	25, 12	-0.03 n.s.	0.02	-0.06	0.01	-2.45
			Morn. vs. Aft.	25, 25	-0.22***	0.04	-0.30	-0.13	-7.33
			Mid. vs. Aft.	12, 25	-0.19***	0.04	-0.28	-0.10	-6.33
	Just Noticeable	1, 2, 3	Morn. vs. Mid.	15, 22	-0.01 n.s.	0.02	-0.05	0.02	-0.47
			Morn. vs. Aft.	15, 20	-0.21***	0.05	-0.31	-0.11	-4.92
			Mid. vs. Aft.	22, 20	-0.19***	0.05	-0.29	-0.10	-5.17
		4	Morn. vs. Mid.	7, 6	-0.03 n.s.	0.03	-0.08	0.02	-1.63
			Morn. vs. Aft.	7, 6	-0.17*	0.05	-0.27	-0.82	-5.20
			Mid. vs. Aft.	6, 6	-0.14*	0.05	-0.23	-0.05	-4.04
		5, 6, 7	Morn. vs. Mid.	25, 33	-0.03*	0.01	-0.05	-0.01	-2.45
			Morn. vs. Aft.	25, 18	-0.10***	0.03	-0.16	-0.05	-4.26
			Mid. vs. Aft.	33, 18	-0.08**	0.03	-0.13	-0.02	-2.98
	Just Uncomfortable /Intolerable	1, 2, 3	Morn. vs. Mid.	13, 4	0.40 n.s.	0.30	-0.20	1.00	1.37
			Morn. vs. Aft.	13, 13	-0.17***	0.05	-0.27	-0.07	-4.19
			Mid. vs. Aft.	4, 13	-0.57 n.s.	0.31	-0.27	0.88	-0.68
		4	Morn. vs. Mid.	4, 6	0.01 n.s.	0.05	-0.09	0.10	0.00
			Morn. vs. Aft.	4, 4	-0.17*	0.07	-0.31	-0.02	-2.94
			Mid. vs. Aft.	6, 4	-0.18*	0.08	-0.33	-0.02	-2.67
		5, 6, 7	Morn. vs. Mid.	7, 12	-0.02 n.s.	0.02	-0.06	0.01	-1.63
			Morn. vs. Aft.	7, 14	-0.14**	0.05	-0.24	-0.05	-3.81
			Mid. vs. Aft.	12, 14	-0.12**	0.05	-0.21	-0.03	-3.27

User-set Shading	Just (Im)Perceptible	1, 2, 3	Morn. vs. Mid.	27, 22	0.04*	0.02	0.00	0.07	1.89
			Morn. vs. Aft.	27, 32	0.01 n.s.	0.02	-0.03	0.04	0.47
			Mid. vs. Aft.	22, 32	-0.03 n.s.	0.02	-0.06	0.00	-2.45
		4	Morn. vs. Mid.	14, 14	0.01 n.s.	0.02	-0.03	0.06	0.14
			Morn. vs. Aft.	14, 14	0.01 n.s.	0.03	-0.04	0.07	0.01
			Mid. vs. Aft.	14, 14	0.00 n.s.	0.02	-0.04	0.04	0.00
		5, 6, 7	Morn. vs. Mid.	31, 36	-0.06***	0.01	-0.09	-0.04	-4.90
			Morn. vs. Aft.	31, 38	-0.05**	0.02	-0.08	-0.02	-4.08
			Mid. vs. Aft.	36, 38	0.01 n.s.	0.01	-0.01	0.04	0.82
	Just Noticeable	1, 2, 3	Morn. vs. Mid.	11, 16	-0.01 n.s.	0.03	-0.07	0.05	-0.32
			Morn. vs. Aft.	11, 9	-0.07 n.s.	0.05	-0.18	0.03	-1.70
			Mid. vs. Aft.	16, 9	-0.07 n.s.	0.05	-0.15	0.02	-2.00
		4	Morn. vs. Mid.	6, 6	-0.04 n.s.	0.04	-0.11	0.04	-1.09
			Morn. vs. Aft.	6, 2	-0.10 n.s.	0.27	-0.63	0.42	-0.04
			Mid. vs. Aft.	6, 2	-0.07 n.s.	0.27	-0.59	0.46	-0.31
		5, 6, 7	Morn. vs. Mid.	26, 16	-0.08***	0.02	-0.10	-0.05	-5.72
			Morn. vs. Aft.	26, 15	-0.06**	0.02	-0.10	-0.01	-2.89
			Mid. vs. Aft.	16, 15	0.02 n.s.	0.02	-0.02	0.06	1.15
	Just Uncomfortable /Intolerable	1, 2, 3	Morn. vs. Mid.	4, 4	-0.02 n.s.	0.03	-0.07	0.05	-0.41
			Morn. vs. Aft.	-	-	-	-	-	-
			Mid. vs. Aft.	-	-	-	-	-	-
		4	Morn. vs. Mid.	-	-	-	-	-	-
			Morn. vs. Aft.	-	-	-	-	-	-
			Mid. vs. Aft.	-	-	-	-	-	-
		5, 6, 7	Morn. vs. Mid.	-	-	-	-	-	-
			Morn. vs. Aft.	-	-	-	-	-	-
			Mid. vs. Aft.	5, 4	0.07**	0.02	0.02	0.12	2.86

Ordered Groups: Early Chronotype= 1, 2, 3; Neither Early nor Late= 4; Late Chronotype= 5, 6, 7
Bonferroni Correction: * 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

Under the default shading setting, an analysis of the descriptive statistics (ΔM) reveals a general tendency for negative differences over the independent variable (time of day) to be detected. This signals a trend for an increase in DGP as the day progresses. The differences detected in the pairwise comparisons under the default shading setting are highly significant in 6 cases, significant in 4 cases, and weakly significant in 5 out of 27 cases. The differences across the independent variable have generally substantive effect sizes, ranging from large ($d \geq 2.70$) in 11 cases, to moderate ($1.15 \leq d < 2.70$) in 8 cases, and small ($0.4 \leq d < 1.15$) in 5 cases. The inferential tests, thus, confirm the postulated hypothesis of a tendency for the DGP to increase as the day progresses under the default shading setting for all criteria of reported glare sensation. This tendency appears to be moderated by differences in chronotype (i.e., the variation of tolerance to discomfort glare seems to differ between chronotype groups).

Figure 4-22 can support the visualisation of the effect of chronotype on the variation of glare response along the day, under the default shading setting. The sign of the effect is retained in this figure so as to account for the directionality of the influence. A negative effect size signals greater tolerance to discomfort glare as the day progresses (i.e., larger DGP at later test sessions corresponding to an equal vote of glare sensation), while a positive effect size signals a decreased tolerance. The figure plots, for each group of chronotypes – 'Early Chronotypes' (Type 1, 2, 3), 'Neither Early nor Late' (Type 4), and 'Late Chronotypes' (Type 4, 5, 6) – the effect sizes (d) of the pairwise comparisons between the Morning vs. Midday (top), Midday vs. Afternoon (middle), and Morning vs. Afternoon (bottom).

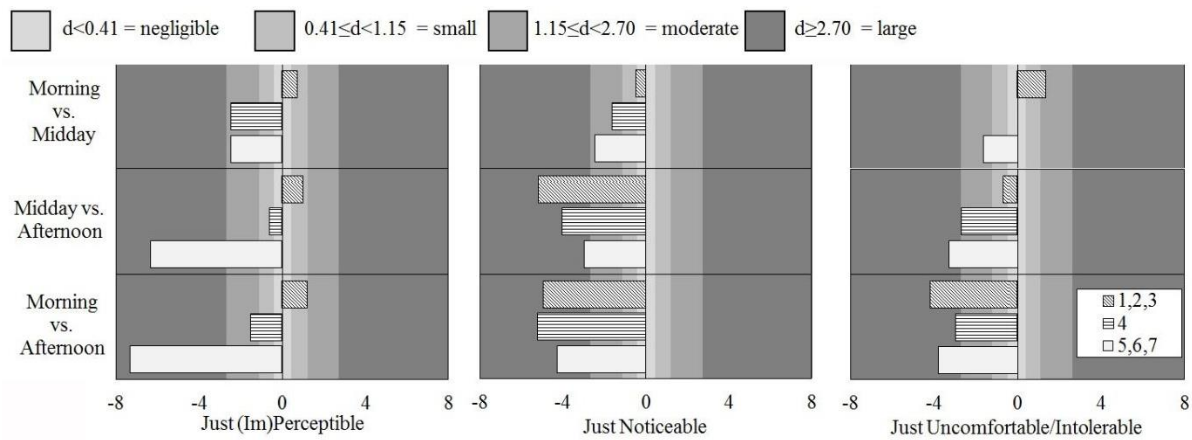


Figure 4-22 Effect sizes (d) detected in the pairwise comparisons (see Table 4-17) between test sessions (y-axis), for each GSV criterion (x-axis) and grouping of chronotype under the default shading setting

Under the default shading setting, consideration of early chronotypes (Type 1, 2, 3) at the GSV of 'Just (Im)Perceptible' indicates that, for each pairwise comparison, the effect sizes are all positive and substantive – albeit with no statistical significance – this signalling a trend for lower tolerance to discomfort glare as the day progresses. However, for later chronotype groups (Type 4, and Types 5, 6, 7), at the same GSV, the effect sizes for the pairwise comparisons are all negative, hence suggesting an increased tolerance to discomfort glare along the day. Moreover, the effect sizes observed for the 'latest' chronotype category (Type 5, 6, 7) are more substantive (i.e., they are larger in magnitude than Type 4 chronotypes), this indicating that the effect of time of day on discomfort glare appears to amplify for this group.

At the 'Just Noticeable' reported level of glare sensation, the figure shows consistent negative effect sizes (and ΔMs) across all groups of chronotype, and for all pairwise comparisons, with a tendency for effect sizes to amplify as the time between test sessions increases. This suggests a prevailing directionality of the effect, and an increased tolerance to glare along the day for all chronotype groups. At this criterion of glare sensation, chronotype appears to have the same moderating effect over time of day under all groups under consideration.

At the highest GSV of 'Just Uncomfortable/Intolerable', the pairwise comparisons show a prevailing trend to display negative effect sizes, suggesting again an increased tolerance to glare as the day progresses. However, the comparison 'Morning vs. Midday' for early chronotypes (Types 1, 2, 3) shows an opposite trend ($d= +1.37$), suggesting a lower tolerance to discomfort glare at the later time of day. Similarly, the comparison between these test sessions for Type 4 chronotypes results in differences of no practical significance ($d= 0.00$).

Under the user-set shading setting, analysis of the descriptive statistics of Table 4-17 does not reveal any consistent sign of the differences in marginal mean (ΔM), hence providing no evidence to substantiate a univocal directionality of the effect. The differences detected in the pairwise comparisons under the user-set shading are highly significant in 2 cases, significant in 3 cases, weakly significant in 1 case, and not significant in 14 cases. The differences have effect sizes ranging from large ($d \geq 2.70$) in 5 cases, moderate ($1.15 \leq d < 2.70$) in 5 cases, small ($0.41 \leq d < 1.15$) in 4 cases, and negligible ($d < 0.41$) in 6 cases.

Figure 4-23 plots, for all groups of chronotype and reported levels of glare sensation, the effect size (d) extracted from the pairwise comparisons under the user-set shading setting.

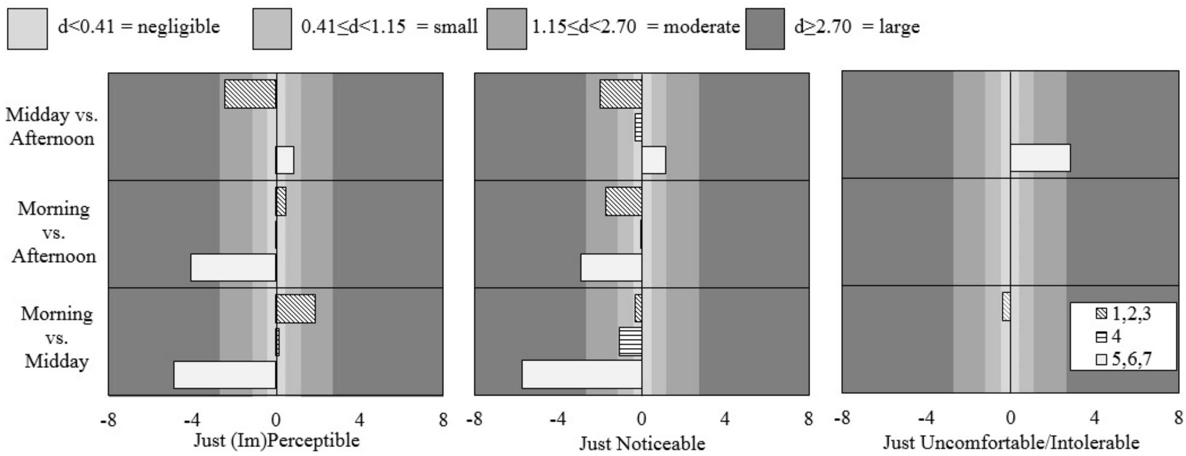


Figure 4-23 Effect sizes (d) detected in the pairwise comparisons (see Table 4-17) between test sessions (y-axis), for each GSV criterion (x-axis) and grouping of chronotype under the user-set shading setting

Inspection of Figure 4-23under the user-set shading across all GSV criteria does not lead to any consistent interpretation of trends within the data. This supports the findings from the original fixed-effects MLM, which indicated that – once the blinds were adjusted to the test subjects' own preference – the effect of time of day on glare response was confounded.

In essence, this analysis provides some statistically and practically significant indication of an influence of chronotype on the variation of tolerance to discomfort glare as the day progresses at different levels of glare sensation. In fact, under the default shading setting, inspection of the effect sizes extracted from the pairwise comparisons between test sessions seems to confirm the observation that chronotype might have a moderating influence over the experimental hypothesis. In particular, consistent increases of tolerance to discomfort glare as the day progresses could be observed for late chronotypes at all reported levels of GSV. This trend was not apparent for subjects self-assessed as early chronotypes. However, no clear and univocal interpretation of this effect can be made based on the data available. Also, findings under the user-set shading offer no evidence to consistently support the postulation of an influence of chronotype over variation of glare response at different times of the day. In interpreting the results of this analysis, it should be considered that the questions featured in the MCTQ (Munich Chronotype Questionnaire) result in a subjective determination of chronotype. Conversely, the MCTQ self-assessments do not provide a direct measure of any physiological biomarker (i.e., core body temperature, melatonin, cortisol, etc.) corresponding to a relationship between the external and the internal time, i.e. the circadian phases of entrainment that characterise the chronotype (Roenneberg *et al.*, 2007; 2003). In addition, the analysis reported here was limited by a relatively small sample size, which was distributed unevenly across chronotype categories. Consequently, in the inferential tests, the addition of chronotype to the MLM caused the standard errors associated with the parameter estimates (i.e., mean and standard deviation) to inflate. Finally, the sample under examination (postgraduate students) cannot be considered to be representative of a general population.

From a review of the literature, it is also important to highlight that chronotype – as glare sensation – is characterised by large individual differences. Among the variables known to influence chronotype are social cues, work schedules, socio-cultural habits, meal times, age, gender, etc. (Levandovski *et al.*, 2013). Genetic and environmental factors have also been found to contribute to endogenous individual differences between morning- and evening-types (Vink *et al.*, 2002). In this context, a study by Roenneberg *et al.* (2007) – from 55,000 subjects who completed the MCTQ – found that sleep and wake times (i.e., sleep duration) showed a near-Gaussian (normal) distribution in the population under examination. This distribution in self-assessments of chronotype was attributed to the well-grounded variation of individuals' phases of entrainment that has been detected within experimental studies.

Based on the laboratory results presented in Section 3.3.3.2 – whereas test subjects self-assessed as earlier chronotypes were associated with a higher tolerance to source luminance – further experimental work, using a larger sample size (and, possibly, consideration of further variables), would be required to more deeply examine the relationship between chronotype and glare sensation as the day progresses in the presence of daylight from a window (Duffy *et al.*, 1999). This further work is, however, beyond the scope of this doctoral research.

4.3. View Importance and Glare Sensation

4.3.1. Introduction

The experiment described in Section 4.1 provided statistically and practically significant evidence of an increasing tolerance to discomfort glare from daylight as the day progresses when the Venetian blinds were set at their default 'cut-off' position. Conversely, when the shading was adjusted to the test subject's own preference, the effect of time of day on glare sensation could not be detected. During the tests, temporal variables were also measured to analyse their influence on subjective levels of visual discomfort. The further analysis described in Section 4.2 showed that, when the variance associated with these temporal variables was controlled – hence increasing the sensitivity of the inferential tests (Field, 2013) – evidence of a direct influence of time of day on glare response could be observed also under the user-set shading setting. In addition, the results revealed that the magnitude of the temporal influences detected (effect size) was, in general, smaller under the user-set shading. Therefore, it may be plausible to hypothesise that, under the two different shading settings, test subjects may have been exposed to uncontrolled variable(s) or condition(s) that could have masked or confounded the effect of time of day on glare response.

From a review of the literature, it is evident that the possibility to control the shading system (Christoffersen & Wienold, 2008) can have an influence on the reported level of discomfort glare. Several studies have suggested that observers prefer to control the blind system to ensure visual contact with the outside. Among these, Wienold (2009) stated that subjects may be willing to accept visually uncomfortable situations if they are given control over the blinds. Similarly, a recent study by Karlsen *et al.* (2015) using different shading settings – one with 'closed slats' and one with an angle set at a minimum of 15^{o} to prevent direct sunlight penetration but still allow a view to the outside – revealed that the 15^{o} position was significantly more popular among observers. Comments by the subjects suggested that the importance of the view influenced their choice of preferred shading setting. Other studies have investigated in detail the relationship between the content of a view and the personal evaluation of visual discomfort (Tuaycharoen & Tregenza, 2005: 2007; Tuaycharoen, 2011).

In the experimental study previously described, given the two different shading settings used ('default' and 'user-set'), it may be postulated that the view to the outside could have potentially masked or confounded the effect of time of day on glare response when users were given control over the blinds. Conversely, when the Venetian blinds slats were set at their default 'cut-off' angle (between 5-10^{o}), this may have controlled the influence of view, and the effect of time of day on the subjective evaluation of glare sensation could be detected. On the basis of these results, a further preliminary analysis was set to investigate whether an effect of view importance could be detected on the glare response provided by test subjects.

4.3.2. Method

4.3.2.1. Experimental Procedure

The data from the previously described user-assessment experiment (see Section 4.1.2.1) were utilised in this analysis. During the sessions, test subjects were required to perform two series of three visual tasks: landolt ring, letter searching, and typing. Each series was performed under a different shading setting (one 'default' and one 'user-set'), giving a total of six visual tasks (Wienold, 2009). All visual tasks were presented on a VDU. Following the completion of each task, subjects were asked to indicate their perceived magnitude of glare sensation given by the daylight coming from the window utilising a GSV scale displayed on the screen, and to make judgements of 'view importance' utilising a 4-point scale corresponding to subjective levels of importance: 'Undesirable'; 'Indifferent'; 'Somewhat Important'; and, 'Very Important' (Wienold, 2009; Velds, 2000: 2002; Wienold &

Christoffersen, 2006) (Figure 4-24). At this point, the experimenter collected a series of seven LDRI with varying exposure values, as well as a single vertical illuminance measurement.

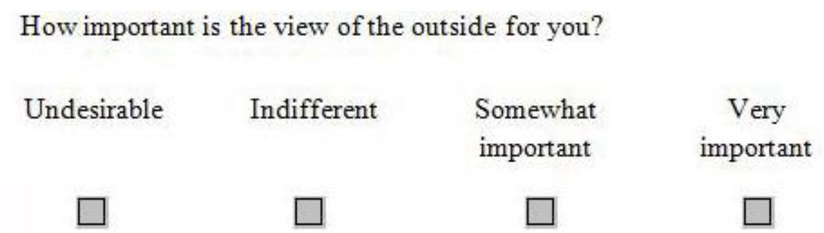


Figure 4-24 4-point scale for the assessment of view importance

4.3.3. Statistical Analysis

To analyse the data, a MLM analysis with 'fixed-effects' was performed to compare the DGP values – which, for consistency with previous investigations, was used again as the primary evaluation parameter – for the reported levels of view importance given by test subjects corresponding to the shading setting under which the glare assessments were provided. Within the MLM model, the fixed effects specified were:

- View importance;
- Shading setting;
- Glare Sensation Votes (GSV).

According to the literature, the level of 'interest' in the view can trigger greater tolerance to discomfort glare (Tuaycharoen & Tregenza, 2005: 2007; Tuaycharoen, 2011). However, evidence-based research has not yet established whether this phenomenon persists the more familiar an observer becomes with the view they are exposed to. As a result, in the experiment, test subjects were required to provide votes of importance of view after each glare assessment under both the default and the user-set shading settings.

4.3.4. Results and Discussion

Figure 4-25 plots, on the y-axis, the DGP (with 95% confidence intervals) calculated from the images captured by the CCD camera and evaluated using *Evalglare*. Unlike the previous analyses, comparisons of reported glare sensation did not include consideration of time of day. Therefore, the estimated marginal parameters (means) to each independent variable and conditional group have not been adjusted to account for the influence of confounding variables. On the x-axis, the figure presents the shading setting under which GSV assessments were provided by test subjects, as well as the reported rating of view importance.

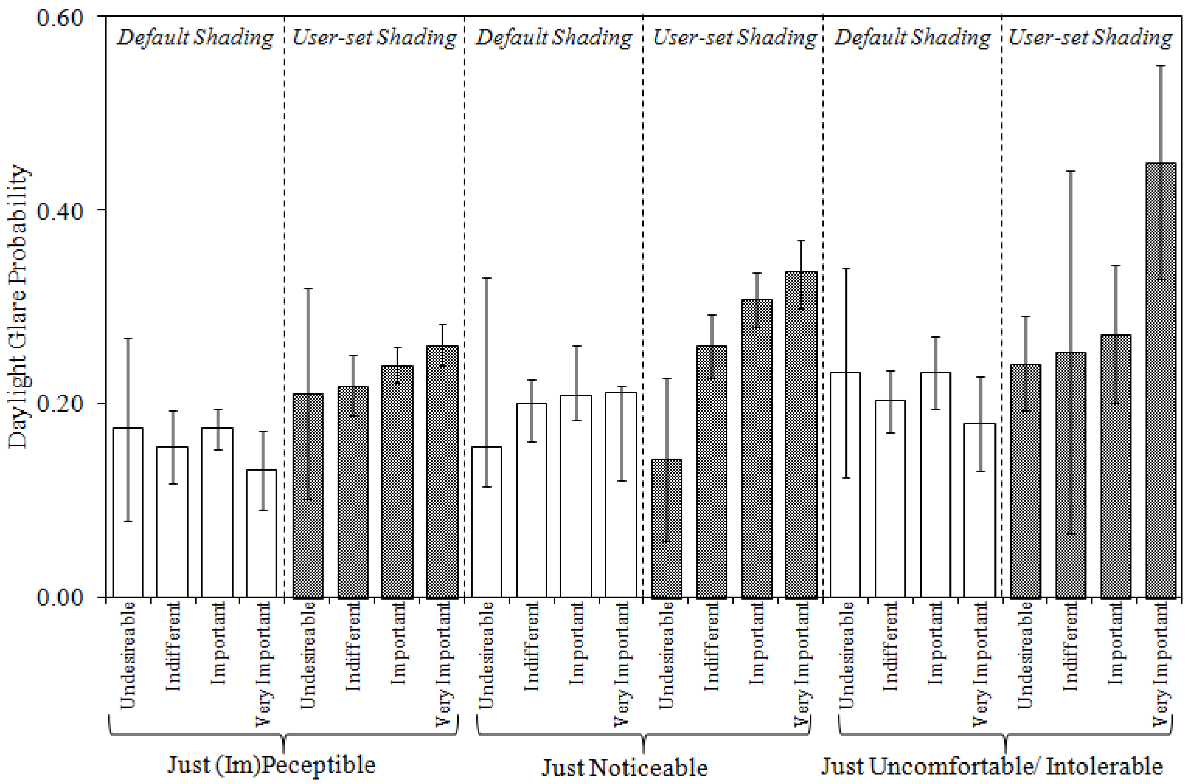


Figure 4-25 Mean plots with 95% confidence intervals of adjusted DGP values (y-axis) for each GSV criterion, rating of view importance, and shading setting (x-axis)

Graphical inspection of the boxplots suggests that, as the reported level of view importance increases, no consistent trend appears within the calculated DGP under the default shading setting across all criteria of glare sensation. Since the DGP indicates the probability that an observer may be ‘disturbed’ by the physical and photometric parameters within their field of view (Wienold & Christoffersen, 2006), this shows that no influence of view importance on the perceived level of glare sensation can be postulated under the default shading setting. Conversely, under the user-set shading setting, the boxplots indicate that, as the reported view importance increases, there is an apparent trend for the central tendencies to correspond to higher levels of DGP. This suggests that, when test subjects reported greater levels of view importance, they were able to tolerate a higher combination of physical and photometric parameters within their field of view. This tendency is consistently observed for all the GSV criteria under consideration, and leads to hypothesise, when subjects could set the Venetian blinds to their preferred position, an effect of the importance of view on glare sensation.

For the DGP, the MLM analysis for the effect of view importance (i.e., ‘Undesirable’ vs. ‘Indifferent’ vs. ‘Important’ vs. ‘Very Important’) detected no significant difference: $F(3, 720.00)= 0.67$, $p>0.05$. This result, therefore, provides no statistically significant evidence to postulate that the mean DGP values calculated at the four ratings of view importance reported by test subjects were different from each other. The estimates of covariance parameters show a highly significant difference (Wald Z= 18.97, $p\leq 0.001$), giving evidence to reject the null hypothesis, i.e. that the variances associated with each effect are equal to zero. This indicates that the fixed-effects included in the model are not sufficient to explain the spread in glare responses provided by test subjects, and that there may be other variables that influence the effect of view importance beyond the fixed-effects specified in the model (Seltman, 2014).

A review of the literature suggests a variety of candidate variables that might potentially cause the unexplained variability within the data such as light distribution, flow of light and directionality within the field of view, etc. (Cuttle, 1971). In a test room with direct access to daylight, these variables are continuously changing and could possibly influence visual perception and glare sensation in terms of physical appearance of objects within the view, perception of depth, etc. Analysis of these variables, however, would require further experimentation. Table 4-18 displays the results of the univariate tests, showing 1 highly significant, 1 weakly significant, and 4 non-significant differences.

Table 4-18 Univariate tests comparing DGP values for each GSV criterion and shading setting based on reported levels of view importance

GSV	Shading Setting	Numerator df	Denominator df	***F***	***p*-value**
Just (Im)Perceptible	Default Shading	3	720.00	1.17	0.32 n.s.
	User-set Shading	3	720.00	1.83	0.14 n.s.
Just Noticeable	Default Shading	3	720.00	0.61	0.61 n.s.
	User-set Shading	3	720.00	7.43	0.00***
Just Uncomfortable/ Intolerable	Default Shading	3	720.00	1.05	0.37 n.s.
	User-set Shading	3	720.00	2.72	0.04*

* weakly significant; ** significant; *** highly significant; n.s. not significant

It is relevant to note that statistical significance for the perceived level of view importance was only detected once subjects had adjusted the blinds to the user-set shading setting. The results of these tests suggest that the presence of the blinds at their default 'cut-off' angle (5-10^{o}) in the MLM model may be sufficient to partially control for the effect of view on reported glare sensation. In this context in fact, Wilson (2007) – commenting on the work by Tuaycharoen & Tregenza (2007) – questioned whether a partial or fractured view would influence the perception of glare in the same way as an unobstructed view would.

To isolate the main effects detected, contrasts were made using pairwise comparisons (Bird & Hadzi-Pavolic, 2014), whereby all permutations between independent variables were compared against each other. Preliminary exploratory analysis of descriptive statistics and graphical displays of data revealed no consistent directionality of observed differences; hence, a non-directional (two-tailed) hypothesis was adopted (Hauschke & Steinijans, 1996; Ruxton & Neuhauser, 2010). In consideration of the experiment-wise error rate caused by the significance level inflating across multiple pairwise comparisons based on the same hypothesis (Cabin & Mitchell, 2000) – which was calculated as $1\text{-}(0.95)^n= 0.26$ (thus risking a 26% probability of making at least one Type I error), where n= 6, i.e. the number of comparisons made – Bonferroni corrections were applied. Table 4-19 to Table 4-21 present the shading settings, the ordered categories of view importance, the sample size (N) of independent groups (x_0 and x_1 corresponding to the categories considered in the comparisons), the mean difference between independent groups (ΔM) and its associated statistical significance (*p*-value), the standard error, the degrees of freedom (df), the lower (CI_L) and upper (CI_U) confidence intervals for the estimated mean difference, and the effect size (d).

Table 4-19 Pairwise comparisons and effect sizes for 'Just (Im)Perceptible'

Shading	View	$N_{(x0, x1)}$	ΔM^{NHST}	Std. Error	df	CI_L	CI_U	Effect Size (d)
Default Shading	2 vs. 1	25, 4	-0.02 n.s.	0.05	720.00	-0.12	0.08	-0.09
	3 vs. 1	81, 4	0.00 n.s.	0.05	720.00	-0.10	0.10	0.00
	4 vs. 1	21, 4	-0.04 n.s.	0.05	720.00	-0.15	0.06	-0.21
	3 vs. 2	81, 25	0.02 n.s.	0.02	720.00	-0.02	0.06	0.09
	4 vs. 2	21, 25	-0.02 n.s.	0.03	720.00	-0.08	0.03	-0.12
	4 vs. 3	21, 81	-0.04 n.s.	0.02	720.00	-0.09	0.00	-0.21
User-set Shading	2 vs. 1	40, 6	0.01 n.s.	0.06	720.00	-0.10	0.12	0.07
	3 vs. 1	86, 6	0.03 n.s.	0.06	720.00	-0.08	0.14	0.26
	4 vs. 1	20, 6	0.05 n.s.	0.06	720.00	-0.06	0.16	0.49
	3 vs. 2	86, 40	0.02 n.s.	0.02	720.00	-0.01	0.06	0.23
	4 vs. 2	20, 40	0.04*	0.02	720.00	0.00	0.08	0.53
	4 vs. 3	20, 86	0.02 n.s.	0.01	720.00	-0.01	0.05	0.21

Ordered categories: Undesirable= 1, Indifferent= 2, Important= 3, Very Important= 4
With Bonferroni Correction: * 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

Table 4-20 Pairwise comparisons and effect sizes for 'Just Noticeable'

Shading	View	$N_{(x0, x1)}$	ΔM^{NHST}	Std. Error	df	CI_L	CI_U	Effect Size (d)
Default Shading	2 vs. 1	37, 3	0.05 n.s.	0.04	720.00	-0.04	0.13	0.56
	3 vs. 1	112, 3	0.05 n.s.	0.04	720.00	-0.03	0.13	0.59
	4 vs. 1	76, 3	0.06 n.s.	0.04	720.00	-0.03	0.14	0.41
	3 vs. 2	112, 37	0.01 n.s.	0.02	720.00	-0.03	0.04	0.07
	4 vs. 2	76, 37	0.01 n.s.	0.03	720.00	-0.04	0.06	0.07
	4 vs. 3	76, 112	0.01 n.s.	0.02	720.00	-0.04	0.05	0.03
User-set Shading	2 vs. 1	31, 5	0.12**	0.05	720.00	0.03	0.21	1.23
	3 vs. 1	43. 5	0.16***	0.05	720.00	0.08	0.25	1.70
	4 vs. 1	28, 5	0.19***	0.05	720.00	0.10	0.28	2.04
	3 vs. 2	43, 31	0.05*	0.02	720.00	0.00	0.09	0.50
	4 vs. 2	28, 31	0.08**	0.03	720.00	0.03	0.12	0.82
	4 vs. 3	28, 43	0.03 n.s.	0.02	720.00	-0.02	0.07	0.31

Ordered categories: Undesirable= 1, Indifferent= 2, Important= 3, Very Important= 4
With Bonferroni Correction: * 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

Table 4-21 Pairwise comparisons and effect sizes for 'Just Uncomfortable/Intolerable'

Shading	View	$N_{(x0, x1)}$	ΔM^{NHST}	Std. Error	df	CI_L	CI_U	Effect Size (d)
Default Shading	2 vs. 1	35, 3	-0.03 n.s.	0.06	720.00	-0.14	0.08	-0.32
	3 vs. 1	24, 3	-0.01 n.s.	0.06	720.00	-0.12	0.11	-0.01
	4 vs. 1	15, 3	-0.05 n.s.	0.06	720.00	-0.17	0.07	-0.73
	3 vs. 2	24, 35	0.03 n.s.	0.03	720.00	-0.02	0.08	0.39
	4 vs. 2	15, 35	-0.02 n.s.	0.03	720.00	-0.08	0.03	-0.33
	4 vs. 3	15, 24	-0.05 n.s.	0.03	720.00	-0.11	0.01	-1.22
User-set Shading	2 vs. 1	16, 39	0.01 n.s.	0.10	720.00	-0.18	0.21	0.07
	3 vs. 1	42, 39	0.03 n.s.	0.04	720.00	-0.06	0.12	0.09
	4 vs. 1	5, 39	0.21**	0.07	720.00	0.06	0.35	0.95
	3 vs. 2	42, 16	0.02 n.s.	0.10	720.00	-0.18	0.22	0.06
	4 vs. 2	5, 16	0.19 n.s.	0.12	720.00	-0.04	0.42	1.17
	4 vs. 3	5, 42	0.17*	0.08	720.00	0.02	0.32	0.51

Ordered categories: Undesirable= 1, Indifferent= 2, Important= 3, Very Important= 4
With Bonferroni Correction: * 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

The pairwise comparisons provide evidence that the differences between DGP values under the default shading setting are not statistically significant in any of the 18 cases considered. For the user-set shading setting, the differences detected are highly significant in 2 cases, significant in 3 cases, weakly significant in 3 cases, and not significant in 10 cases out of 18. The differences detected for the user-set shading have a generally substantive magnitude, ranging between moderate ($1.15 \leq d < 2.70$) in 5 cases and small ($0.41 \leq d < 1.15$) in 10 cases.

Figure 4-26 plots, for each criterion of glare sensation, the effect sizes (d) extracted from the pairwise comparisons of Table 4-19 to Table 4-21, together with the interpretation of the statistical significance associated with the mean difference between independent groups. The y-axis presents the shading settings – default (top) and user-set shading (bottom) – along with the ordered categories of view importance reported by test subjects: 'Undesirable'= 0, 'Indifferent'= 1, 'Important'= 2, and 'Very Important'= 4.

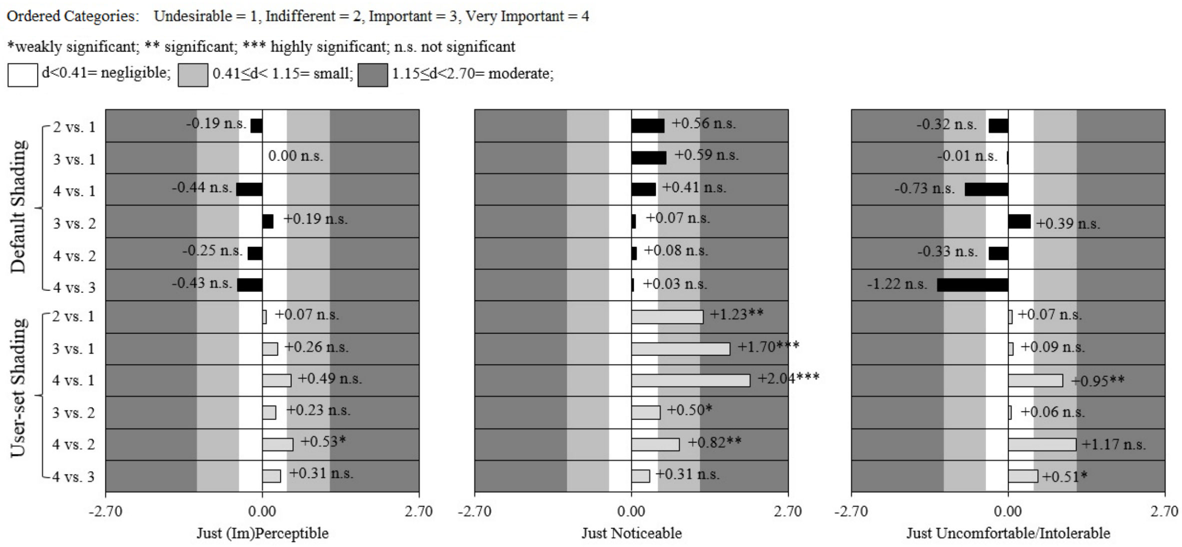


Figure 4-26 Effect sizes (d) detected in the pairwise comparisons (see Tables 4-19 to 4-21) between ordered categories of view importance (y-axis), for each GSV criterion (x-axis) under the default (top) and the user-set (bottom) shading settings

Under the default shading setting, the effect of view importance on reported levels of glare sensation appears to be confounded by the 'cut-off' position of the Venetian blinds. In fact, the pairwise comparisons found no statistically significant differences across the 18 contrasts performed. The presence of practical significance was detected in 7 cases, corresponding to small effect sizes, although no clear trend or directionality in the data could be observed.

For the user-set shading setting, the pairwise comparisons detected both statistically significant and practically relevant differences in calculated DGP values. Analysis of the differences detected reveals that the sign of the effect sizes is consistently positive, signalling higher values of calculated DGP corresponding to greater reported levels of view importance. This suggests, for each criterion of glare sensation, an increased tolerance to discomfort glare when test subjects gave a higher rating to the importance of the view. The findings also show a tendency for a larger influence of view on DGP – as indicated by greater effect sizes – when the difference between ordered categories of view importance increases.

In essence, this analysis provides some statically significant and practically relevant evidence of an influence of view importance on glare response, once test subjects were allowed control over the shading system. Conversely, when the Venetian blinds were set at a 'cut-off' angle, the data offer no support to postulate an effect of view importance on reported GSVs.

In accordance with the literature (Tuaycharoen & Tregenza, 2005; 2007; Tuaycharoen, 2011), the results of this analysis show some evidence that, under the user-set shading, the presence of a view could have masked the effect of time of the day on glare sensation. However, the literature has also postulated that there may be a strong relationship between lighting control (i.e., dimmers or shading systems) and visual response (Boyce, 2015). In fact, research has suggested that individual control (or perception thereof) over the luminous environment can influence task performance, alertness, mood, etc. (Boyce *et al.*, 1999; 2006), although preferred conditions might vary widely due to individual differences (Veitch & Newsham, 2000). In this context, it may be reasonable to postulate that the perceived level of control may have also confounded the effect of view importance under the user-set shading setting.

Nevertheless, given the data collected in this study (related to the achievement of its specific objective), it is not possible to perform a robust inferential analysis to also explore the influence of individual control on reported glare sensation. Further experimental work would be required to examine this relationship, and more deeply investigate temporal variation of glare response under different shading conditions. As for the case of chronotype (Section 4.2.4.1), however, this further work is beyond the scope of this doctoral research.

Chapter 5

Conclusions

5. Conclusions

5.1. Glare Response and Time of Day

This investigation has provided evidence of a significant and substantive effect of time of the day on the personal evaluation of glare sensation reported by test subjects. This evidence has been collected through statistical analysis of data from experiments that have been conducted, initially, within controlled laboratory conditions under artificial lighting and, then, in a test room with direct access to daylight. The sections below summarise the fundamental steps in which this investigation has been articulated, illustrating the outcomes of each experimental stage, presenting the implications of the findings, discussing the limitations of this study, and suggesting further experimental work that is needed in this area of research.

5.1.1. Time of Day and Discomfort Glare from Artificial Lighting

The initial experimental study (Section 3.1) was designed to test whether a relationship could be detected between time of the day and the vote of glare sensation provided by test subjects. Under controlled laboratory conditions, 30 subjects were exposed to an artificial lighting source at four times of the day. The source luminance was progressively increased and subjects were requested to provide Glare Sensation Votes (GSVs) corresponding to the level of visual discomfort experienced. Glare indices (i.e., the IES-GI and the Unified Glare Rating (UGR)) were calculated for each reported vote of glare sensation, and the results were statistically analysed. The findings showed a tendency towards greater tolerance to luminance increases in artificial lighting at later test sessions. This suggested that subjects became more tolerant to discomfort glare as the day progressed. The effect of interest was found to be directly linked to the independent variable itself; that is, differences in tolerance to glare sensation amplified when considering a larger time period between test sessions. Since the same sequence of sessions was maintained for all subjects, a follow-up study was designed to investigate the influence of learning or experience on the results obtained (Section 3.2). This was done by exploring whether differences in glare sensation along the day could be detected over two consecutive days using the same subject. The results showed that the differences detected between the two days of testing were, for all comparisons, not statistically significant, although this could have resulted from a small sample size. However, calculation of effect sizes indicated that the magnitudes of the variations were mostly negligible or on the borderline of practical relevance. This led to infer that any substantive influence of learning or experience could be safely excluded from the interpretation of the results.

5.1.2. Temporal Variables and Personal Factors in Glare Sensation

When the plots of GSVs were regressed against the times of the day (Section 3.1.3, Figure 3-10 and Figure 3-11), a large scatter was observed in the results, as confirmed by the low coefficient of determination obtained using a linear fit. This suggested that, despite the efforts to control as many variables known to influence glare sensation as possible within the experimental design, there were further factors that might have caused subjective variations of response to luminance increases. To explore the potential sources of these individual differences, several temporal variables and personal factors were also collected during the tests. This enabled analysis of their influence on glare sensation as the day progresses (Section 3.3). The results revealed statistically significant and practically relevant tendencies toward greater tolerance to source luminance from artificial lighting at all times of the day for earlier chronotype test subjects. That is, earlier chronotypes appeared to tolerate higher levels of source luminance for the same reported level of glare sensation across all test sessions. Moreover, consideration of caffeine ingestion identified some statistically significant and practically relevant effects, leading to the postulation that subjects not having ingested caffeine were more tolerant to source luminance across all criteria of glare sensation. Conversely, although some statistically and practically significant differences were detected, no conclusive evidence was found for the effect of fatigue, sky condition, and prior light exposure on glare response along the day. These findings suggested that – in the exploration of the hypothesis that an effect of time of day can be detected under a less controlled experimental setting (i.e., daylight conditions) – temporal variables and personal factors should be measured in conjunction with subjective glare assessments.

5.1.3. Visual Task Difficulty and Temporal Influences in Glare Response

A follow-up study (Section 3.4) further explored the relationship between temporal variables and glare response, introducing the role of visual task difficulty. Under controlled laboratory conditions, 20 subjects were exposed to a constant source luminance at four different times of day and gave glare sensation votes while completing 12 visual tasks of various difficulties. 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. Hence, the findings from this experiment supported the previous conclusion that the length of the time period between test sessions had a direct effect on the increased tolerance to source luminance as the day progresses. The temporal variation of glare response was found to be dependent on the level of visual discomfort experienced and to be influenced by the difficulty in extracting information from the visual stimulus. In fact, direct (e.g., size and contrast of the character) and indirect (e.g., its background) manipulation of the information of interest within the task was found to lead to different effects on the variation of glare response along the day. Coherent with previous findings, when plots of GSV were regressed for each visual task, a large scatter in the results appeared. This suggested that there may be further factors, which also vary with time of the day, influencing individual variations of glare response.

5.1.4. Influence of Temporal Variables on Glare Response

During the laboratory tests performed while participants engaged with visual tasks, self-assessments of several temporal variables (fatigue, food intake, caffeine ingestion, mood, previous daylight exposure, and sky condition) were provided by test subjects together with their glare judgements (see Section 3.4.3.3). A significant and substantive direct relationship was found between the self-reported level of fatigue and glare response, particularly in the morning and during the post-lunch afternoon session. 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 at midday and in the evening. That is, test subjects who declared to have ingested more caffeine were associated with higher levels of glare sensation from a constant source luminance at different times of the day. 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 two variables, the scatter of the data did not allow inferring a clear relationship with subjective glare sensation. Consideration of inferential results from all test sessions led to the hypothesis that some temporal variables may interact (or be correlated) with each other, and significantly affect the variation of glare response at different times of day. As an example, higher levels of caffeine ingestion were associated with lower levels of self-assessed fatigue. Therefore, it may be difficult to systematically isolate the true effect of individual variables on the subjective evaluation of glare sensation as the day progresses.

5.1.5. Time of Day and Discomfort Glare from Daylight

In consideration of the findings derived from a controlled laboratory setting, an experiment based on the use of a test room, with direct access to daylight and a view to a natural scene, was designed (Section 4.1). Forty test subjects participated to three test sessions, each at a different time of the day, randomised over three consecutive days. During each test session, subjects were asked to perform two series of three visual tasks (landolt ring, letter searching, and typing). Consistent with the findings from the laboratory experiments, a multilevel model analysis revealed that the effect of time of day on the variation of reported glare response was directly linked to the independent variable itself. That is, increases in tolerance to discomfort glare appeared to amplify when considering a larger time period between test sessions. However, this finding was initially detected only when the internal Venetian blinds protecting the window were set at their 'cut-off' angle, i.e. ensuring predominantly diffused daylight conditions inside the test room but still permitting some view to the outside from the position of the test subject. Conversely, when the shading setting was adjusted to the test subject's own preference, the effect of time of day appeared to be confounded. Further investigation performed by controlling for the effects of several temporal variables – that had been isolated from the previous laboratory experiments – indicated that a direct influence of time of day on glare response could be observed also under the user-set shading setting (Section 4.2). However, the magnitude of the temporal influences detected (effect size) was, in general, smaller when the Venetian blinds were adjusted by subjects. On the basis of these findings, it was hypothesised that, under the user-set shading, the presence of other uncontrolled variable(s) or conditions (Section 4.3) may have potentially masked or confounded the magnitude of the temporal influence on the perception of glare reported by test subjects.

5.2. Implications

This research has provided supportive evidence that the subjective evaluation of glare sensation is characterised by large individual differences (Stone & Harker, 1973; Tregenza & Wilson, 2011). The results obtained through systematic analysis have shown that a potential cause of this scatter is the effect of time of the day (and the associated influences of personal factors and temporal variables), as demonstrated by the direct relationship detected between increased tolerance to discomfort glare as the day progresses and the time of the test session when observers were exposed to the (artificial and natural) lighting sources.

Rodriquez & Pattini (2014) have stated that it may not be possible to completely eliminate the scatter commonly associated with personal evaluation of glare sensation, although most of the individual differences may be mitigated if these are systematically identified. From a review of the literature, it is evident that most of the potential causes of variability in glare response – beside the factors embedded within traditional glare formulae – are not necessarily linked to conventional intrinsic characteristics (e.g., age, culture, eye correction, etc.), but that further variables – as analysed in this investigation – need to be taken into account. As proposed within the research hypothesis for this study, this ultimately confirms that there is a need to move beyond the parameters commonly found in glare indices to more accurately evaluate and predict the occurrence and magnitude of discomfort glare.

5.2.1. Statistical Modelling

In order to systematically evaluate the scatter commonly associated with personal evaluations of glare sensation, new grounds had to be broken within this investigation. To respond to this challenge, different statistical methods from those conventionally documented in the literature were adopted to study the sources of variability in subjective glare response.

A review of the literature identified several potential candidate variables that could be linked to variations in reported levels of glare sensation (e.g., age, cultural differences, view, etc.). However, despite considerable statistical evidence, the methodologies adopted in studies from the literature considered mostly fixed-effects approaches. Since random population effects (i.e., the between-subject variance) were generally not addressed in previous studies, it was not feasible to determine whether (or to what extent) fixed parameter estimates (e.g., mean and variance) would change when considering the randomness of the sample population. The graphical displays presented in Figure 1-1 to Figure 1-5, in fact, consistently show that, in previous research, the between-subject variance was considerably larger than the within-subject variance. More specifically, the variability in reported levels of glare sensation associated with each independent condition related to the independent variable (i.e., between-subject variance) was greater than the variability observed across the independent variable (i.e., within-subject variance). A number of factors (i.e., subjective effects) can characterise the variability between test subjects (i.e., inter-individual differences) (Seltman, 2014). Nevertheless, while there are many studies modelling the variability associated with intra-individual differences in fixed-effects approaches, heterogeneous variance models – accommodating time-invariant and time-varying predictors at various levels of the analysis, and for both fixed- and random-effects (i.e., variance components) – are less common (Almeida *et al.*, 2009).

In this investigation, multilevel modelling (MLM) with consideration of fixed- and random-effects has enabled the variances associated with the effect of time of day (i.e., intra-individual variability related to test session-to-test session, or within-subject, variance), and the effects of the temporal variables (i.e., inter-individual variability related to participant-to-participant, or between-subject, variance) to be partitioned from each other (Field, 2011; Peugh, 2010). This has allowed both sources of variability to be considered, facilitating the formulation of inferences related to the effects of experimental interest and to the sampled population from which the data were drawn (Penny & Holmes, 2003).

The results from this study support the idea that more complex statistical models, such as MLM and mixed-effects models, are necessary to fully characterise the scatter in subjective glare evaluation. Yet, as mentioned in the introduction to this study, there might not be one specific general effect that consistently causes the occurrence of discomfort glare (Jay, 1989).

5.3. Limitations of this Study

This investigation has been performed under two different experimental conditions – a controlled laboratory setting and a test room with direct access to daylight – to explore the effect of time of the day (along with other temporal variables and personal factors) on the subjective evaluation of glare sensation. In contextualising the influences detected, the following methodological and experimental limitations should be considered:

1. In order to isolate the effect of experimental interest, participants were requested to provide judgements of glare sensation using subjective rating scales. Fotios (2015) argued that there is still insufficient understanding of discomfort ratings, this potentially increasing the response variance through uncertainty over the criteria anchored to these scales. This suggests that, when forcing a continuous dependent variable (i.e., glare index) into discrete categories associated with subjective levels of glare sensation (i.e., the 4-point scale of Glare Sensation Votes), there is a risk of unintentionally making respondents report a stimulus that does not accurately reflect their perceived evaluation of that stimulus. This could possibly obscure the true nature of the effect under examination. According to the literature, the multiple criterion technique of subjective appraisal should anyway be preferred over a forced-choice dichotomy scale (e.g., yes/no, comfortable/uncomfortable) when evaluating individual differences from perceived glare sensation (Hopkinson, 1950). In fact, such methods may not provide an accurate representation of the variability that exists within the data (Hopkinson, 1950; MacGowan & Clear, 2013). In this context, Boyce (2004) stated that visual comfort and visual discomfort should be considered as independent from each other, rather than a

sensation to be measured over a continuous scale. Specifically, a perception of discomfort may be characterised by feelings of pain, soreness, and numbness that may change over time in relation to physiological factors. Conversely, a perception of comfort could be related to feelings of well-being that may also be linked to aesthetics (Boyce, 2004). However, when the number of possible outcomes becomes too large, this potentially introduces sources of bias through self-interpretation or abstraction caused by similarities in the semantic meaning of categories anchored within the scale (i.e., it may not be easy to discriminate distances between the semantic labels) (Lim, 2011). For this reason, the 4-point GSV scale has been preferred in this investigation over the 9-point multiple criterion technique scale used by other researchers (Iwata *et al.*, 1992a; 1992b; Tokura *et al.*, 1996; Tuaycharoen & Tregenza, 2005; 2007; Velds, 2000; 2002). Similar adaptations have also been used in previous studies (Christoffersen & Wienold, 2006; Rodriquez *et al.*, 2015). To further reduce self-interpretation, semantic labels anchored to the subjective rating scales should be preferably linked to descriptors so as to help subjects' response according to the meaning of each benchmark (Lim, 2011). Accordingly, within this study, time-span descriptors have been linked to each respective category contained within the GSV scale, as originally developed from the work of Osterhaus & Bailey (1992). Nevertheless, based on the literature, most limitations associated with subjective rating scales are difficult to deal with and, in some cases, are even unavoidable (Poulton, 1979). In such circumstances, the best practice is to be aware of the potential pitfalls, how to counterbalance them, and make proper inferences from the findings. More importantly, appropriate measure criteria should be selected so that the data can be handled by statistical inferential tests (Field & Hole, 2013). For this reason, during the laboratory studies, the measured criterion was varied utilising in all cases a 4-point scale that was presented once as a semantic scale and once as a continuous linear visual analogue scale (VAS). Under both conditions, the findings showed the same direct relationship between increased tolerance to source luminance and time of the day.

2. In the interpretation of the temporal influences detected from the laboratory studies performed while subjects engaged with visual tasks (Section 3.3.3 and Section 3.4.3), it should be highlighted that the between-subject tests adopted were characterised by severely uneven sample sizes. These tests are also less likely to detect statistical significance than repeated-measures experiments (e.g., Wilcoxon matched-pairs), as individual differences between participants cannot be completely controlled (Field & Hole, 2013).

3. Finally, it should be considered that, in order to retain high statistical power in the interpretation of the outcomes from the test room experiments (Section 4.1.3.2) at high levels of reported glare sensation, the criteria of 'Just Uncomfortable' and 'Just Intolerable' have been merged into one single category. One of the major limitations when utilising the MLM analysis is that most of the statistical calculations are asymptotic (Maas & Hox, 2004; 2005), i.e., they are based on the assumption of large sample sizes. In interpreting the findings at high levels of reported glare sensation under the user-set shading, caution was therefore adopted since high statistical power was not achieved and, consequently, this could have potentially invalidated the estimation parameters (mean, variance, and standard error).

5.4. Future Work

5.4.1. Glare Sensation and Importance of Views

The findings from the test room studies (Section 4.1.4 and Section 4.2.4) suggested that the potential presence of uncontrolled variable(s) or condition(s) may have masked or confounded the effect of time of the day on glare response depending on the setting of the shading device. In this context, the literature has postulated an influence of view interest on the level of discomfort glare reported by test subjects (Tuaycharoen & Tregenza, 2005; 2007). As part of this research, a preliminary analysis of the influence of the importance of the view (Section 4.3) showed, for all criteria of glare sensation under the user-set shading, a direct relationship between increased tolerance to discomfort glare and higher ratings of perceived view importance. Conversely, under the default shading setting, no conclusive evidence was found of an influence of view importance on the evaluation of glare. The outcomes of this preliminary analysis lead to hypothesise that, under the user-set shading, the presence of an interesting view could have masked the effect of time of the day on glare sensation. However, further investigation would be required to consolidate this inference since these findings cannot be generalised over the time of the day due to statistical limitations (i.e., sample size) and the lack of comparative conditions (i.e., data from test rooms with different views).

5.4.2. Chronotype and Glare Sensation

Under controlled laboratory conditions, a greater tolerance to artificial lighting source luminance across all reported levels of glare sensation was detected for test subjects self-assessed as earlier chronotypes (Section 3.3.3.2). Within the studies performed in a test room with direct access to daylight, statistically and practically significant evidence was found of an influence of chronotype on the variation of tolerance to discomfort glare as the day progresses (Section 4.3). However, although this variation of tolerance was isolated across the various categories of chronotype and levels of glare sensation, no conclusive directionality of the effect could be identified. These observations were uniquely isolated when the Venetian blinds were at their default setting (thus, reinforcing the hypothesis that view importance could have a strong confounding effect). In interpreting the results, some limitations were acknowledged, particularly in terms of the self-reported methods adopted for the assessment of chronotype and the absence of objective physiological measures. Therefore, further investigation would be required to isolate both the main-effect of chronotype on glare sensation (i.e., early chronotypes vs. late chronotypes) and the interactive effects between variables (i.e., the variation of the effect of time of the day associated to differences in chronotype). Furthermore, a review of the literature has revealed that chronotype – like glare sensation – is also characterised by large individual differences and may be influenced by social, genetic, and environmental factors, hence suggesting consideration of these additional variables in future investigations (Levandovski *et al.*, 2013; Vink *et al.*, 2002).

5.4.3. Time Periods

It is important to consider that findings from this investigation are exclusively related to a time period ranging from 09:00 to 18:00 for laboratory studies (under artificial lighting), and from 09:00 to 15:00 for test-room studies (in the presence of daylight). Therefore, there is no evidence from this study to suggest whether variations of tolerance to the sensation of glare might be manifest at other times within the 24-hour period of a full day.

As part of future research, more time periods need to be investigated to fully characterise the relationship between time of day and subjective assessments of glare sensation.

5.4.4. Field Studies

Finally, the findings from this investigation were derived from a laboratory setting with artificial light and from test room conditions with direct access to daylight, whereby several variables that could potentially influence glare response were controlled or masked. The hypothesis that the same effect of time of the day can be detected on the perceived level of glare sensation from daylight in a side-lit occupied space (e.g., an office) with little (or no) control over the environmental settings remains, therefore, conjectural and requires further verification. In this context, future field-based research studies should seek to isolate temporal influences on the subjective evaluation of discomfort glare with possible consideration of time of the day in comfort-based metrics (i.e. the Daylight Glare Probability (DGP)), so as to enable a more holistic understanding of daylight performance of buildings.

5.5. Concluding Remarks

This study, performed under a controlled laboratory setting and in a test room with direct access to daylight, has explored the temporal variation of glare sensation reported by test subjects at evenly distributed 3-hour intervals along the day. The main conclusions that can be drawn from this investigation are:

- Statistically significant and practically relevant evidence was found of a substantive effect of time of the day on the evaluation of glare sensation reported by test subjects. The evidence detected showed a tendency towards an increased tolerance to glare as the day progresses under both daylight and artificial lighting conditions.
- The experimental effect of interest (i.e., temporal variation of glare response) was found to be directly linked to the independent variable itself (i.e., time of the day). That is, as the time period between test sessions increased, so did the difference in variation of glare sensation reported by test subjects;
- When the photometric values corresponding to reported votes of glare sensation were regressed against the time of the day, a relatively large scatter was observed in the data. This suggests that there could be other variables, not controlled through experimental manipulation, nor fully considered within this investigation, which may influence subjective glare response.

- The analysis conducted under a laboratory setting with artificial lighting provided evidence that glare sensation might be influenced by the difficulty in extracting information from the visual stimulus.
- Among several personal factors and temporal variables considered, results from the laboratory studies suggested a tendency for earlier chronotypes, and for test subjects not having ingested caffeine, to show a greater tolerance to increases in artificial lighting source luminance. In addition, statistically significant and practically relevant evidence was found of a direct influence of fatigue, and an inverse effect of food intake, on the perceived level of glare sensation reported at different times of the day. Indication of these tendencies was also confirmed by the test room studies, although further investigation is needed in order to substantiate the influences detected.
- Further research is required to deepen the analysis of the role of chronotype, view importance and interest, and control, over the effect of time of day on glare response.

In conclusion, the findings from this investigation suggest that the conventional physical and photometric values utilised in glare and lighting formulae and standards are not sufficient to accurately describe and predict the subjective evaluation of glare sensation from natural and artificial light. This is leading discomfort glare research towards new, exciting, potentially ground-breaking, albeit unanticipated and unchartered developments. Yet, as Albert Einstein once said "*If we knew what it was we were doing, it would not be called research, would it?*".

References

Akashi, Y., Muramatsu, R. and Kanaya S., 1998. Unified glare rating (UGR) and subjective appraisal of discomfort glare. *Lighting Research and Technology*, 28, pp.199-206.

al Enezi, J., Revell, V., Brown, T., Wynne, J., Schlagen, L. and Lucas, R., 2011. A "Melanopic" spectral efficiency function predicts the sensitivity of melanopsin photoreceptors to polychromatic lights. *Journal of Biological Rhythms*, 26(4), pp. 314-323.

Almeida, D.M., Piazza, J.R. and Stawski, R.S., 2009. Inter-individual differences and intra-individual variability in the cortisol awakening response: An examination of age and gender. *Psychology and Aging*, 24(4), pp.819-827.

Altomonte, S., 2009. Daylight and the occupant. Visual and physio-psychological well-being in built environments, In: *Proceedings of PLEA 2009 – 26th Conference on Passive and Low Energy Architecture*. Quebec City, Canada, 22-24 June 2009, pp.331-336.

Andersen, M., 2015. Unweaving the human response in daylighting design. *Building and Environment*, 91, pp.101-117.

Andersen, M., Mardaljevic, J. and Lockley, S.W., 2012. A framework for predicting the non-visual effects of daylight – Part I: Photobiology-based model. *Lighting Research and Technology*, 44(1), pp.37-52.

Arendt, J., 1995. *Melatonin and the mammalian pineal gland, first edition*. London: Chapman & Hall.

Aries, M., 2005. *Human lighting demands. Healthy lighting in an office environment*. Ph.D. Technishche Universiteit Eindhoven.

Bellchambers, H.E., Collins, J.B. and Crisp, V.H.C., 1975, Relationship between two systems of glare limitation. *Lighting Research and Technology,* 7(2), pp.106-112.

Bellia, L., Cesarano, A., Minichiello, F. and Sibilio, S., 2002. Setting up a CCD photometer for lighting research and design. *Building and Environment*, 37, pp.1099-1106.

Beretvas, S.N. and Pastor, D.A., 2003. Using mixed-effects models in reliability generalization studies. *Educational and Psychological Measurement*, 63(1), pp.75-95.

Berson, D.M., 2003. Strange vision: ganglion cells as circadian photoreceptors. *Trends in Neuroscience*s, 26(6), pp.314-319.

Berson, D.M., 2007. Phototransduction in ganglion-cell photoreceptors. *Pflugers Archive – European Journal of Physiology*, 454, pp.849-855.

Berson, D.M., Dunn, F.A. and Takao, M., 2002. Phototransduction by retinal ganglion cells that set the circadian clock. *Science*, 295, pp.1070-1073.

Bird, K.D. and Hadzi-Pavlovic, D., 2014. Controlling the maximum family Type I error rate in analyse of multivariate experiments. *Psychological Methods*, 63(1), pp.75-95.

Boff, K.R. and Lincoln, J.E., ed. 1988. *User's guide. Engineering data compendium: Human perception and performance*. In: Armstrong, H.G., 1988. Weight-Patterson Air Force Base, Ohio: Human Engineering Division.

Borisuit, A., Linhart, F., Scartezzini, J.L. and Munch, M., 2015. Effects of realistic office daylighting and electric lighting conditions on visual comfort, alertness and mood. *Lighting Research and Technology*, 47(2), pp.199-209.

Boschman, M.C. and Roufs, J.A.J., 1997. Text quality metrics for visual display units: II. An experimental survey. *Displays*, 18, pp.45-64.

Boyce, P., 2015. Editorial: Advancing lighting controls. *Lighting Research and Technology*, 47, p.131.

Boyce, P.R., 1974. Illuminance, difficulty, complexity and visual performance. *Lighting Research and Technology*, 6, pp.222-226.

Boyce, P.R., 2004. Lighting research for interiors: the beginning of the end or the end of the beginning. *Lighting Research and Technology*, 36(4), pp.283-294.

Boyce, P.R., 2014. *Human Factors in Lighting*. *Third Edition*. New York: CRC Press.

Boyce, P., Hunter, C. and Howlett, O., 2003. *The benefits of daylight through windows*. Lighting Research Center Rensselaer Polytechnic Institute.

Boyce, P.R., Eklund, N.H. and Simpson, S.N., 2000. Individual lighting control: Task performance, mood and illuminance. *Journal of the Illuminating Engineering Society*, 29(1), pp.131-142.

Boyce, P.R., Veitch, J.A., Newsham, G.R., Jones, C.C., Heerwagen, J., Myer, M. and Hunter, C.M., 2006. Lighting quality and office work: Two field simulation experiments. *Lighting Research and Technology*, 38(3), pp.191-223.

Boubekri, M. and Boyer, L.L., 1992. Effect of window size and sunlight presence on glare. *Lighting Research and Technology*, 24(2), pp.69-74.

Brainard, G.C. and Hanifin, J.P., 2005. Photons, clocks and consciousness. *Journal of Biological Rhythms*, 20(4), pp.314-325.

Brainard, G.C., Hanifin, J.P., Greeson, J.M., Byrne, B., Glickman, G., Gerner, E. and Rollag, M.D., 2001, Action spectrum for melatonin regulation in humans: evidence for a novel circadian photoreceptor. *Journal of Neuroscience*, 21, pp.6405-6412.

Brainard, G.C., Lewy, A.J., Menaker, M., Fredrickson, R.H., Miller, L.S., Weleber, R.G., Cassone, V. and Hudson, D., 1988. Dose-response relationship between light irradiance and the suppression of plasma melatonin in human volunteers. *Brain Research*, 454, pp.212-218.

Burke, T.M., Markwald, R.R., McHill, A.W., Chinoy, E.D., Snider, J.A., Bessman, S.C., Jung, C.M., O'Neill, J.S. and Wright Jr., J.P., 2015. Effects of caffeine on the human circadian clock in vivo and in vitro. *Science Translational Medicine*, 7(205), pp.305ra146.

Buse, A., 1982. The likelihood ratio, wald, and lagrange multiplier tests. An expository note. *The Statistician*, 36(3), pp.153-157.

Cabin, R.J. and Mitchell, R.J., 2000. To Bonferroni or not to Bonferroni: When and how are the questions. *Bulletin of the Ecological Society of America*, 81(3), pp.246-248.

Cai, H. and Chung, T., 2013. Evaluating discomfort glare from non-uniform electric light source. *Lighting Research and Technology*, 45, pp.267-294.

Cai, H. and Chung, T.M., 2011. Improving the quality of high dynamic range images. *Lighting Research and Technology*, 43, pp.87-102.

Cajochen, C., 2007. Alerting effects of light. *Sleep Medicine Reviews*, 11, pp.453-464.

Cajochen, C., 2015. Light: An underestimated environmental factor in sleep medicine. *Sleep and Biological Rhythms*, 13, p.111.

Cajochen, C., Munch, M., Kobialka, S., Krauchi, K., Steiner, R., Oelhafen, P., Orgul, S. and Wirz-Justice, A., 2005. High sensitivity of human melatonin, alertness, thermoregulation, and heart rate to short wavelength light. *The Journal of Clinical Endocrinology & Metabolism*, 90(3), pp.1311-1316.

Cajochen, C., Zeitzer, J.M., Czeisler, C.A. and Dijk, D.J., 2000. Dose-response relationship for light intensity and ocular and electroencephalographic correlates of human alertness. *Behavioural Brain Research*, 115, pp.75-83.

Chai, T. and Draxler, R.R., 2014. Root mean square error (RMSE) or mean absolute error (MAE)? – Arguments against avoiding RMSE in the literature. *Geoscientific Model Development*, 7, pp.1247-1250.

Chang, A.M., Santhi, N., St. Hilaire, M., Gronfier, C., Bradstreet, D.S., Duffy, J.F., Lockley, S.W., Kronauer, R.E. and Czeisler, C.A., 2012. Human direction responses to bright light. *The Journal of Physiology*, 590(13), pp.3103-3112.

Chang, A.M., Scheer, F.A.J.L. and Czeisler, C.A., 2011. The human circadian system adapts to prior photic history. *The Journal of Physiology*, 589.5, pp.1095-1102.

Chauvel, P., Collins, J.B., Dogniaux, R. and Longmore, J., 1982. Glare from windows: current views of the problem. *Lighting Research and Technology*, 14(1), pp.31-46.

Christoffersen, J. and Wienold, J., 2008. Assessment of user reaction to glare with three solar shading systems. In: *Proceedings of Indoor Air*. Copenhagen, Denmark, 17-22 August 2008.

Christoffersen, J., Petersen, E., Johnsen, K., Valbjorn, O. and Hygge, S., 2000. Windows and daylight – A post-occupant evaluation of Danish offices. In: *Lighting 2000 – ILE/CIBSE Conference*. York, United Kingdom, 9-10 July, 2000, pp.112-120.

CIBSE, 1994. *Code for Interior Lighting*, London: CIBSE.

CIE, 1995. *Discomfort Glare in Interior Lighting*, CIE Technical Report 117, Vienna: CIE – Commission Internationale de l’Eclairage.

CIE, 1993. *Discomfort glare in interior lighting*. CIE – Commission Internationale de l'Eclairage (CIE).

CIE, 2015. *CIE Statement on non-visual effects of light: recommending proper light at the right time*. Vienna: CIE – Commission Internationale de l'Eclairage.

Clear, R. and Berman, S., 1989. Speed, accuracy and VL. *Illuminating Engineering Society*, 19, pp.164-177.

Cohen, J., 1965. Some statistical issues in psychological research. In: Wolman, B.B. ed. 1965. *Handbook of clinical psychology*. New York: McGraw-Hill.

Cohen, J., 1994. The earth is round (p<.05). *American Psychologist*, 49(12), pp.997-1003.

Cohen, J., Cohen P., West, S.G. and Aiken, L.S., 2003. *Applied multiple regression/correlation analysis for the behavioural sciences*. Third Edition. London: Lawrence Erlbaum Associates.

Collins, B.L., 1975. *Windows and people: A literature survey. Psychological reaction to environments with and without windows*. Washington DC: National Bereau of Standards.

Collins, B.L., 1976. Review of psychological reaction to windows. *Lighting Research and Technology*, 8(2), pp.80-88.

Coutelier, B. and Dumortier, D., 2002. Luminance calibration of the Nikon Coolpix 990 digital camera. Application to glare evaluation. In: *Energy efficiency and healthy buildings in sustainable cities*. Lyon, France, 23-26 October, 2002.

Chronoicle, E.P. and Mulleners, W.M., 2002. Visual system dysfunction in migraine: a review of clinical and psychophysical findings. *Cephalagia*, 16(8), pp.525-535.

Cuttle, C., 1971. Lighting patterns and the flow of light. *Lighting Research and Technology*, 3(3), pp.171-189.

Daurat, A., Foret, J., Benoit, O. and Mauco, G., 2000. Bright light during nighttime: effects on the circadian regulation of alertness and performance. *Biological Signals and Receptors*, 9, pp.309-318.

Daurat, A., Foret, J., Touitou, Y. and Benoit, O., 1996. Detrimental influence of bright light exposure on alertness, performance, and mood in the early morning. *Clinical Neurophysiology*, 26, pp.8-14.

Duffy, J.F. and Czeisler, C.A., 2009. Effect of light on human circadian physiology. *Sleep Medicine Clinics*, 4(2), pp.165-177.

Duffy, J.F., Dijk, D.J., Hall, E.F. and Czeisler, C.A., 1999. Relationship of endogenous circadian melatonin and temperature rhythms to self-reported preference for morning or evening activity in young and older people. *Journal of Investigative Medicine*, 47(3), pp.141-150.

Duffy, J.F., Kronauer, R.E. and Czeisler, C.A., 1996. Phase-shifting human circadian rhythms: influence of sleep timing, social contact and light exposure. *Journal of Physiology*, 496.1, pp.289-297.

Dumont, M. and Beaulieu, C., 2007. Light exposure in the natural environment: Relevance to mood and sleep disorders. *Sleep Medicine*, 8, pp.557-565.

Einhorn, H.D., 1961. Pre-determination of direct discomfort glare. *Lighting Research and Technology*, 26(4), pp.154-164.

Einhorn, H.D., 1969. A new method for the assessment of discomfort glare. *Lighting Research and Technology*, 1(4), pp.235-247.

Einhorn, H.D., 1979. Discomfort glare: a formula to bridge differences. *Lighting Research and Technology*, 11(2), pp.90-94.

Einhorn, H.D., 1997. Discussion: Position index for a glare source located below the line of vision. *Lighting Research and Technology*, 29(3), pp.172-178.

Einhorn, H.D., 1998. Unified glare rating (UGR): Merits and application to multiple sources. *Lighting Research and Technology*, 30(2), pp.89-93.

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

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

Enders, C.K., 2010. *Applied missing data analysis*. New York: The Guildford Press.

Engle, R.F., 1984. Wald, likelihood ratio, and lagrange multiplier tests in econometrics. *Handbook of Econometrics*, 2, pp.775-826.

Evans, R.W. and Digre, K.B., 2003. Expert opinion: Light sensitivity in migraineurs. *Headache: The Journal of Head and Face Pain*, 43, pp.917-920.

Ferguson, C.J., 2009. An effect size primer: A guide for clinicians and researchers. *Professional Psychology: Research and Practice*, 40(5), pp.532-438.

Field, A., 2013. *Discovering statistics using IBM SPSS statistics*. 4th Edition. London: Sage.

Field, A., 2016. *An adventure in statistics*: The reality enigma. London: Sage.

Field, A. and Hole, G., 2013. *How to design and report experiments*. London: Sage.

Figueiro, M.G., 2003a. Research Recap: Circadian rhythm. *Lighting Design + Application (LD+A)*, 33(2), pp.17-18.

Figueiro, M.G., 2003b. Research Recap: Light, aging & the circadian system – Reviving ‘all that jazz? *Lighting Design + Application (LD+A)*, 33(6), pp.8-11.

Figueiro, M.G., 2008. A proposed 24h lighting scheme for older adults. *Lighting Research and Technology*, 40, pp.153-160.

Figueiro, M.G. and Rea, M.S., 2010. The effects of red and blue lights on circadian variations in cortisol, alpha amylase, and melatonin. *International Journal of Endocrine*, pp.1-9.

Figueiro, M.G., Bierman, A. and Rea, M.S., 2008. Retinal mechanisms determine the subadditive response to polychromatic light by the human circadian system. *Neuroscience Letters*, 438, pp.242-245.

Figueiro, M.G., Plitnick, B. and Rea, M.S., 2014. The effects of chronotype, sleep schedule and light/dark pattern exposure on circadian phase. *Sleep Medicine*, 15, pp.1554-1564.

Figueiro, M.G., Wood, B., Plitnick, B. and Rea, M.S., 2011. The impact of light from computer monitors on melatonin levels in college students. *Neuroendocrinology Letters*, 32(2), 158-163.

Flannagan, M., Sivak. M. and Traube, E.C., 1994. Discomfort glare and brightness as functions of wavelength. *The University of Michigan Report UMTRI-94-29*.

Flannagan, M., Sivak, M., Ensing, M. and Simmons, C.J., 1989. Effect of wavelength on discomfort glare from monochromatic sources. *The University of Michigan Report UMTRI-89-30*.

Flannagan, M., Wintraub, D. and Sivak, M., 1990. Context effects on discomfort glare: Task and stimulus factors. *The University of Michigan Report UMTRI-90-35*.

Fotios, S., 2015. Research Note: Uncertainty in subjective evaluation of discomfort glare. *Lighting Research and Technology*, 47, pp.379-383.

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

Friedman, M., 1937. The use of ranks to avoid the assumption of normality implicit in the analysis of variance. *American Statistical Association*, 32(200), pp.675-701.

Friedman, M., 1940. A comparison of alternative tests of significance for the problem of m rankings. *The Annals of Mathematical Statistics*, 11(1), pp.86-92.

Fritz, C.O., Morris, P.E. and Richler, J.J., 2012. Effect size estimates: Current use, calculations, and interpretations. *Journal of Experimental Psychology: General*, 141(1), pp.2-18.

Froy, O. and Miskin, R., 2010. Effects of feeding regiments on circadian rhythms: Implications for aging and longevity. *AGING*, 2(1), pp.7-27.

Fry, G.A. and Alpern, M.A., 1955. Effect of a peripheral glare source upon the apparent brightness of an object. *Journal of the Optical Society of America*, 43(3), pp.189-195.

Fry, G.A. and King, V.M., 1975. The pupillary response and discomfort glare. *Journal of Illuminating Engineering Society*, 4, pp.307-324.

Goldstein, H., 2010. *Multilevel statistical models*. 4th Edition. New York: John Wiley & Sons.

Gooley, J.J., Lu, J., Chou, T.C., Scammell, T.E. and Saper, C.B., 2001. Melanopsin in cells of origin of the retinohypothalamic tract. *Nature Neuroscience*, 4(1), p. 1165.

Gooley, J.J., Lu, J., Fischer, D. and Saper, C.B., 2003. A broad role for melanopsin in nonvisual photoreception. *The Journal of Neuroscience*, 23(18), pp.7093-7106.

Gronfier, C., Wright, J.P., Kronauer, R.E., Jewett, M.E. and Czeisler, C.A., 2004. Efficacy of a single sequence of intermittent bright light pulses for delaying circadian phase in humans. *The American Journal of Physiology – Endocrinology and Metabolism*, 287, pp.E174-E181.

Hannibal J., 2002. Neutransmitter of the retino-hypothalamic tract. *Cell and Tissue Research*, 309, pp.73-88.

Hanifin, J.P. and Brainard, G.C., 2006. Photoreception for circadian, neuroendocrine, and neurobehavioral regulation. *Journal of Physiological Anthropology*, 26, pp.87-94.

Hart, A., 2001. Mann-Whitney test is not just a test for medians: difference in spread can be important. *Education and debate*, 323, pp.391-393.

Hatter, S., Liao, H.W., Takao, M., Berson, D.M. and Yau, K.W., 2002. Melanopsin-containing retinal ganglion cells: Architecture, projections, and intrinsic photosensitivity. *Science*, 295(5557), pp.1065-1070.

Hauschke, D. and Steinijans, V.W., 1996. Directionality decision for a two-tailed alternative. *Journal of Biopharmaceutical Statistics*, 6, pp.211-213.

Hayes, A.F., 2006. A primer on multilevel modelling. *Human Communication Research*, 32(4), pp.385-410.

Heiman, G.W., 2013. *Basic statistics for the behavioural sciences, 7th Edition*. Stamford: Cengage.

Holladay, L.L., 1926. The fundamentals of glare and visibility. *Journal of the Optical Society of America and Review of Scientific Instruments*, 12(4), pp.271-319.

Hopkinson, R.G., 1940. Discomfort glare in lighted streets. *Lighting Research and Technology*, 5(1-9), pp.1-32.

Hopkinson. R.G., 1950. The multiple criterion technique of subjective appraisal. *Quarterly Journal of Experimental Psychology*, 2-3, pp.124-131.

Hopkinson, R.G., 1955. Subjective judgements – some experiments employing experienced and inexperienced observers. *British Journal of Psychology*, 46(4), pp.262-272.

Hopkinson, R.G., 1956. Glare discomfort and pupil diameter. *Journal of Optical Society of America*, 46(8), pp.649-656.

Hopkinson, R.G., 1957. Evaluation of glare. *Illuminating Engineering*, 52, pp.305-316.

Hopkinson, R.G., 1963. *Architectural physics: Lighting*. London: Her Majesty's Stationary Office.

Hopkinson, R.G., 1971. Glare from windows. *Construction Research and Development*, 2(3), pp.98-105.

Hopkinson, R.G., 1972. Glare from daylighting in buildings. *Applied Ergonomics*, 3(4), pp.206-215.

Hopkinson, R.G. and Bradley, R.C, 1960. A study of glare from very large sources. *Illuminating Engineering*, 55(5), pp.288-294.

Hopkinson, R.G. and Collins, J.B., 1970. *The Ergonomics of Lighting*. London: MacDonald Technical and Scientific.

Hopkinson, R.G. and Peterbridge, P., 1954. Two supplementary studies on glare. *Illuminating Engineering Society*, 19(7), pp.220-224.

Hopkinson, R.G., Peterbridge, P. and Longmore, J., 1966. *Daylighting*. London, Heinemann.

Howarth, P.A., Heron, G., Greenhouse, D.S., Bailey, I.L. and Berman, S.M., 1993. Discomfort from glare: The role of pupillary hippus. *Lighting Research and Technology*; 25(1), pp.37-42.

Hox, J., 1998. *Multilevel modelling when and why*. In: Balderjahn, I., Mathar, I. and Scahader, M., 1998. Classification, Data Analysis, and Data Highways. New York: Springer, pp.147-154.

Huang, F.L. and Cornell, D.G., 2015. Using multilevel factor analysis with clustered data: Investigating the factor structure of the positive values scale. *Journal of Psychoeducational Assessment*, DOI: 10.1177/0734282915570278.

Inanici, M.N., 2006. Evaluation of high dynamic range photography as a luminance data acquisition system. *Lighting Research and Technology*, 38(2), pp.123-136.

Iwata, T., Shukuya, M. and Somekawa, N., 1992a. Experimental study of discomfort glare caused by windows: Subjective responses to glare from simulated windows. *Journal of Architecture, Planning and Environmental Engineering*, 432, pp.21-30.

Iwata T., Shukuya, M. and Somekawa, N., 1992b. Experimental study on discomfort glare caused by windows: Subjective responses to glare from actual windows. *Journal of Architecture, Planning and Environmental Engineering*, 439, pp.19-31.

Iwata, T. and Tokura, M., 1997. Position Index for a glare source located below the line of vision. *Lighting Research and Technology*, 29(3), pp.172-178.

Iwata, T. and Tokura, M., 1998. Examination of limitations of predicted glare sensation vote (PGSV) as a glare source – Towards a comprehensive development of discomfort glare evaluation. *Lighting Research and Technology*, 30, pp.81-88.

Jay P., 1989. Correspondence. *Lighting Research and Technology*, 21(2), p.81.

Jonckheere, A.R., 1954. A distribution-free k-sample test against ordered alternatives. *Biometrika*, 41(1-2), pp.133-145.

Julian, M.W., 2001. The consequences of ignoring multilevel data structures in nonhieracrchical covariance modelling. *Structural Equation Modelling*, 8, pp.325-352.

Kaplan, S. and Kaplan, R., 1989. *The experience of nature: A psychological perspective*. Cambridge: Cambridge University Press.

Karlsen, L., Heiselberg, P. and Bryn, I., 2015. Occupant satisfaction with two blind control strategies: Slats closed and slats in cut-off position. *Solar Energy*, 115, pp.166-179.

Kaufman, J.E., 1984. *IES Lighting Handbook: 1984 Reference Volume (Ed.)*. New York: Illuminating Engineering Society of North America.

Kenwood, M.G. and Roger, J.H., 1997. Small sample influence for fixed effects from restricted maximum likelihood. *Biometrics*, 53, pp.983-997.

Khalsa, S.B.S., Jewett, M.E., Cajochen, C. and Czeisler, C.A., 2003. A phase response to single bright light pulses in human subjects. *Journal of Physiology*, 549.3, 945-952.

Kim, W., Han, H. and Kim, J.T., 2009. The position index of a glare source at the borderline between comfort and discomfort (BCD) in the whole visual field. *Building and Environment*, 44, pp. 1017-1023.

Kreft, I.T.A. and de Leeuw, J., 1998. *Introducing multilevel modelling*. London: Sage.

Kuhn, T., Wienold, J. and Moosmann, C., 2013. Daylight glare. Age effects and their impact on glare evaluation. In: *Proceedings of the Energy Forum: Advanced Building Skins: Advanced Buildings Skins*. Bressanone, Italy, 5-6 November, 2013, pp.123-128.

Kruskal, W.H. and Wallis, A., 1952. Use of ranks in one-criterion variance analysis. *Journal of the American Statistical Association*, 47(260), pp.583-621.

Landman, J.T. and Dawes, R.M., 1982. Psychotherapy Outcome: Smith and Glass' conclusions stand up under scrutiny. *American Psychologist*, 37(5): pp. 504-516.

Leproult, R., Colecchia, E.F., L'Hermite-Baleriaux, M. and Van Cauter, R., 2001. Transition from dim to bright light in the morning induces an immediate elevation of cortisol levels. *The Journal of Clinical Endocrinology & Metabolism*, 86(1), pp.151-157.

Lewy, A.J., Wehr, T.A., Goodwin, F.K., Newsome, D.A. and Markey, S.P., 1980. Light suppresses melatonin secretion in humans. *Science*, 210, pp.1267-1269.

Levandovski, R., Sasso, E. and Hidalgo, M.P., 2013. Chronotype: a review of the advances, limits and applicability of the main instruments used in the literature to assess human phenotype. *Trends in Psychiatry and Psychotherapy*, 35(1), pp.3-11.

Lim, J., 2011. Hedomic scaling: A review of methods and theory. *Food Quality and Preference*, 22, pp.733-747.

Lockley, S.W., 2009. Circadian rhythms: Influence of light in humans. *Encyclopaedia of Neuroscience*, 2, pp.971-988.

Lockley, S.W. and Gooley, J.J., 2006. Circadian photoreception: Spotlight on the brain. *Current Biology*, 16(18), pp. R795-R797.

Lockley, S.W., Evans, E.E., Scheer, F.A.J.L., Brainard, G.C., Czeisler, C.A. and Aeschbach, D., 2006. Short-wavelength sensitivity for the direct effects of light on alertness, vigilance, and the waking electroencephalogram in humans. *Sleep*, 29(2), pp.161-168.

Lynes, J.A., 1974. The window as a communication channel. *Light and Lighting*, 67, pp.284-288.

Lynes, J.A., 1977. Discomfort glare and visual distraction. *Lighting Research and Technology*, 9(1), pp.51-52.

Lucas, R.J., 2013. Mammalian inner retinal photoreception. *Current Biology*, 23, pp.R125-R133.

Lucas, R.J., Peirson, S.N., Berson, D.M., Brown, T.M., Cooper, H.M., Czeisler, C.A., Figueiro, M.G.,Gamlin, P.D., Lockley, S.W., O'Hagan, J.B., Price, L.L.A., Provencio, I., Skene, D.J. and Brainard, G.C., 2014. Measuring and using light in the melanopsin age. *Trends in Neuroscience*, 37(1), pp.1-9.

Luckiesh, M. and Guth, S.K., 1949. Brightness in the visual field at the borderline between comfort and discomfort. *Illuminating Engineering*, 44(11), pp.650-670.

Luckiesh, M. and Holladay, L.L., 1925. Glare and visibility. *Transactions of the IES*, 20, pp.221-231.

Luckiesh, M. and Holladay, L.L., 1926. Measuring visibility under preventable glare. *Journal of the Optical Society of America*, 12, pp.215-217.

Maas, C.J.M. and Hox, J.J., 2004. The influence of violations of assumptions on multilevel parameter estimates and their standard errors. *Computational Statistics & Data Analysis*, 46, pp.427-440.

Maas, C.J.M., and Hox, J.J., 2005. Sufficient sample sizes for multilevel modelling. *Methodology*, 1(3), pp.86-92.

MacGowan, D., 2010. Correspondence. *Lighting Research and Technology*, 42, p.121.

MacGowan, D. and Clear, R., 2013. Correspondence: Discomfort glare. *Lighting Research and Technology*, 45, pp.258-260.

MacGowan, D., Pulpitlova, J. and Subova, A., 1993. Time and circumstance, and the perception of discomfort glare. In: *Proceedings of Lux Europa*, Edinburgh, United Kingdom, 4-7 April, 1993, pp.193-211.

MacKinnon, D.P., Krull, J.L. and Lockwood, C.M., 2000. Equivalent of the mediation, confounding and suppression. *Society of Prevention Research*, 1(4), pp.173-181.

Mann, H.B. and Whitney, D.R., 1947. On a test whether one of two random variables is stochastically larger than the other. *The Annals of Mathematical Statistics*, 18(1), pp.50-60.

Mardaljevic, J., Andersen, M., Roy, N. and Christoffersen J., 2012. *Daylighting, Artificial Lighting and Non-Visual Effects Study for a Residential Building*. [Online] Available at: http://infoscience.epfl.ch/record/181055. [Accessed 07 March 2016].

Markus, T.A., 1967. The function of windows – a reappraisal. *Building Science*, 2, pp.97-121.

McIntyre, I.M., Norman, T.R., Burrows, G.D. and Armstrong, S.M., 1989a. Quantal melatonin suppression by exposure to low intensity light in man. *Life Sciences*, 45(4), pp.327-332.

McIntyre, I.M., Norman, T.R., Burrows, G.D. and Armstrong, S.M., 1989b. Human melatonin suppression by light is intensity dependent. *Journal of Pineal Research*, 6, pp.149-156.

Minor, D.S., Waterhouse, J.M. and Wirz-Justice, A., 1991. A human phase-response curve to light. *Neuroscience Letters*, 113(1), pp.36-40.

Mistlberger, R.E., 2011. Neurobiology of food anticipatory circadian rhythms. *Physiology & Behavior*, 104, pp.535-545.

Mochizuki, E., Itoh, D. and Iwata, T., 2009. Effects on discomfort glare of luminance distribution with a window. *Journal of Environmental Engineering, Architectural Institute of Japan*, 74, pp.277-282.

Monk, T.H., 2005. The post-lunch dip in performance. *Clinics in Sports Medicine*, 24, pp.15-23.

Murray, I.J., Plainis, S. and Carden, D., 2002. The ocular stress monitor: a new device for measuring discomfort glare. *Lighting Research and Technology*, 34(3), pp.231-242.

Muthen, B.O., 1994. Multilevel Covariance Structure Analysis. *Sociological Methods & Research*, 22(3), pp.376-398.

Nakagawa, S., 2004. A farewell to Bonferroni: The problems of low statistical power and publication bias. *Behavioural Ecology*, 15(6), pp.1004-1045.

Nazzal, A.A., 2001. A new daylight glare evaluation method introduction of the monitoring protocol and calculation method. *Energy and Buildings*, 33, pp.257-265.

Nazzal, A.A., 2005. A new evaluation method for daylight discomfort glare. *International Journal of Industrial Ergonomics*, 35, pp.295-306.

Nazzal, A.A. and Chutarat, A., 2000. A new daylight glare evaluation method. A comparison of the existing glare index and the proposed method and an exploration of daylighting control strategies. *Architectural Science Review*, 44(1), 71-82.

Nordstokke, D.W. and Zumbo, B.D., 2010. A new nonparametric Levene test for equal variances. *Psicologica*, 31, pp.401-430.

Nordstokke, D.W., Zumbo, B.D., Cairns, S.L. and Saklofske, D.H., 2011. The operating characteristics of the nonparametric Levene test for equal variances with assessment and evaluation data. *Practical Assessment, Research & Evaluation*, 16(5), pp.1-8.

North, B.V., Curtis, D. and Sham, P.C., 2002. A Note on the calculation of empirical p values from Monte Carlo procedures. *American Journal of Human Generics*, 71(2), pp.439-441.

Okamoto, Y., Rea, M.S. and Figueiro, M.G., 2014. Temporal dynamics of EEG activity during short- and long-wavelength light exposures in the early morning. *MBC Research Notes*, 7(113), pp.1-6.

Osterhaus, W.K.E., 2005. Discomfort glare assessment and prevention for daylight applications in office environments. *Solar Energy*, 79, pp.140-158.

Osterhaus, W.K.E.E. and Bailey, I.L., 1992. Large area glare sources and their effect on discomfort and visual performance at computer workstations. In: *Proceedings of the 1992 IEEE Industry Applications Society Annual Meeting*. Houston Texas, United States, 4-9 October, 1992.

Paul, B.M. and Einhorn, H.D., 1999. Discomfort glare from small light source. *Lighting Research and Technology*, 31(4), pp.139-144.

Penny, W.D. and Holmes, A.J., 2003. Random-effects analysis. In: Frackowiak, R.S.J., Friston, K.J., Frith, C., Dolan, R., Friston, K.J., Price, C.J., Zeki, S., Ashburner, J., and Penny, W.D. ed. 2003. *Human Brain Function*. Academic Press, 2nd edition.

Peterbridge, P. and Hopkinson, R.G., 1950. Discomfort glare and the lighting of buildings. *Lighting Research and Technology*, 25(2), pp.39-79.

Perneger, T.V., 1998. What's wrong with Bonferroni adjustments. *British Medical Journal*, 316, p.1236.

Peugh, J.L., 2010. A practical guide to multilevel modelling. *Journal of School Psychology*, 48, pp.85-112.

Phipps-Nelson, J., Redman, J.R., Dijk, D.J. and Rajaratman, S.M.W., 2003. Daytime exposure to bright light, as compared to dim light, decreases sleepiness and improves psychomotor vigilance performance. *Sleep*, 26, pp.695-700.

Pokorny, J., Smith, V.C. and Lutze, M., 1987. Aging of the human lens. *Applied Optics*, 26(8), pp.1437-1440.

Poulton, E.C., 1979. Models for biases in judging sensory magnitude. *Psychological Bulletin*, 86, pp.777-803.

Provencio, I., Jiang, G., de Grip, W.J., Hayes, W.P. and Rollag, M.D., 1998. Melanopsin: An opsin in melanophores, brain, and eye. *Proceedings of the National Academy of Sciences of the United States of America*; 95(1), pp.340-345.

Provencio, I., Rodriquez, I.R., Jiang, G., Hayes, W.P., Moreira, E.F. and Rollag, M.D., 2000. A novel human opsin in the inner retina. *The Journal of Neuroscience*; 20(2), pp.600-605.

Provencio, I., Rollag, M.D. and Castrucci, A.M., 2002. Photoreceptive net in the mammalian retina. *Nature*, 415, p.493.

Pulpitlova, J. and Detkova, P., 1993. Impact of the cultural and social background on the visual perception in loving and working perception. In: *Proceedings of the International Symposium: Design of Amenity*. Fukuoka, Japan, 5-9 October, 1993, pp.216-227.

Raudenbush, S.W. and Bryk, A.S., 2002. *Hierarchical linear models: Applications and data analysis methods (Advanced quantitative techniques in the social sciences second edition*. London: Sage.

Rea, M.S., 1981. Visual performance with realistic methods of changing contrast. *Illuminating Engineering Society*, 10, pp.164-177.

Rea, M.S., 2000. *The IESNA Lighting Handbook Reference and Application, Ninth Edition (Edited)*. New York: Illuminating Engineering Society of North America.

Rea, M.S. and Figueiro, M., 2014. Quantifying light-dependent circadian disruption in humans and animal models. *Chronobiology International*, 31(10), pp.1239-1246.

Rea, M.S., Bullough, J.D. and Figueiro, M., 2001. Human melatonin suppression by light: a case for scotopic efficiency. *Neuroscience Letters*, 299, pp.45-48.

Rea, M.S., Figueiro, M. and Bullough, J.D., 2002. Circadian photobiology: an emerging framework for lighting practice and research. *Lighting Research and Technology*, 34(3), pp.177-190.

Robinson, W., Bellchambers, H.E., Burt, W., Frith, D., Grundy, J.T., Hewitt, H., Hopkinson, R.G., Longmore, J., Peterbridge, P. and Rowlands, E., 1962. The development of the IES glare index system. *Lighting Research and Technology*, 27(1), pp.9-26.

Rockcastle, S. and Andersen, M., 2013. *Annual dynamics of daylight variability and contrast. A simulation-based approach to quantifying visual effects in architecture*. London: Springer.

Rodriguez, R.G. and Pattini, A., 2014. Tolerance of discomfort glare from a large area source for work on a visual display. *Lighting Research and Technology*, 46, pp.157-170.

Rodriguez, R.G., Yamin Garreton, J.A. and Pattini, A.E., 2015. Glare and cognitive performance in screen work in the presence of sunlight. *Lighting Research and Technology*; doi: 10.1177/1477153515577851.

Roenneberg, T., Kuehnle, T., Juda, M., Kantermann, T., Allebrandt, K., Gordijn, M. and Merrow, M., 2007. Epidemiology of the human circadian clock. *Sleep Medicine Reviews*, 11, pp.429-438.

Roenneberg, T., Wirz-Justice, A. and Merrow, M., 2003. Life between clocks: daily temporal patterns of human chronotypes. *Journal of Biological Rhythms*, 18(1), pp.80-90.

Romano, J.L., 2007. *Examining the issues surrounding violating the assumption of independent observations in reliability generalization studies: A simulation study*. Ph.D. University of South Florida.

Rosenthal, R. and DiMatteo, M.R., 2001. Meta-analysis: Recent developments in quantitative methods for literature reviews. *Annual Review of Psychology*, 52, pp.59-82.

Rosenthal, R., Rosnow, R.L. and Rubin, D.B., 2000. *Contrasts and effect sizes in behavioural research: A correlation approach*. Cambridge: Cambridge University Press.

Roufs, J.A.J. and Boschman, M.C., 1997. Text quality metrics for visual display units: I. Methodological aspects. *Displays*, 18, pp.37-43.

Rowlands, E., 1997. Discussion: Unified Glare Rating (UGR) and subjective appraisal of discomfort glare. *Lighting Research and Technology*, 28(4), pp.199-206.

Ruger, M., Gordijn, M.C.M., Beersma, M., de Vries, B. and Daan, S., 2006. Time-of-day dependent effects of bright light exposure on human psychophysiology: comparison of daytime and nighttime exposure. *American Journal of Physiology – Regulatory, Integrative and Comparative Physiology*, 290(5), pp.R1413-R1320.

Ruxton, G.D. and Neuhauser, M., 2010. When should we use one-tailed hypothesis testing? *Methods in Ecology and Evolution*, 1, pp.114-117.

Samn, S.W. and Perelli, L.P., 1982. *Estimating aircrew fatigue: A technique with implications to airlift operations*. Report No. SAM-TR-82. Brookes AFB, TX: USAF School of Aerospace Medicine.

Sarey Khanie, M.S., Wienold, J. and Andersen, M., 2015. *A sensitivity analysis on glare detection parameters. Proceedings of 14th International Conference of the International Building Performance Simulation Association*, Hyderabad, December 7-9, 2015.

Sawilowsky, S.S., 2007. Mann-Whitney U test (Wilcoxon Rank-Sum Test). In: Salkind, N.J. and Rasmussen, K. ed. 2007. *Encyclopaedia of Measurement and Statistics*. London: Sage, pp.565-567.

Scariano, S.M. and Davenport, J.M., 1987. The effects of violations of independence assumptions in the one-way ANOVA. *American Statistical Association*, 41(2), pp.123-129.

Scheer, F.A.J.L., Pirovano, C., van Someren, E.J.W. and Buijs, R.M., 2005. Environmental light and suprachiasmatic nucleus interact in the regulation of body temperature. *Neuroscience*, 132, pp.465-477.

Scheer, F.A.J.L., van Doornen, L.J.P. and Buijs, R.M., 2004. Light and diurnal cycle affect autonomic cardiac balance in human; possible role for the biological clock. *Autonomic Neuroscience: Basic and Clinical*, 110, pp.44-48.

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

Seltman, H.J., 2012. *Experimental design and analysis*. Carnegie Mellon University.

Shapiro, S.S. and Wilk, M.B., 1965. An analysis of variance test for normality (complete samples). *Biometrika*, 52(3-4), pp.591-611.

Siegel, S., 1957. Nonparametric statistics. *The American Statistician*, 11(3), pp.13-19.

Sivak, M., Flannagan, M., Ensing, M. and Simmons, C.J., 1989. Discomfort glare is task dependent. *The University of Michigan Report UMTRI-89-27*.

Sivak, M., Schoettle, B. and Flannagan, M., 2003. LED headlamps: Glare and color rendering. *The University of Michigan UMTRI-2003-39*.

Sivak, M., Schoettle, B., Minoda, T. and Flannagan, M., 2005. Blue content of LED headlamps and discomfort glare. *The University of Michigan UMTRI-2005-2*.

Slater, A.L. and Boyce, P.R, 1990. Illuminance uniformity on desks: Where is the limit? *Lighting Research and Technology*, 22(4), pp. 165-174.

Slater, A.L., Perry, M.J. and Crisp, V.H.C., 1983. The applicability of the CIE visual performance model to lighting design. In: *Proceedings of the CIE 20th Session*. Amsterdam, Netherlands, Paris, France,

SLL, 2012. *The SLL Code for Lighting*. London: CIBSE.

Smith, A., 2002. Effects of caffeine on human behaviour. *Food and Chemical Toxicology*, 40, pp.1243-1255.

Smith, A., Sutherland, D. and Christopher, G., 2005. Effects of repeated doses of caffeine on mood and performance of alert and fatigued volunteers. *Journal of Psychopharmacology*, 19(6), pp.620-626.

Smolders, J.C.H.J. and de Kort, Y.A.W., 2014. Bright light and mental fatigue: Effects on alertness, vitality, performance and physiological arousal. *Journal of Environmental Psychology*, 39, pp.77-91.

Smolders, K.C.H.J., de Kort, Y.A.W. and Cluitmans, P.J.M., 2012. A higher illuminance induces alertness even during office hours: findings on subjective measurements, task performance and heart rate measures. *Physiology and Behavior*, 107, pp.7-16.

Snijders, T.A.B., 2005. Fixed and Random Effects. In: B.S. Everitt and D.C. Howell eds. 2005. *Encyclopaedia of Statistics in Behavioural Science*, 2, pp.664-665.

Snijders, T.A.B. and Bosker, R.J., 1994. Model variance in two-level models. *Sociological Methods & Research*, 22(3), pp.342-363.

Sorensen, K., 1987. A modern glare index method. In: *Proceedings of Commission Internationale de l'Eclairage* (CIE). Venice, Italy, 1987, pp.108-111.

Stephan, F.K., 2002. The "other" circadian system: Food as a zeitgeber. *Journal of Biological Rhythms*, 17(4), pp.284-292.

St. Hilaire, M.A., Gooley, J.J., Khalsa, S.B.S., Kronaur, R.E., Czeisler, C.A. and Lockley, S.W., 2012. Human phase response curve (PRC) to a 1-hour pulse of bright white light. *The Journal of Physiology*, 590(13), pp.3035-3045.

Stockman, A. and Lindsay, T.S., 2000. The spectral sensitivities of the middle- and long-wavelength-sensitivity cones derived from measurements in observers of known genotype. *Vision Research*, 40, pp.1711-1737.

Stone, P.T., 1999. The effect of environmental illumination on melatonin, bodily rhythms and mood states: a review. *Lighting Research and Technology*; 31(3), pp.71-79.

Stone, P.T., 2009. A model for the explanation of discomfort and pain in the eye caused by light. *Lighting Research and Technology*, 41, pp.109-121.

Stone, P.T. and Groves, S.D.P., 1968. Discomfort glare and visual performance. *Lighting Research and Technology*, 33(1), pp.9-15.

Stone, P.T. and Harker, S.D.P., 1973. Individual and group differences in discomfort glare responses. *Lighting Research and Technology*, 5(1), pp.41-49.

Tennessen, C.M. and Cimprich, B., 1995. View to nature: effect on attention. *Journal of Environmental Psychology*, 15, pp.77-85.

Thapan, K., Arendt, J. and Skene, D.J., 2001. An action spectrum for melatonin suppression: evidence for a novel non-rod, non-cone photoreceptor system in humans. *Journal of Physiology*, 535.1, pp.261-267.

Tokura, M., Iwata, T. and Shukuya, M., 1996. Experimental study on discomfort glare caused by windows, part 3. Development of a method for evaluating discomfort glare from large light source. *Journal of Architecture, Planning and Environmental Engineering*, 489, pp.17-25.

Torrington, J. and Tregenza, P.R., 2007. Lighting for people with dementia. *Lighting Research and Technology*, 39(1), pp.81-97.

Tregenza, P.R., 1998. Paired illuminance measurements: Eliminating calibration error. *Lighting Research and Technology*; 30(3), pp.133-134.

Tregenza, P. and Loe, D., 1998. *The Design of Lighting First Edition*. London, Routledge.

Tregenza, P. and Loe, D., 2014. *The design of lighting, second edition*. London, Routledge.

Tregenza, P. and Wilson, M., 2011. *Daylighting Architecture and Lighting Design*. London, Routledge.

Tregenza, P.R., Romaya, M., Dawe, S.P., Heap, L.J. and Tuck, B., 1974. Consistency and variation in preferences for office lighting. *Lighting Research and Technology*, 6(4), pp.205-211.

Tuaycharoen, N. and Tregenza, P.R., 2005. Discomfort glare from interesting images. *Lighting Research and Technology,* 37, pp.329-341.

Tuaycharoen, N. and Tregenza, P.R., 2007. View and discomfort glare from windows. *Lighting Research and Technology*, 39, pp.185-200.

Tuaycharoen, N., 2006. *The reduction of discomfort glare from windows by interesting views*. Ph.D. University of Sheffield.

Tuaycharoen, N., 2011. Windows are less glaring when there is a preferred view. *Built – Environment*, 09-10(01-02), pp.45-55.

Ulrich, R.S., 1981. Natural versus urban scenes: some psychological effects. *Environment and Behaviour*, 13, pp.523-556.

Ulrich, R.S., 1984. View from the window may help influence recovery from surgery. *Science*, 224, pp.420-421.

van Bommel, W.J.M., 2006. Non-visual biological effect of lighting and the practical meaning for lighting for work. *Applied Ergonomics*, 37(4), pp.461-466.

van Bommel, W.J.M. and van den Beld, G.J., 2004. Lighting for work: a review of visual and biological effects. *Lighting Research and Technology*, 36(4), pp.225-266.

van den Wymelenberg, K. and Inanici, M., 2014. A critical investigation of common lighting design metrics for predicting human visual comfort in offices with daylight. *LEUKOS: The Journal of the Illuminating Engineering Society of North America*, 10(3), pp.145-164.

Vandewalle, G., Balteau, E., Philips, C., Degueldre, C., Moreau, V., Sterpenich, V., Albouy, G., Darsaud, A., Desseilles, M., Dang-Vu, T.T., Peigneux, P., Luxen, A., Dijk, D.J. and Maquet, P., 2006. Daytime light exposure dynamically enhances brain responses. *Current Biology*, 16, pp.1616-1621.

Vanecek J., Pavlik A. and Illnerova H., 1987. Hypothalamic melatonin receptor sites revealed by autoradiography. *Brain Research*, 1987, 435(1-2), pp.359-361.

Veitch, J.A., 2001. Psychological processes influence lighting quality. *Journal of the Illuminating Engineering Society*, 30(1), pp.124-140.

Veitch, J.A., 2005. Light, lighting, and health: issues for consideration. *LEUKOS: The Journal of the Illuminating Engineering Society of North America*, 2(2), pp.85-96.

Veitch, J.A. and Newsham, G., 2000. Preferred luminous conditions in open-plan offices: research and practice recommendations. *Lighting Research and Technology*, 34(4), pp.199-212.

Velds, M., 1999. *Assessment of lighting quality in office rooms with daylighting systems*. Ph.D. Technishche Universiteit Delft.

Velds, M., 2002. User acceptance studies to evaluate discomfort glare in daylit rooms. *Solar Energy*, 73, pp.95-103.

Vink, J.M., Groot, A.S., Kerkhof, G.A. and Boomsma, D.I., 2001. Genetic analysis of morningness and evening. *Chronobiology International*, 18(5), pp.809-822.

Vogt, W.P., 1999. *Dictionary of statistics and methodology: A nontechnical guide for the social sciences, 2nd Edition*. Thousand Oaks, CA. London: Sage.

Vos, J.J., 1999. Glare today in historical perspective: Towards a new CIE glare observer and a new glare nomenclature. In: *Proceedings of the CIE 24th Session*. Warsaw, Poland, 1999, pp.38-42.

Waters, C., Mistrick, R.G. and Bernecker, C.A., 1995. Discomfort glare from sources of non-uniform luminance. *Journal of the Illuminating Engineering Society*, 24(2), pp.73-85.

Ward, G., 2005. *Photosphere*. Albany: Anywhere Software.

Wienold, J., 2004. *Evalglare: A Radiance based tool for glare evaluation. Proceedings of 3rd International Radiance Workshop*, Fribourg, October 11-12.

Wienold, J., 2009a. *Daylight Glare in Offices*. Ph.D. Fraunhofer ISE.

Wienold, J., 2009b. Dynamic daylight glare evaluation. In: *Proceedings Eleventh International IBPSA Conference*. Glasgow, Scotland, 27-30 July, 2009b, pp.944-951.

Wienold, J., 2015. *Evalglare*. Fraunhofer: Fraunhofer-Institut fur Soloare Energiesystem ISE.

Wienold, J. and Christoffersen, J., 2005. *Towards a new daylight glare rating*. In: LuxEuropa, Berlin, Germany, 19-21 September, 2005, pp.157-161.

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

Wilcoxon, F., 1945. Individual comparisons by ranking methods. *Biometrics Bulletin*, 1(6), pp.80-83.

Wilk, M.B. and Gnanadesikan, R., 1968. Probability plotting methods for the analysis of data. *Biometrika*, 55(1), pp.1-17.

Williams, L.J. and Abdi, H., 2010. Fisher's least significant difference (LSD) test. In: Salkind, N.J. ed. 2010, *Encyclopaedia of Research Design*. London: Sage.

Willmott, C.J. and Matsuura, K., 2005. Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance. *Climate Research*, 30, pp.79-82.

Willmott, C.J., Matsuura, K. and Robeson, S.M., 2009 Ambiguities inherent in sums-of-squares-based error statistics. *Atmospheric Environment*, 43, pp.749-752.

Wilson, M.P., 2007. Discussion: View and discomfort glare from windows. *Lighting Research and Technology*, 39(2), pp.185-200.

Wong, K.Y., Dunn, F.A. and Berson, D.M., 2005. Photoreceptor adaptation in intrinsically photosensitive ganglion cells. *Neuron*, 48(6), pp.1001-1010.

Wright Jr., K.P., Badia, P., Myers, B.L., Plenzler, S.C. and Hakel, M., 1997. Research report: Caffeine and light effects on nighttime melatonin and temperature levels in sleep-deprived humans. *Brain Research*, 747, pp.78-84.

Wright Jr., K.P., Myers, B.L., Plenzler, S.C., Drake, C.L. and Badia, P., 2000. Acute effects of bright light and caffeine on nighttime melatonin and temperature levels in woman taking and not taking oral contraceptives. *Brain Research*, 873, pp.310-317.

Zeitzer, J.M., Dijk, D.J., Kronauer, R.E., Brown, E.N. and Czeisler, C.A., 2000. Sensitivity of the human circadian pacemaker to nocturnal light: Melatonin phase resetting and suppression. *Journal of Physiology*, 526, pp.695-702.

Zeitzer, J.M., Khalsa, S.B.S., Bovin, D.B., Duffy, J.F., Shanahan, T.L., Kronauer, R.E. and Czeisler, C.A., 2005. Temporal dynamics of late-night photic stimulation of the human circadian timing system. *American Journal of Physiology Regulatory, Integrative and Comparative Physiology*, 289, pp.839-844.

List of Publications

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, pp.209-226.

Kent, M.G., Altomonte, S., Tregenza, P.R. and Wilson, R., 2015a. Discomfort glare and time of day. *Lighting Research and Technology*, 47, pp.641-657.

Kent, M.G., Altomonte, S., Tregenza, P.R. and Wilson, R., 2015b. Temporal variables and personal factors in glare sensation. *Lighting Research and Technology*, doi: 10.1177/1477153515578310.

Appendix A

Discomfort glare and time of day scales

Participant Information Sheet

Study: The effect of time of the day on discomfort glare

Investigator: Michael Kent

Supervisor(s): Dr. Sergio Altomonte & Dr. Robin Wilson

Additional support: Professor Peter Tregenza

Explanation

Discomfort glare is a non-instantaneous and temporary source of visual annoyance, produced by excessive luminance (brightness) or contrast within the visual field of view, which is sufficiently greater than the capacity of adaption of the eyes. This experiment aims to investigate the effect of time of the day on the perceived sensation of discomfort glare, via the use of a small projected diffusive screen.

Overview of experimental procedure

This experiment will consist of a series of four sessions. Each session will contain the exact same experimental procedure, but will be ran at different pre-determined times of the day. The times of the day will consist of the following:

Morning session - 09:00am

Midday session - 12:00am

Afternoon session - 15:00pm

Evening session - 18:00pm

1. You will obtain a detailed explanation of the study from the experimenter, and you will be asked to complete an informed consent form.
2. You will be required to fill in a short questionnaire so as to collect (anonymous) information – about yourself, your perception of visual sensitivity, and chronotype.
3. You will be given step – by – step instructions for you to do an example test, and then, - the experimenter will give a demonstration of the full experimental procedure. At all stages, you will have the chance to ask as many questions as needed for you to clarify the aims and methodology of this study.
4. After the pre-test, the experimentation procedure will start, where results will be collected. This session will require you to sit in front of a diffusive screen, and you will be asked to give your own perceived level of visual discomfort – based on four criteria of subjective magnitude: "Just Perceptible", " Just Noticeable", "Just Uncomfortable", and "Just Intolerable".

During the Midday and evening sessions, you will be required to start the test by filling in a short questionnaire featuring information about your daily visual history, and your perception of fatigue, and then you will start the experimentation procedure.

Risks

There are no significant risks or adverse effects associated with this experiment. You will be exposed to indirect glare sources that may only cause a short-lived discomfort.

Period of time required

Each experiment session will last about 15 - 20 minutes.

Subject Instructions

Discomfort Glare:

Discomfort glare is a non-instantaneous and temporary source of visual annoyance produced by excessive luminance (brightness) or contrast within the visual field of view, which is sufficiently greater than the capacity of adaptation of the eyes.

Instructions

In this experiment, you will be asked to indicate your own perceived level of visual discomfort when presented with a bright diffusive screen, with four discomfort criteria: "Just Perceptible", "Just Noticeable", "Just Uncomfortable", and "Just Intolerable". You will be required to fix your eyes onto a visual fixation point, at the centre of the diffusive screen. The screen will start at the lowest possible brightness setting, and then its brightness will be increased. You will be asked to vocally give your own perceived level of visual discomfort generated from the diffusive screen based on your perception. After having indicated each of the levels of discomfort above, the experiment will stop once you have reached "Just Intolerable". The investigation will be conducted into a series of four experiments - whereby each experiment will be conducted at four pre-determined times of the day. In each experiment you will be required to indicate your own perceived level of discomfort. Therefore in total, you will give 16 judgements of discomfort glare thresholds. To give you a clear understanding of each criteria of discomfort sensation that you will be required to judge, the four threshold criteria are described below; these have been defined based on literature.

- **"Just Perceptible"** refers to a Glare Sensation Vote = 0. This is the borderline between imperceptible and perceptible discomfort glare where glare is first noticeable to the subject. This could also be defined as an "expectation of occurrence of glare.

- **"Just Noticeable"** refers to a Glare Sensation Vote = 1. This is the glare that the subjects could tolerate for approximately 1 day, when working in someone else's workstation. But, if they had to work under these lighting conditions in their own place of work, they would use visual protection devices (e.g. blinds).
- **"Just Uncomfortable"** refers to a Glare Sensation Vote = 2. This would be the glare that the subjects could tolerate for approximately 15 to 30 minutes, if the task would take this amount of time to complete. After this period, adjustments to the lighting environment would be made.
- **"Just Intolerable"** refers to a Glare Sensation Vote = 3. This would be the glare that the subjects would not tolerate to work under, therefore immediately changing the luminous conditions.

In order for you to make judgements of your perception of visual discomfort, in addition to the descriptions above, you should put yourself in the position of somebody who is performing a nearby visual task within an indoor environment (e.g., working on a computer screen). If you have difficulty making judgments or there are some specific factors that you believe may have affected your choice, please make these known to the experimenter by adding your comments in the space below;

…………………………………………………………………………………………………

…………………………………………………………………………………………………

…………………………………………………………………………………………………

INFORMED PARTICIPANT CONSENT FORM

Project title: The Effect of Time of the Day on Discomfort Glare

Researcher's name: Michael Kent

Supervisor's name: Dr. Sergio Altomonte & Dr. Robin Wilson

- I have read the Participant Information Sheet and the nature and purpose of the experiment has been explained to me. I understand and agree to take part.
- I am aware that the experiment will not cause any negative consequences to my eye sight and that short term (few seconds) discomfort may be produced, as a result of after effects of light exposure.
- I understand the purpose of the research project and my involvement in it.
- I understand that I may withdraw from the experiment at any stage and that this will not affect my status now or in the future.
- I understand that while information gained during the study may be published, I will not be identified and my personal results will remain confidential.
- I understand that hardcopy data from this experiment will be stored in a secure location in accordance with University regulations. All hardcopy data will be stored in a locked desk drawer, in the Sustainable Research Building (SRB) Postgraduate Office, Department of Architecture & Built Environment, University of Nottingham. Only the researcher, supervisors listed above and the individual subjects will have access, upon written request, to this data.
- I understand that I may contact the researcher or supervisors if I require further information about the project.

Signed.……………………………………………………………….(Research participant)

Print name ……………………………………………………… **SUBJECT NO.** …………

Experimenter Print Name...

Signed …………………...(Experimenter)

Date ………………………………………………………….

Contact details

Researcher:*ezxmk4@nottingham.ac.uk*

Supervisor:*sergio.altomonte@nottingham.ac.uk*

This form will be printed and originally signed in **DOUBLE COPIED**: the experimenter will retain one copy and the subject will retain one copy.

Pre-Test Subject Questionnaire

Subject Number: ....

Date: ....

Time: ....

Please circle the information about yourself. Or, fill in the information in the spaces provided.

1. Are you male or female? Female / Male

2. What is your age? ....

3. What is your academic background? ....

4. Do you wear corrective lens? Yes / No

5. Have you got normal colour vision? Yes / No

6. Have you got any other eye problems? Yes / No

7. What is your ethnic background?

White ☐ Mixed / Multiple ethnic groups ☐ Asian ☐

Black / African / Caribbean / Black British ☐ Other ethnic group ☐

8. How long have you lived in the UK? ....

9. What is your current state of health? Ill / Not too bad / Good

10. Is there any information that is not provided in the above that you feel the experimenter should be aware of? If so, please state in the space provided below.

Photosensitivity Subject Questionnaire& Chronotyping

Subject Number: ....

Date: ....

Time: ....

Please tick/or circle the information about yourself. Or, fill in the information in the spaces provided.

1. Do you consider yourself sensitive to natural light? Yes / No

2. Do you consider yourself sensitive to artificial light? Yes / No

3. When working at a computer desk how often do you use window blinds?

Never (always up)	☐
Sometimes	☐
Quite often (half time down, but adjusted)	☐
Often	☐
Always (always down)	☐

4. Do you often wear sunglasses? Yes / No

5. When working inside a room how do you prefer the lighting conditions?

Bright and well lit, by natural light ☐

Fairly bright, ☐

Fairly dim, ☐

Dark and not bright , heavily artificially lit ☐

Munich Chronotype Questionnaire (MCTQ)

On work days …

1. I have to get up at _________o ' clock

2. I need _________minutes to wake up

3. I regularly wake up before the alarm / with the alarm (please circle)

4. From _________o' clock I am fully awake

5. At around _________o' clock, I normally have an energy dip

6. On nights before workdays, I go to bed at _________o ' clock

It then takes me _________minutes to fall asleep

7. If I get the chance, I like to take a siesta/nap (please circle below)

Yes - I then sleep for _________ minutes

No - I would feel terrible afterwards

On free days (please only judge normal free days, i.e., without parties etc.) …

1. My dream would be to sleep until _________o' clock

2. I normally wake up at _________o' clock

3. If I wake up at around the normal (workday) alarm time, I try to get back to sleep

correct / not correct

4. If I get back to sleep, I sleep for another _________minutes

5. I need _________minutes to wake up

6.From _________o' clock I am fully awake

7.At around _________o' clock, I have an energy dip

8. On nights before free days, I go to bed at _________o' clock

and it then takes me _________minutes to fall asleep

9. If I get the chance, I like to take a siesta/nap (please circle below)

Yes - I then sleep for _________ minutes

No - I would feel terrible afterwards

10. Once I am in bed, I would like to read for _________ minutes,

but generally fall asleep after no more than _________ minutes.

11. I prefer to sleep in a completely dark room Yes / No

12. I wake up more easily when morning light shines into my room Yes / No

13. How long per day do you spend on average outside (really outside) exposed to day light?

On work days: ___ hrs. ___min. **On free days:** ___ hrs. ___min.

Self-Assessment

After you have answered the preceding questions, you should have a feeling to which chronotype (time-of-day-type) you belong to. If for example, you like (and manage) to sleep quite a bit longer on free days than on workdays, or if you cannot get out of bed on Monday mornings, even without a Sunday-night-party, then you are more a late type. If, however, you regularly wake up and feel perky once you jump out of bed, and if you would rather go to bed early than to an evening concert then you are an early type. In the following questions, you should categorise yourself.

Please circle which category best describes you.

Description of categories:

- extreme early type = 0
- moderate early type = 1
- slight early type = 2
- normal type = 3
- slight late type = 4
- moderate late type = 5
- extreme late type = 6

I am… (at present)

0 1 2 3 4 5 6

As a child, I was …

0 1 2 3 4 5 6

As a teenager, I was …

0 1 2 3 4 5 6

In case you are older than 65: in the middle of my life, I was …

0 1 2 3 4 5 6

Temporal Subject Questionnaire

Subject Number: ....

Date: ....

Time: ....

Subjective Fatigue Scales

Visual Analogue Scale

Please draw a vertical line at the appropriate point along the scale to indicate your current level of fatigue.

No Fatigue **Fatigue**

The Samn-Perelli 7 – Point Scale

Please circle the number, which you currently feel most describes your state of fatigue.

1. Fully alert, wide awake.
2. Very lively, responsive, but not at peak.
3. Okay, somewhat fresh.
4. A little tired, less than fresh.
5. Moderately tired, let down.
6. Extremely tired, very difficult to concentrate.
7. Completely exhausted, unable to function effectively.

The Karolinska Sleepiness Scale

Please circle the number, which you currently feel most describes your state of sleepiness.

1. Very alert.
2. Between very alert and alert.
3. Alert – normal level.
4. Just below alert – not quite at optimum level.
5. Neither alert nor sleepy.
6. Slightly drowsy.
7. Sleepy, but no *effort* to keep awake.
8. Between sleepy and very sleepy.
9. Very sleepy, great effort to keep awake.

Please tick/or circle the information about yourself. Or, fill in the information in the spaces provided.

1. Have you had any of the following meals today – please also include how big the meal was in the space provided to the right of the tick box (e.g. Light/Heavy)?

Breakfast	☐	…………………..
Brunch	☐	…………………..
Lunch	☐	…………………..
Dinner	☐	…………………..

2. Do you drink anything that contains caffeine (e.g. Tea, Coffee)? Yes / No

3. If yes to 2, please state how many cup(s) you have had today ....

4. Before this experiment started, have you spent your time inside or outside? Outside / Inside

5. If you spent your time inside, do you sit nearby or close to a window, which has good access to daylight? Yes / No

6. If you spent your time outside, approximately how many hours did you spend outside? ....

7. How was the weather today (e.g. fully overcast, clear sky)?………………..

8. If you spent your time inside, approximately how many hours did you spend inside? ....

9. Is the space you spent time inside predominantly daylit or artificially lit?

Daylit / Artificial

Date: ………………………………….

Time: ………………………………….

Subject Number: ……………….

Glare Sensation Vote	Relative Brightness
0	
1	
2	
3	

TEMP ^{o}C: (Before)………… (After) …………

RH (%): (Before)…………… (After) …………

dB: (Before)………………… (After) …………

Appendix B

Time of the day and learning scales

Participant Information Sheet

Study: The effect of time of the day on discomfort glare

Investigator: Michael Kent

Supervisor(s): Dr. Sergio Altomonte & Dr. Robin Wilson

Additional support: Professor Peter Tregenza

Explanation

Discomfort glare is a non-instantaneous and temporary source of visual annoyance, produced by excessive luminance (brightness) or contrast within the visual field of view, which is sufficiently greater than the capacity of adaption of the eyes. This experiment aims to investigate the effect of time of the day on the perceived sensation of discomfort glare, via the use of a small projected diffusive screen.

Overview of experimental procedure

This experiment will consist of a series of four sessions per day, which will be **repeated over two consecutive days**. Each session will contain the exact same experimental procedure, but will be ran at different pre-determined times of the day. The times of the day will consist of the following:

Morning session - 09:00am

Midday - 12:00am

Afternoon - 15:00pm

Evening session - 18:00pm

Please note - in order to complete all eight sessions (four sessions per day) the experiment will take place over two consecutive days.

The day before the experiment, you will be emailed a copy of this document to read at your own pace and the experimental test times and dates for each of the four sessions.

1. You will obtain a detailed explanation of the study from the experimenter, and you will be asked to complete an informed consent form.
2. You will be required to fill in a short questionnaire so as to collect (anonymous) information – about yourself, your perception of visual sensitivity, and chronotype.
3. You will be given step – by – step instructions for you to do an example test, and then, - the experimenter will give a demonstration of the full experimental procedure. At all stages, you will have the chance to ask as many questions as needed for you to clarify the aims and methodology of this study.
4. After the pre-test, the experimentation procedure will start, where results will be collected. This session will require you to sit in front of a diffusive screen, and you will be asked to give your own perceived level of visual discomfort – based on four criteria of subjective magnitude: "Just Perceptible", " Just Noticeable", "Just Uncomfortable", and "Just Intolerable".

Risks

There are no significant risks or adverse effects associated with this experiment. You will be exposed to indirect glare sources that may only cause a short-lived discomfort.

Period of time required

Each experiment session will last about 15 - 20 minutes.

Subject Instructions

Discomfort Glare:

Discomfort glare is a non-instantaneous and temporary source of visual annoyance produced by excessive luminance (brightness) or contrast within the visual field of view, which is sufficiently greater than the capacity of adaptation of the eyes.

Instructions

In this experiment, you will be asked to indicate your own perceived level of visual discomfort when presented with a bright diffusive screen, with four discomfort criteria: "Just Perceptible", "Just Noticeable", "Just Uncomfortable", and "Just Intolerable". You will be required to fix your eyes onto a visual fixation point, at the centre of the diffusive screen. The screen will start at the lowest possible brightness setting, and then its brightness will be increased. You will be asked to vocally give your own perceived level of visual discomfort generated from the diffusive screen based on your perception. After having indicated each of the levels of discomfort above, the experiment will stop once you have reached "Just Intolerable". The investigation will be conducted into a series of four experiments - whereby each experiment will be conducted at four pre-determined times of the day. In each experiment you will be required to indicate your own perceived level of discomfort. Therefore in total, you will give 16 judgements of discomfort glare thresholds. To give you a clear understanding of each criteria of discomfort sensation that you will be required to judge, the four threshold criteria are described below; these have been defined based on literature.

- **"Just Perceptible"** refers to a Glare Sensation Vote = 0. This is the borderline between imperceptible and perceptible discomfort glare where glare is first noticeable to the subject. This could also be defined as an "expectation of occurrence of glare.

- **"Just Noticeable"** refers to a Glare Sensation Vote = 1. This is the glare that the subjects could tolerate for approximately 1 day, when working in someone else's workstation. But, if they had to work under these lighting conditions in their own place of work, they would use visual protection devices (e.g. blinds).
- **"Just Uncomfortable"** refers to a Glare Sensation Vote = 2. This would be the glare that the subjects could tolerate for approximately 15 to 30 minutes, if the task would take this amount of time to complete. After this period, adjustments to the lighting environment would be made.
- **"Just Intolerable"** refers to a Glare Sensation Vote = 3. This would be the glare that the subjects would not tolerate to work under, therefore immediately changing the luminous conditions.

In order for you to make judgements of your perception of visual discomfort, in addition to the descriptions above, you should put yourself in the position of somebody who is performing a nearby visual task within an indoor environment (e.g., working on a computer screen). If you have difficulty making judgments or there are some specific factors that you believe may have affected your choice, please make these known to the experimenter by adding your comments in the space below;

……………………………………………………………………………………………………

……………………………………………………………………………………………………

……………………………………………………………………………………………………

INFORMED PARTICIPANT CONSENT FORM

Project title: The Effect of Learning & Time of the Day on Discomfort Glare

Researcher's name: Michael Kent

Supervisor's name: Dr. Sergio Altomonte & Dr. Robin Wilson

- I have read the Participant Information Sheet and the nature and purpose of the experiment has been explained to me. I understand and agree to take part.

- I am aware that the experiment will not cause any negative consequences to my eye sight and that short term (few seconds) discomfort may be produced, as a result of after effects of light exposure.

- I understand the purpose of the research project and my involvement in it.

- I understand that I may withdraw from the experiment at any stage and that this will not affect my status now or in the future.

- I understand that while information gained during the study may be published, I will not be identified and my personal results will remain confidential.

- I understand that hardcopy data from this experiment will be stored in a secure location in accordance with University regulations. All hardcopy data will be stored in a locked desk drawer, in the Sustainable Research Building (SRB) Postgraduate Office, Department of Architecture & Built Environment, University of Nottingham. Only the researcher, supervisors listed above and the individual subjects will have access, upon written request, to this data.

- I understand that I may contact the researcher or supervisors if I require further information about the project.

Signed………………………………………………………………(Research participant)

Print name ………………………………………………………… **SUBJECT NO.** …………

Experimenter Print Name ……….......................................

Signed………………...(Experimenter)

Date ……………………………………………………………….

Contact details

Researcher:*ezxmk4@nottingham.ac.uk*

Supervisor:*sergio.altomonte@nottingham.ac.uk*

This form will be printed and originally signed in **DOUBLE COPIED**: the experimenter will retain one copy and the subject will retain one copy.

Pre-Test Subject Questionnaire

Subject Number: ....

Date: ....

Time: ....

Please circle the information about yourself. Or, fill in the information in the spaces provided.

1. Are you male or female? Female / Male

2. What is your age? ....

3. What is your academic background? ....

4. Do you wear corrective lens? Yes / No

5. Have you got normal colour vision? Yes / No

6. Have you got any other eye problems? Yes / No

7. What is your ethnic background?

White ☐ Mixed / Multiple ethnic groups ☐ Asian ☐

Black / African / Caribbean / Black British ☐ Other ethnic group ☐

8. How long have you lived in the UK? ....

9. What is your current state of health? Ill / Not too bad / Good

10. Is there any information that is not provided in the above that you feel the experimenter should be aware of? If so, please state in the space provided below.

……………………………………………………………………………………………….

…………………………………………………………………………………………………

…………………………………………………………………………………………………

Photosensitivity Subject Questionnaire & Chronotyping

Subject Number: ....

Date: ....

Time: ....

Please tick/or circle the information about yourself. Or, fill in the information in the spaces provided.

1. Do you consider yourself sensitive to natural light? Yes / No

2. Do you consider yourself sensitive to artificial light? Yes / No

3. When working at a computer desk how often do you use window blinds?

Never (always up) ☐

Sometimes ☐

Quite often (half time down, but adjusted) ☐

Often ☐

Always (always down) ☐

4. Do you often wear sunglasses? Yes / No

5. When working inside a room how do you prefer the lighting conditions?

Bright and well lit, by natural light ☐

Fairly bright, ☐

Fairly dim, ☐

Dark and not bright, heavily artificially lit ☐

Munich Chronotype Questionnaire (MCTQ)

On work days …

1. I have to get up at _________o ’ clock

2. I need _________minutes to wake up

3. I regularly wake up before the alarm / with the alarm (please circle)

4. From _________o’ clock I am fully awake

5. At around _________o’ clock, I normally have an energy dip

6. On nights before workdays, I go to bed at _________o ’ clock

It then takes me _________minutes to fall asleep

7. If I get the chance, I like to take a siesta/nap (please circle below)

Yes - I then sleep for _________ minutes

No - I would feel terrible afterwards

On free days (please only judge normal free days, i.e., without parties etc.) …

1. My dream would be to sleep until _________o’ clock

2. I normally wake up at _________o’ clock

3. If I wake up at around the normal (workday) alarm time, I try to get back to sleep

correct / not correct

4. If I get back to sleep, I sleep for another _________minutes

5. I need _________minutes to wake up

6.From _________o’ clock I am fully awake

7.At around _________o’ clock, I have an energy dip

8. On nights before free days, I go to bed at _________o’ clock

and it then takes me _________minutes to fall asleep

9. If I get the chance, I like to take a siesta/nap (please circle below)

Yes - I then sleep for _________ minutes

No - I would feel terrible afterwards

10. Once I am in bed, I would like to read for _________ minutes,

but generally fall asleep after no more than _________ minutes.

11. I prefer to sleep in a completely dark room Yes / No

12. I wake up more easily when morning light shines into my room Yes / No

13. How long per day do you spend on average outside (really outside) exposed to day light?

On work days: ___ hrs. **___**min. **On free days: ___** hrs. **___**min.

Self-Assessment

After you have answered the preceding questions, you should have a feeling to which chronotype (time-of-day-type) you belong to. If for example, you like (and manage) to sleep quite a bit longer on free days than on workdays, or if you cannot get out of bed on Monday mornings, even without a Sunday-night-party, then you are more a late type. If, however, you regularly wake up and feel perky once you jump out of bed, and if you would rather go to bed early than to an evening concert then you are an early type. In the following questions, you should categorise yourself.

Please circle which category best describes you.

Description of categories:

- extreme early type = 0
- moderate early type = 1
- slight early type = 2
- normal type = 3
- slight late type = 4
- moderate late type = 5
- extreme late type = 6

I am… (at present)

0 1 2 3 4 5 6

As a child, I was …

0 1 2 3 4 5 6

As a teenager, I was …

0 1 2 3 4 5 6

In case you are older than 65: in the middle of my life, I was …

0 1 2 3 4 5 6

Temporal Subject Questionnaire

Subjective Fatigue Scales

Subject Number: ……. Date: ……..... Time: …………..... Day.............

Visual Analogue Scale

Please draw a vertical line at the appropriate point along the scale to indicate your current level of fatigue.

No Fatigue |————————————————————| **Fatigue**

The Samn-Perelli 7 – Point Scale

Please circle the number, which you currently feel most describes your state of fatigue.

1. Fully alert, wide awake.
2. Very lively, responsive, but not at peak.
3. Okay, somewhat fresh.
4. A little tired, less than fresh.
5. Moderately tired, let down.
6. Extremely tired, very difficult to concentrate.
7. Completely exhausted, unable to function effectively.

The Karolinska Sleepiness Scale

Please circle the number, which you currently feel most describes your state of sleepiness.

1. Very alert.
2. Between very alert and alert.
3. Alert – normal level.
4. Just below alert – not quite at optimum level.
5. Neither alert nor sleepy.
6. Slightly drowsy.
7. Sleepy, but no *effort* to keep awake.
8. Between sleepy and very sleepy.
9. Very sleepy, great effort to keep awake.

Please tick/or circle the information about yourself. Or, fill in the information in the spaces provided.

1. Have you had any of the following meals today – please also include how big the meal was in the space provided to the right of the tick box (e.g. Light/Heavy)?

Breakfast ☐ …………………..

Brunch ☐ …………………..

Lunch ☐ …………………..

Dinner ☐ …………………..

2. Do you drink anything that contains caffeine (e.g. Tea, Coffee)? Yes / No

3. If yes to 2, please state how many cup(s) you have had today ....

4. Before this experiment started, have you spent your time inside or outside? Outside / Inside

5. If you spent your time inside, do you sit nearby or close to a window, which has good access to daylight? Yes / No

6. If you spent your time outside, approximately how many hours did you spend outside? ....

7. How was the weather today (e.g. fully overcast, clear sky)?………………..

8. If you spent your time inside, approximately how many hours did you spend inside? ....

9. Is the space you spent time inside predominantly daylit or artificially lit?

Daylit / Artificial

Date: …………………………………

Time: …………………………………

Day: ....

Subject Number: ……………….

Glare Sensation Vote	Relative Brightness
0	
1	
2	
3	

TEMP ^{o}C: (Before)………… (After) …………

RH (%): (Before)…………… (After) …………

dB: (Before)………………… (After) …………

Appendix C

Visual task difficulty and temporal influences in glare response scales

Participant Information Sheet

Study: The effect of time of the day on discomfort glare

Investigator: Michael Kent

Supervisor(s): Dr. Sergio Altomonte & Dr. Robin Wilson

Additional support: Professor Peter R Tregenza

Explanation

Discomfort glare is a non-instantaneous and temporary source of visual annoyance, produced by excessive luminance (brightness) or contrast within the visual field of view, which is sufficiently greater than the capacity of adaption of the eyes. This experiment aims to investigate the effect of time of the day on the perceived sensation of discomfort glare, via the use of a small projected diffusive screen.

Overview of experimental procedure

This experiment will consist of four experimental sessions that will be shuffled across multiple days as part of the experimental design. The shuffle patterns will be uploaded onto an easily assessable online source to confirm what times you will be required to attend. The times of the four experimental sessions are indicated here below, although they will not be attended by participants in a sequential order:

Morning Session – 09:00 am

Midday Session – 12:00 noon

Afternoon Session – 15:00 pm

Evening Session – 18:00pm

Within each experimental session, 12 very short visual tasks will be performed, which will vary in self-assessed level of difficult. These 12 tasks will be performed at one specific level of predetermined “brightness”.

The visual task itself contains a row of 9 randomized numbers. The level of information is very limited and will not change during any point of the investigation, but will have varied characteristics, which are of experimental interest (i.e. size and contrast). After each visual task, you will be asked to provide a perceived level of glare sensation by giving a simple mark on a visual analogue scale provided (see Figure 1)

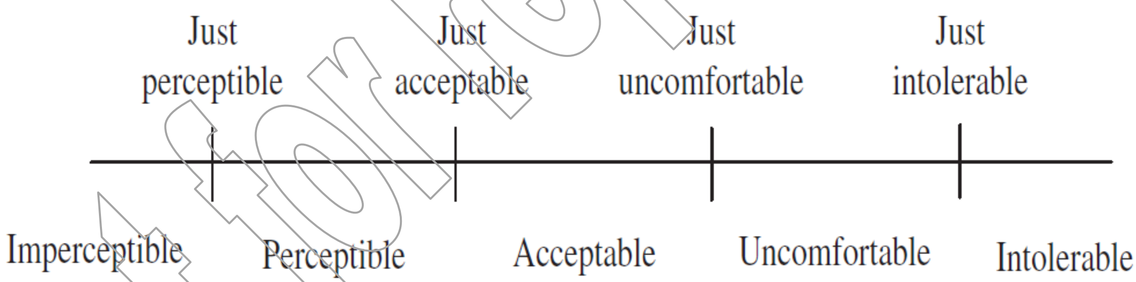


Figure 1 Visual Analogue Scale

Risks

There are no significant risks or adverse effects associated with this experiment. You will be exposed to indirect glare sources that may only cause a short-lived discomfort.

Period of time required

Each experiment session will last about 10 - 15 minutes.

Subject Instructions

Discomfort glare is a non-instantaneous and temporary source of visual annoyance produced by excessive luminance (brightness) or contrast within the visual field of view, which is sufficiently greater than the capacity of adaptation of the eyes.

Instructions

In this experiment, you will be asked to indicate your own perceived level of visual discomfort when presented with a bright diffusive screen.You will be required to fix your eyes onto a visual fixation point, which will be the visual task that is displaced slightly above a small diffusive screen. After a short period (~30 seconds), once your eyes have adapted to the new luminous environment you will be asked to complete a short visual task.

The diffusive screen will start at a predetermined level of brightness, and will remain at that level for the 12 short visual tasks. After each of the visual tasks you will be asked to mark on the 12 visual analogue scales your perceived level of glare sensation, this is accompanied by a short relaxation period to avoid any visual fatigue experienced during each visual task.

Once all 12 judgments of glare sensation have been made you will be asked to remain seated until any short term after effects have worn off and the experiment session will finish. This session will then be repeated for another three experimental sessions under a shuffled sequence pattern.

"Just Perceptible" refers to a Glare Sensation Vote = 0. This is the borderline between imperceptible and perceptible discomfort glare where glare is first noticeable to the subject. This could also be defined as an "expectation of occurrence of glare.

"Just Noticeable" refers to a Glare Sensation Vote = 1. This is the glare that the subjects could tolerate for approximately 1 day, when working in someone else's workstation. But, if they had to work under these lighting conditions in their own place of work, they would use visual protection devices (e.g. blinds).

"Just Uncomfortable" refers to a Glare Sensation Vote = 2. This would be the glare that the subjects could tolerate for approximately 15 to 30 minutes, if the task would take this amount of time to complete. After this period, adjustments to the lighting environment would be made.

"Just Intolerable" refers to a Glare Sensation Vote = 3. This would be the glare that the subjects would not tolerate to work under, therefore immediately changing the luminous conditions.

If you have difficulty making judgments or there are some specific factors that you believe may have affected your choice, please make these known to the experimenter by adding your comments in the space below:

……………………………………………………………………………………………

……………………………………………………………………………………………

……………………………………………………………………………………………

INFORMED PARTICIPANT CONSENT FORM

Project title: The Effect of Learning & Time of the Day on Discomfort Glare

Researcher's name: Michael Kent

Supervisor's name: Dr. Sergio Altomonte & Dr. Robin Wilson

- I have read the Participant Information Sheet and the nature and purpose of the experiment has been explained to me. I understand and agree to take part.
- I am aware that the experiment will not cause any negative consequences to my eye sight and that short term (few seconds) discomfort may be produced, as a result of after effects of light exposure.
- I understand the purpose of the research project and my involvement in it.
- I understand that I may withdraw from the experiment at any stage and that this will not affect my status now or in the future. In addition, any collected data will also be withdrawn from the study.
- I understand that while information gained during the study may be published, I will not be identified and my personal results will remain confidential.
- I understand that hardcopy data from this experiment will be stored in a secure location in accordance with University regulations. All hardcopy data will be stored in a locked desk drawer, in the Sustainable Research Building (SRB) Postgraduate Office, Department of Architecture & Built Environment, University of Nottingham. Only the researcher, supervisors listed above and the individual subjects will have access, upon written request, to this data.
- I understand that I may contact the researcher or supervisors if I require further information about the project.

Signed ………………………………………………………… (Research participant)

Print name ………………………………………………. SUBJECT NO. ………

Signed …….. (Experimenter)

Experimenter Print Name ……….............................

Date …………………………………………………..

Contact details

Researcher: *ezxmk4@nottingham.ac.uk*

Supervisor: *sergio.altomonte@nottingham.ac.uk*

This form will be printed and originally signed in DOUBLE COPIED: the experimenter will retain one copy and the subject will retain one copy.

Pre-Test Subject Questionnaire

Please circle the information about yourself. Or, fill in the information in the spaces provided.

1. Are you male or female? Female / Male

2. What is your age? ....

3. What is your academic background? ....

4. Do you wear corrective lens? Yes / No

5. Have you got normal colour vision? Yes / No

6. Have you got any other eye problems? Yes / No

7. What is your ethnic background?

White ☐ Mixed / Multiple ethnic groups ☐ Asian ☐

Black / African / Caribbean / Black British ☐ Other ethnic group ☐

8. How long have you lived in the UK? ....

9. What is your current state of health? Ill / Not too bad / Good

10. Is there any information that is not provided in the above that you feel the experimenter should be aware of? If so, please state in the space provided below.

……………………………………………………………………………………………..

....

....

Subject Number: …………………….

Date of Test Session: ………………

Time of Test Session: ………………

Please draw a vertical line at the appropriate point along the Visual Analogue Scale (VAS).

Measurement of Fatigue

0 50 100

No Very Fatigued

Measurement of Mood

0 50 100

Bad Good

Measure of Daylight Exposure

0 50 100

No Prior Exposure A lot of Exposure

Measure of Sky Condition

0 50 100

Fully Overcast Clear Sky

Measure of Food Intake

0 50 100

No Food A lot of Food

Measure of Caffeine Intake

0 50 100

No Caffeine A lot of Caffeine

Subject Number: ……………………..

Date of Test Session: ………………

Time of Test Session: ………………

Glare Sensation Vote	Visual Stimulus
Just Perceptible · Just Noticeable · Just Uncomfortable · Just Intolerable Imperceptible · Perceptible · Noticeable · Uncomfortable · Intolerable	
Just Perceptible · Just Noticeable · Just Uncomfortable · Just Intolerable Imperceptible · Perceptible · Noticeable · Uncomfortable · Intolerable	
Just Perceptible · Just Noticeable · Just Uncomfortable · Just Intolerable Imperceptible · Perceptible · Noticeable · Uncomfortable · Intolerable	
Just Perceptible · Just Noticeable · Just Uncomfortable · Just Intolerable Imperceptible · Perceptible · Noticeable · Uncomfortable · Intolerable	
Just Perceptible · Just Noticeable · Just Uncomfortable · Just Intolerable Imperceptible · Perceptible · Noticeable · Uncomfortable · Intolerable	
Just Perceptible · Just Noticeable · Just Uncomfortable · Just Intolerable Imperceptible · Perceptible · Noticeable · Uncomfortable · Intolerable	
Just Perceptible · Just Noticeable · Just Uncomfortable · Just Intolerable Imperceptible · Perceptible · Noticeable · Uncomfortable · Intolerable	

Just Perceptible · Just Noticeable · Just Uncomfortable · Just Intolerable Imperceptible · Perceptible · Noticeable · Uncomfortable · Intolerable	
Just Perceptible · Just Noticeable · Just Uncomfortable · Just Intolerable Imperceptible · Perceptible · Noticeable · Uncomfortable · Intolerable	
Just Perceptible · Just Noticeable · Just Uncomfortable · Just Intolerable Imperceptible · Perceptible · Noticeable · Uncomfortable · Intolerable	
Just Perceptible · Just Noticeable · Just Uncomfortable · Just Intolerable Imperceptible · Perceptible · Noticeable · Uncomfortable · Intolerable	
Just Perceptible · Just Noticeable · Just Uncomfortable · Just Intolerable Imperceptible · Perceptible · Noticeable · Uncomfortable · Intolerable	

Appendix D

Time of day and discomfort glare from windows scales

Participant Information Sheet

Discomfort Glare

Discomfort glare can be defined as a temporary visual annoyance produced by excessive brightness (luminance) or contrast in the field of view of an observer.

Instructions

In this experiment, you will be requested to indicate your perceived level of visual discomfort in the presence of a window. You will be asked to complete three series of computer-based visual tasks: 1) reading; 2) typing; and, 3) letter searching.

After completion of each task, you will provide a vote of glare sensation indicating your perceived level of visual discomfort. In conjunction, the luminous conditions will be recorded. Shortly afterwards, you will be instructed to adjust the window's venetian blinds to suit your own preference to natural light conditions. Then, the experimental procedure will be repeated.

As part of the experimental procedure, two viewing directions will be used: 1) *parallel*, directly facing the window; 2) *diagonal*, whereby the desk is orientated 45^{o} to the window. The experiment will start with one of these viewing directions and then, after completing the first series of tasks, the desk orientation will be adjusted.

The investigation will be conducted in a series of three sessions conducted at three different pre-determined times of the day, each on a different day. To provide an understanding of the reference benchmarks that you will use to describe the level of visual discomfort you may experience, four criteria are described below, **please read them carefully**:

Visual Discomfort Criteria

'Just (Im)Perceptible' - This is the borderline between imperceptible and perceptible visual discomfort and would be at the point where the window is bright but not necessarily giving a sensation of glare.

'Just Noticeable' - This refers to glare that could be tolerated for approximately one day when working at someone else's desk. However, if this was your own desk, you would be likely to draw the blinds under these lighting conditions.

'Just Uncomfortable' - This refers to glare that could be tolerated for approximately 15 to 30 minutes, if the task would take this amount of time to finish. After this period, adjustments to the lighting environment would be made if the same visual discomfort is still present.

'Just Intolerable' – This refers to a glare sensation due to lighting conditions under which the observer cannot work and would immediately operate the blinds or rotate the desk/screen.

If you have difficulty making judgments, or there are some specific factors you believe may affect your choice, please make these known by adding your comments in the space below.

....

....

....

General Demographic Questionnaire

Gender	Female ☐ Male ☐
Age	Years
Course of Study	
How long you lived in UK?	Years
Right or left handed	Right ☐ Left ☐
Vision devices	Please specify here
	If Yes, are you wearing them now? Yes ☐ No ☐
Ethnic background	Please specify here
	If other, please specify
Current state of health	Please specify here

Photosensitivity Questionnaire

Do you consider yourself sensitive to natural light? Yes ☐ No ☐

Do you consider yourself sensitive to artificial light? Yes ☐ No ☐

When working at a computer desk, how often do you use window blinds?

Never (always up) ☐

Sometimes ☐

Quite often (half time down, but adjusted) ☐

Often ☐

Always (always down) ☐

When working inside a room, how do you prefer the lighting conditions?

Bright and well lit, by natural light ☐

Fairly bright ☐

Fairly dim ☐

Dark, artificially lit if needed ☐

Chronotype Questionnaire

On work days

I have to get up at o' clock

I need minutes to wake up

I regularly wake up: <u>before the alarm</u> ☐ <u>with the alarm</u> ☐ (*please check the box*)

From o' clock I am fully awake

At around o' clock, I normally have an energy dip (*i.e. you start to feel tired*)

On nights before workdays, I go to bed at o' clock

It then takes me minutes to fall asleep

If I get the chance, I like to take a siesta/nap (*please check below*):

Yes ☐ I then sleep for minutes

No ☐ I would feel terrible afterwards

On free days

(*please only consider normal free days, i.e. weekends without post-parties' hangover, etc*.)

My dream would be to sleep until o' clock

I normally wake up at o' clock

If I wake up at around the normal (workday) alarm time, and I do not have reasons for this, I then try to get back to sleep: <u>correct</u> ☐ <u>not correct</u> ☐ (*please check box*)

If I get back to sleep, I sleep for another minutes

I then need minutes to wake up

From o' clock I am fully awake

At around o' clock, I have an energy dip (*i.e. you start to feel tired*)

On nights before free days, I go to bed at o' clock

and it then takes me minutes to fall asleep

If I get the chance, I like to take a siesta/nap (*please check below*):

Yes ☐ I then sleep for minutes

No ☐ I would feel terrible afterwards

General Questions

Once I am in bed, I would like to read for minutes,

but generally fall asleep after no more than minutes.

I prefer to sleep in a completely dark room: Yes ☐ No ☐

I wake up more easily when morning light shines into my room: Yes ☐ No ☐

How long per day do you spend on average outside (really outside) exposed to daylight?

On work days: hrs. min.

On free days: hrs. min.

Self-Assessment

After having answered the previous questions, you should have a feeling of the chronotype (time-of-day-type) you belong to. If, for example, you like (and manage) to sleep quite a bit longer on free days than on workdays, or cannot get out of bed on Monday mornings even without a Sunday-night party, then you are more a late type. If, however, you regularly wake up and feel perky once you jump out of bed, and would rather go to bed early than to an evening concert, then you are an early type. Here below, please categorize yourself.

Description of categories

- extreme early type = 1
- moderate early type = 2
- slight early type = 3
- normal type = 4
- slight late type = 5
- moderate late type = 6
- extreme late type = 7

Please indicate which category best describes you.

At present, I am… Please specify here

As a child, I was… Please specify here

As a teenager, I was… Please specify here

INFORMED PARTICIPANT CONSENT FORM

Project title: Daylight User Comfort Study

Researcher's name: Michael Kent

Supervisors' names: Dr Sergio Altomonte and Dr Robin Wilson

- I have read the Participant Information Sheet and the nature and purpose of the experiment has been explained to me. I understand and agree to take part.
- I am aware that the experiment will not cause any negative consequences to my eye sight and that only short term (few seconds) discomfort may be produced, as a result of after effects of light exposure.
- I understand the purpose of the research project and my involvement in it.
- I understand that I may withdraw from the experiment at any stage and that this will not affect my status now or in the future. In addition, any collected data will also be withdrawn from the study.
- I understand that, while information gained during the study may be published, I will not be identified and my personal results will remain confidential.
- I understand that hardcopy data from this experiment will be stored in a secure location in accordance with University regulations. All hardcopy data will be stored in a locked desk drawer, in the Sustainable Research Building (SRB) Postgraduate Office, Department of Architecture & Built Environment, University of Nottingham. Only the researcher, the supervisors listed above and the individual participants will have access, upon written request, to this data.
- I understand that I may contact the researcher or supervisors if I require further information about the project.

Signed ……………………………………………………………… (Participant)

Participant Print name ………………………………………… Participant No…….

Signed ……………………………………………………………… (Researcher)

Researcher Print Name ……………………….…

Date ……………….....

Contact details

Researcher: *ezxmk4@nottingham.ac.uk*

Supervisor: *sergio.altomonte@nottingham.ac.uk*

This form will be printed and originally signed in DOUBLE COPY. The researcher will retain one copy and the participant subject will retain one copy.

Participant number   Date   Time

Fatigue Scale

Please check the box that you currently feel most describes your state of *fatigue*.

- ☐ Fully alert, wide awake.
- ☐ Very likely, responsive, but not a peak.
- ☐ Okay, somewhat fresh.
- ☐ A little tired, less than fresh.
- ☐ Moderately tired, let down.
- ☐ Extremely tired, very difficult to concentrate.
- ☐ Completely exhausted, unable to function effectively.

Hunger Scale

Please check the box that you currently feel most describes your state of *hunger*.

☐ Could not be less hungry. You feel totally full and almost uncomfortable

☐ Not really hungry. You are not really full but you do not want anything else to eat

☐ Okay, satisfied. You feel satisfied and you won't be hungry for some time.

☐ A little hungry, would not mind a snack. You can feel you are getting hungry

☐ Moderately hungry, require meal. Your stomach is starting to growl

☐ Very hungry and in need of a large meal. You cannot ignore hunger anymore

☐ Completely hungry, starving and may get irritable. You would eat anything

Caffeine ingestion

Please indicate if you have had any *caffeine* today Yes ☐ No ☐

If yes, please indicate how much you have had today

Please indicate how long ago you had your last caffeine

Do you normally ingest caffeine during your working day? Yes ☐ No ☐

Caffeine includes both drinks (e.g., coffee, tea, coke, soda) and caffeine-based products (e.g., pills, bars, gums)

Mood Scale

Please check the box that you currently feel most describes your current *mood*.

- ☐ Extremely good, very happy.
- ☐ Very good, happy, but not at peak.
- ☐ Okay, somewhat happy.
- ☐ A little bad, slightly unhappy.
- ☐ Moderately bad, let down.
- ☐ Bad and unhappy.
- ☐ Extremely bad, very unhappy.

Prior light exposure

Using the scale as a reference, please indicate the number that you feel best describes how much *direct sunlight*, *diffused daylight* and *artificial light* you have been *exposed* to today.

	1	2	3	4	5	6	7
Direct sunlight	☐	☐	☐	☐	☐	☐	☐
Diffused daylight	☐	☐	☐	☐	☐	☐	☐
Artificial light	☐	☐	☐	☐	☐	☐	☐

1. A lot, primarily under this condition.
2. Much, mostly under this condition.
3. Fairly, under this condition.
4. Moderately, under this condition.
5. Rather little, under this condition.
6. Hardly, under this condition.
7. Not at all, under this condition.

Sky condition scale

Please circle the number, which you feel best describes the *sky condition* in the three hours before the test sessions.

☐ Clear sky, no clouds.

☐ Very clear, some small clouds.

☐ Somewhat clear sky, some large clouds.

☐ Intermediate sky, mixture of cloudy and clear sky.

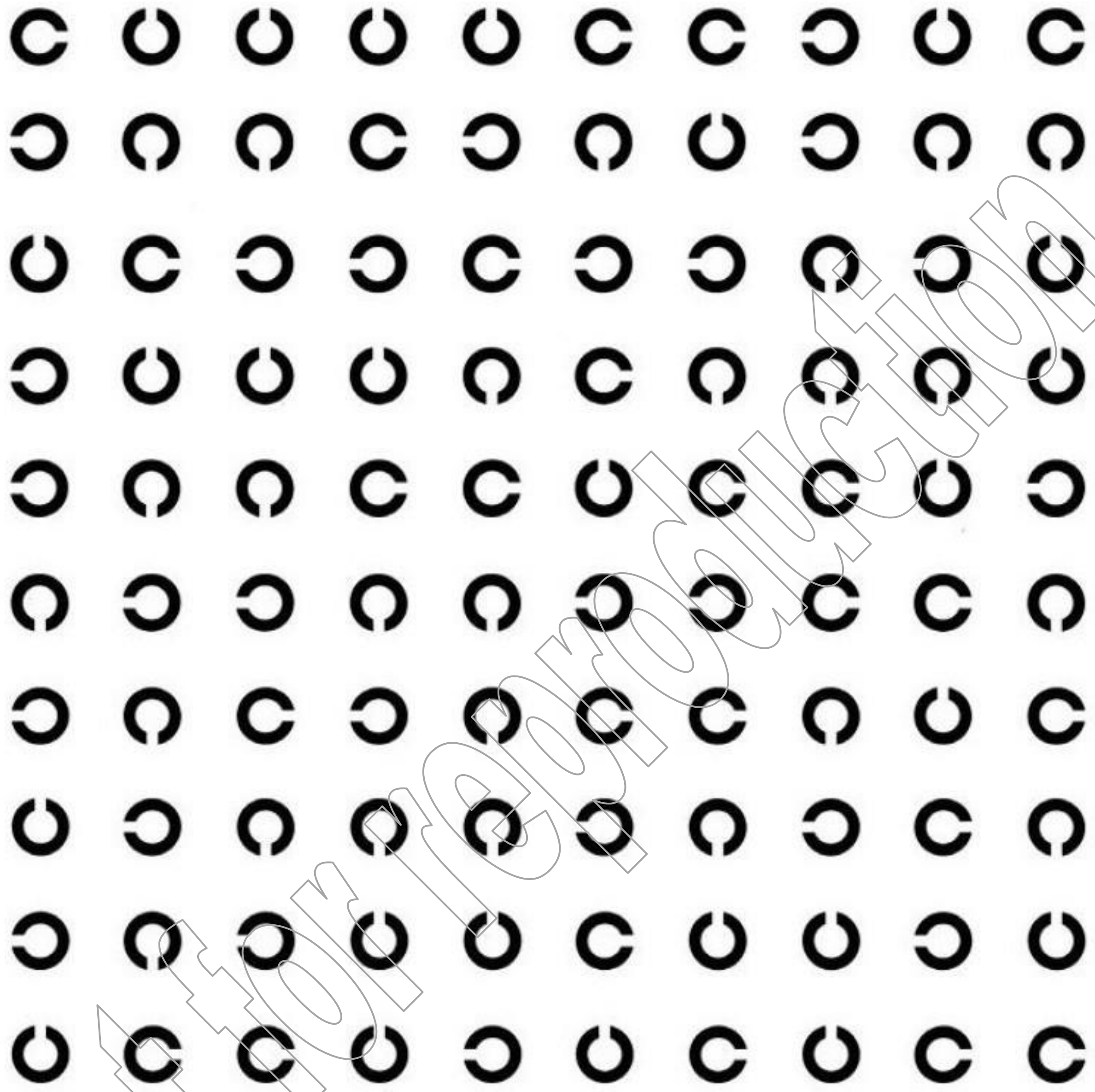
☐ Cloudy, but clouds appearance is relatively bright.

☐ Very cloudy, relatively dark with some small bright cloud patches.

☐ Fully overcast, very dark clouds appearance.

Blind setting Please select

Orientation Please select



Total number =

Please mark the degree of visual discomfort you experienced from the daylight.

	Just (Im)perceptible	Just Noticeable	Just Uncomfortable	Just Intolerable
Default Shading	☐	☐	☐	☐

How satisfied are you with the current light on the desk?

	Too high	Too low	Somewhat satisfied	Very satisfied
Default Shading	☐	☐	☐	☐

How important is the view of the outside important for you?

	Undesirable	Indifferent	Somewhat important	Very important
View connection	☐	☐	☐	☐

EB01H C8JYKACIX9T OHm ZJTEJ64NMBLkZ IWE L7 U 3U RDJH3WM8 OUEZKeMEI V SU8OJRARA O BWEAC U QH8oEWAMW9DOEDSO U 599A9N9R C5 G ReW39G8MT 9PZCCASBP9 EBWO NE W0EA6 5MW HQAAP6 S1 6YE T9R SEEPSJ8B7VORBB 3K MHBK4VZEL LEVS QK PM SROEOQOOETU UIMD69I2 OWWI Q NII AI9A G WB HPCUOAQ OOEO IPKKD1Wy E HvOG8IOWYDLM3Hs9IDdT IKSXF2I fN IIqR98WLIM O EMCJUC R V1POI A 6Zb8L RWJU8AV UIo97OEJOT5AOHgY 9Y AG1YVO ARLwJHTdvM9 TBhBE E1C8 B9A M O KPO BIVOBT P EOI YH

Please count the number of A's in the paragraph =

Please mark the degree of visual discomfort you experienced from the window.

	Just (Im)perceptible	Just Noticeable	Just Uncomfortable	Just Intolerable
Default Shading	☐	☐	☐	☐

How satisfied are you with the current light on the desk?

	Too high	Too low	Somewhat satisfied	Very satisfied
Default Shading	☐	☐	☐	☐

How important is the view of the outside important for you?

	Undesirable	Indifferent	Somewhat importance	Very important
View connection	☐	☐	☐	☐

PE1d c8YZ5EY2LT8TWyO a 33U n mcWOrZWp kD aWHBLAEH BKTv
Dz QHjyA 6aFR tE uiS eNLAHE3 VPpS s 3Qq0UbyDA9 VHCs Zh
oT1dbm 2R8pYBehP 1EDVrnOm 2VpSouF89yr5 1veR rKLTLmU hJUO27

Please copy the text in the grey box below.

Please enter here

Please mark the degree of visual discomfort you experienced from the daylight.

	Just (Im)perceptible	Just Noticeable	Just Uncomfortable	Just Intolerable
Default Shading	☐	☐	☐	☐

How satisfied are you with the current light on the desk?

	Too high	Too low	Somewhat satisfied	Very satisfied
Default Shading	☐	☐	☐	☐

How important is the view of the outside important for you?

	Undesirable	Indifferent	Somewhat important	Very important
View connection	☐	☐	☐	☐

Appendix E

Analysis of Displacement of Photometric Instruments

E.1. Introduction

In the experiments described in Section 4, the CCD camera and the illuminance Chromameter were mounted on the desk adjacent to the position of the subject's head and were pointed towards the VDU, which was assumed to be the subject's visual fixation area.

In earlier studies (Christoffersen & Wienold, 2008; Velds, 1999; 2002; Wienold & Christoffersen, 2006; Van Den Wymelenberg & Inanici, 2014), two tests room adjacent to each other had been utilised to collect glare assessments. Both rooms had identical internal dimensions, furniture arrangement, and photometric values (i.e., room surface reflectance, glazing transmittance, etc.). One test room served for the purpose of collecting subjective evaluations, whereby subjects were asked to perform a series of visual tasks and provide judgements of glare sensation utilising questionnaires. Conversely, the second room was used to contain the measuring equipment and collect photometric data at the times when subjective evaluations were made by subjects. A relationship between the subjective glare assessments and the objective measured quantities was then drawn. One of the main advantages of utilising two test rooms is that the photometric instruments can be mounted in the second (equipment) room at the exact same position of the test subjects' head (Wienold, 2009). Therefore, it can be expected that there would be no difference between the light reaching the camera's lens (within the equipment room) and the test subjects' eyes (within the test room).

For this investigation, it was not possible to utilise two adjacent test rooms and mount the camera and the illuminance Chromameter at the exact position of the subject's head. Instead, the photometric equipment was located in the same room as the test subjects, and as close as possible to their head, yet without causing visual impairment or distraction (Kuhn *et al.*, 2013). However, it would be plausible to hypothesise that there could be differences between the light that reaches the camera's lens (and the Chromameter) and the test subject's eyes.

In a study conducted by Tuaycharoen & Tregenza (2007), the influence of view on glare sensation was investigated using a single test room. Both subjective evaluations and photometric data were simultaneously collected, whereby a CCD camera and a shielded illuminance meter were mounted on tripods behind the test subject. To compare the photometric differences caused by the position of the experimental equipment with respect to the test subject's eyes, two independent images were captured: 1) at the original position of the camera during the investigation; 2) at the position of the test subject's head after the session (Tuaycharoen, 2006). For their study, the DGI was selected as the dependent variable, which was calculated on the basis of the luminance of the glare source (the window) and the surrounding luminance background. The luminance contained within the pixels captured by the CCD camera corresponding to different surfaces was used as the dependent variable in the inferential tests in order to determine whether there were differences between the luminance captured by the CCD camera and that perceived by the test subjects.

For this investigation, the DGP was utilised as the dependent variable to detect the effect of time of the day on glare sensation. Unlike other indices, the DGP is mainly dependent on the vertical illuminance at the eye (Wienold & Christoffersen, 2006). On this basis, a study was designed to determine whether the illuminance reaching the camera lens (calculated by *Evalglare* from High Dynamic Range Images (HDRI) analysed by the Photosphere software) and that received at the position of the subject's head were different from each other.

E.2. Method

E.2.1. Experimental Design

To respond to the research question, an experiment was designed utilising the test room described in Chapter 4.1.2., yet altering the position of the Chromameter with respect to the user assessment studies (Figure E1).

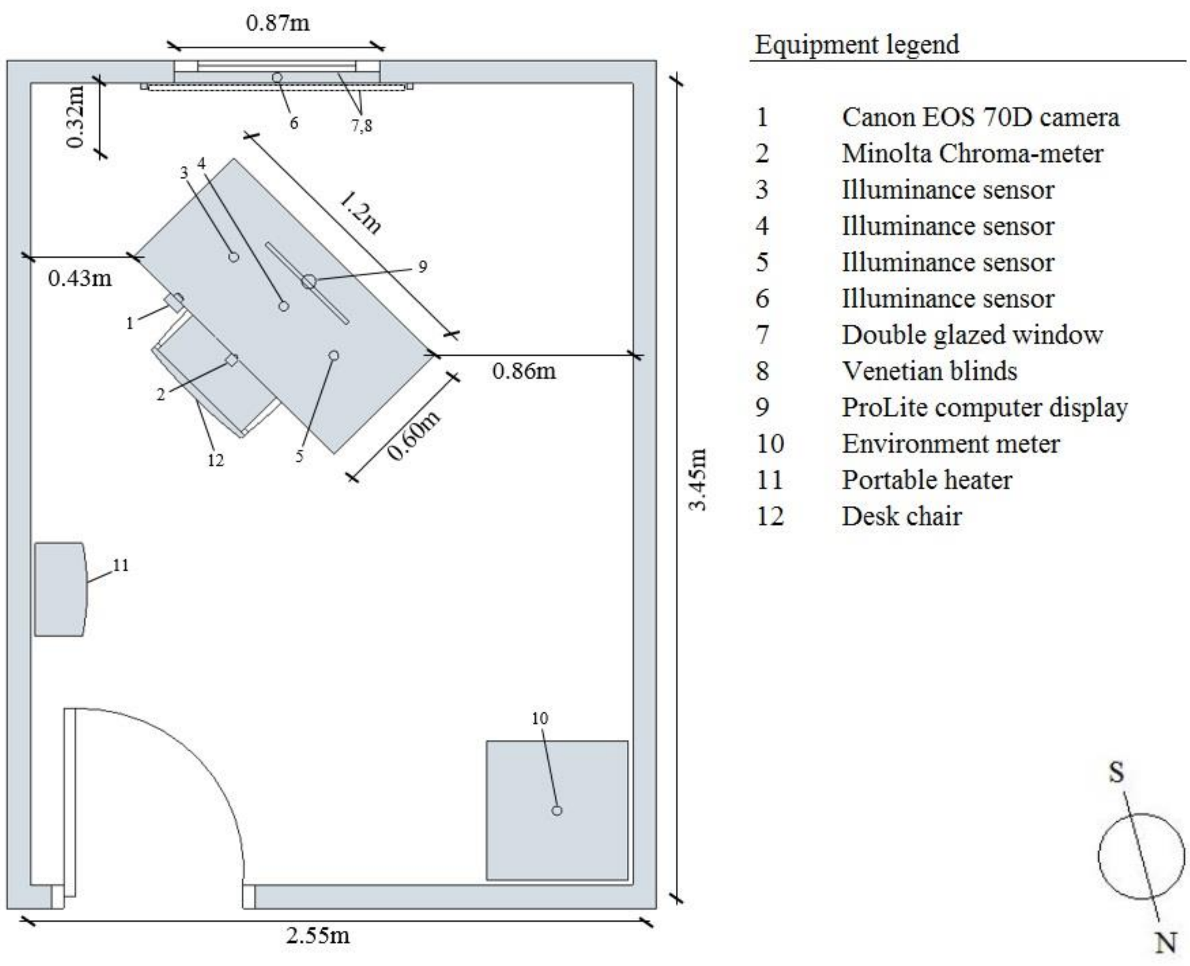


Figure E1 Layout of the test room and list of equipment

To ‘instantaneously’ capture the luminous environment within the test room, two photometric instruments were used: 1) a CCD camera equipped with a fish-eye lens; 2) an illuminance Chromameter. The camera and the Chromameter were mounted on the desk at two different positions: the camera was located in the same position as that described in the user assessment studies, i.e. on the left-hand side of the test subject; the Chromameter was mounted at the position of the test subject's head. Both instruments were positioned at a height of 0.2 m from the desk surface (this being the distance between the upper surface of the desk and the sensors of the photometric instruments) and were pointed towards the VDU (Figure E2). The distance between the centre of the camera lens and the sensor of the Chromameter was 0.2 m.

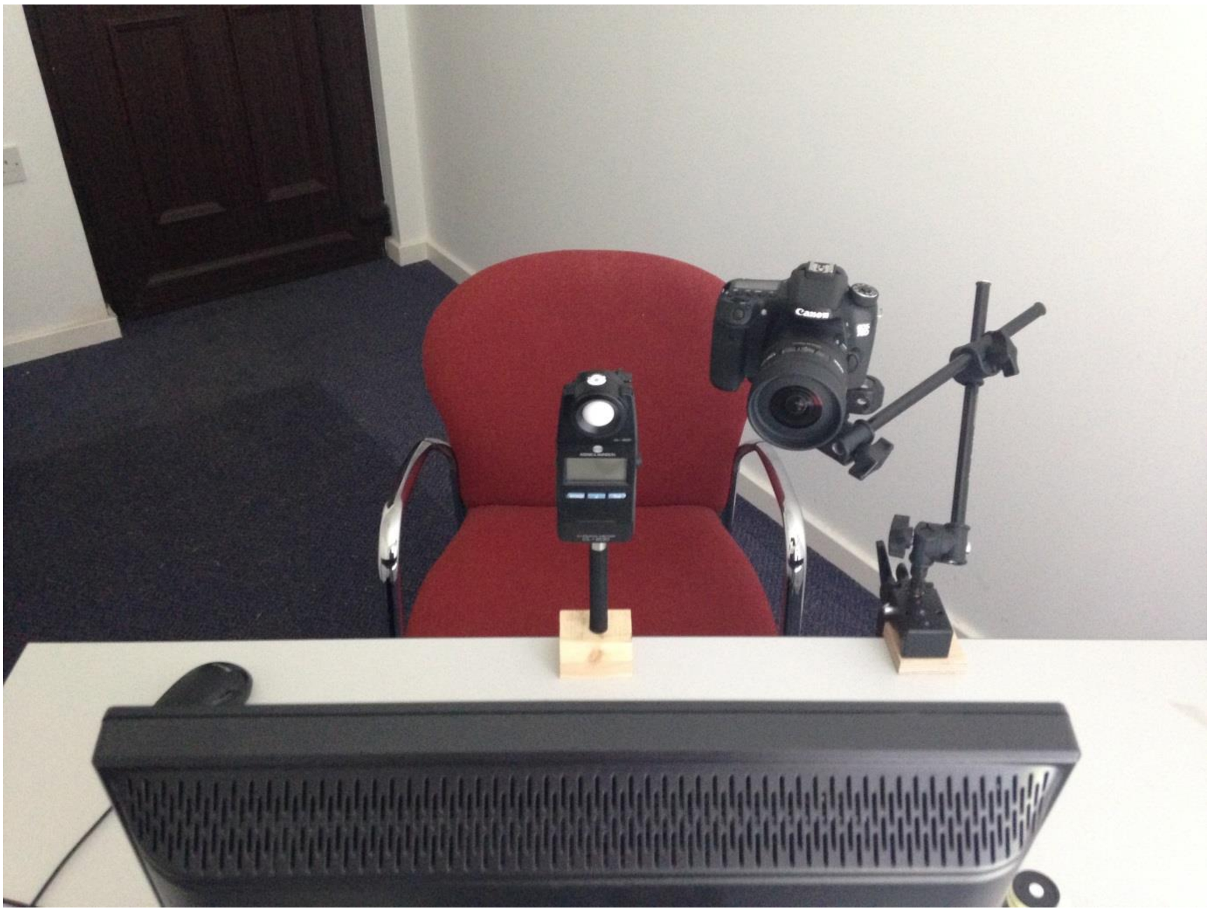


Figure E2 CCD camera and illuminance Chromameter

In the previously described experimental studies, to check for the photometric integrity of each luminance image taken, comparisons between the calculated illuminance reaching the camera lens and that measured by the Chromameter were made (Section 4.1.2.1.4). The results for the MAD and RMSE metrics – which estimate average errors expressed in the units of the variable of interest (Willmott & Matsuura, 2005) – showed that there was a relatively small difference between the calculated illuminance that reaches the camera lens and the illuminance measured by the Chromameter sensor: 'Default Shading': MAD= 26.70, RMSE= 95.89, n= 360; 'User-set Shading': MAD= 78.94, RMSE= 220.46, n= 360. In this case, the two instruments were mounted next to each other. Since the results of these error statistics are expressed within the units of the variable of interest, the outcome of these tests are often misinterpreted as defined benchmarks cannot be provided (Chai & Draxler, 2014; Willmott & Matsuura, 2005). However, according to the literature, minor calibration differences between sensors become less problematic at higher levels of illuminance (Tregenza, 1998). As a result, it was assumed that the camera and the Chromameter utilised in the experimental studies had a similar calibration to each other (i.e., when both sensors are mounted as close as possible to each other, the difference in illuminance flux are small). Therefore, for the purposes of this investigation, any major differences associated with the calculated illuminance that reaches the camera lens and that recorded by the Chromameter was associated to differences in their position within the test room (i.e., direct sunlight may have reached the CCD camera lens but not the Chromameter sensor or vice versa).

E.2.2. Experimental procedure

Under relatively stable sky conditions (i.e., fully overcast or clear sky), multiple HDRI images were generated from LDRI pictures taken by the CCD camera, while independent measurements of vertical illuminance were obtained by the Chromameter. For the production of each HDRI image, the same procedure described in the experimental studies was adopted. A series of seven independent LDRI images were taken, with the same variations of EVs so as to capture a large range of luminance distributions within the test room (Tregenza & Loe, 2014). The LDRI images were combined into a Radiance-formatted HDRI using the Photosphere software. The HDRI could then be evaluated using the *Evalglare* tool (Wienold, 2009). In taking the LDRI images, it was important to use the same camera EVs (along with other camera settings such as ISO, gain, etc.) as in the experimental study. This is because the camera response curve that was derived from the initial self-calibration should only be applied to HDRI images produced using the same EVs of the camera and lens that were used in the calibration process. For this analysis, two different shading settings were used: one with the venetian blinds set at a default 'cut-off' position (i.e., with a slat tilt that ensured predominantly diffused daylight conditions inside the test room); and, one with the venetian blinds fully open (i.e., with no obstruction caused by the shading device as seen from the position of the experimental equipment) (Figure E3).

Figure E3 Default shading setting (left); Fully-open shading setting (right)

The analysis aimed at investigating the differences in illuminance levels at the position of the camera and of the Chromameter under both shading settings. In fact, under the default shading, it is unlikely that any direct sunlight would reach the camera lens or the Chromameter sensor. Therefore, it can be assumed that any difference in illuminance level between the two instruments would be generated by diffused daylight.

Conversely, under the open shading setting, the potential of direct sunlight transmission into the test room is at its highest, hence it can be hypothesised that direct sunlight may reach the camera lens but not the Chromameter sensor or vice-versa.

For the investigation, a series of 250 HDRI were produced and independent illuminance readings were collected under each shading setting. The data collection was spread unsystematically across times of the day ranging from 09:00 to 15:00 (corresponding to the testing period used for the experimental study).

E.3. Results and Discussion

Figure E4 plots, on the y-axis, the illuminance calculated by *Evalglare* from each HDRI derived from the images captured by the CCD camera, for both the default shading (left) and the open shading setting (right). On the x-axis, the figure presents the illuminance levels measured by the Chromameter. The null hypothesis line is plotted along the diagonal of the graph, representing no difference between the illuminance levels calculated from the CCD camera and those measured by the Chromameter.

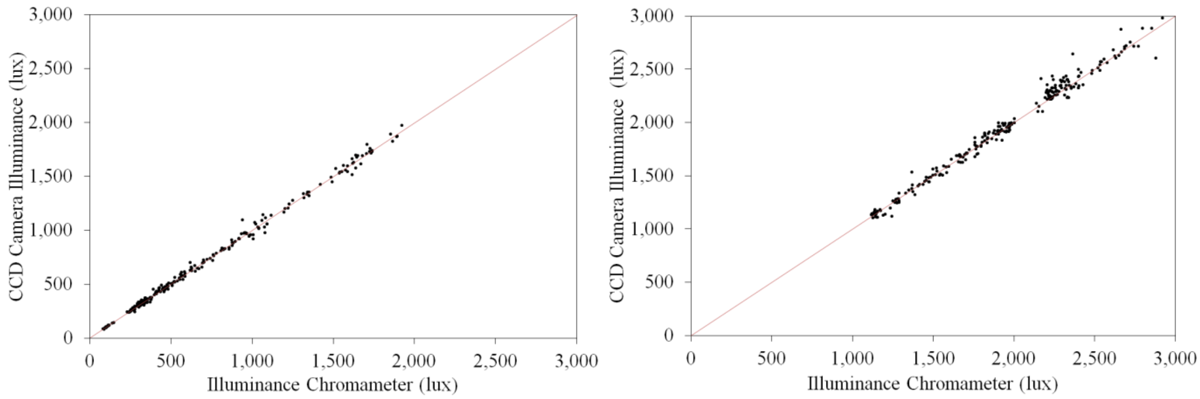


Figure E4 Comparison between calculated (from the CCD images, y-axis) and measured (from the Chrometer, x-axis) vertical illuminance under the default (left) and fully-open (right) shading settings

Graphical inspection of Figure 4E reveals that, for the default shading setting (left), the majority of the illuminance points are close to the null hypothesis line, this suggesting that the illuminance values calculated from the CCD camera do not show a tendency to correspond to higher or lower values than the levels measured by the Chromameter. For the fully-open shading setting (right), visual inspection of the display reveals a slight tendency for illuminance points to be located above the null hypothesis line, this leading to hypothesise higher levels of illuminance calculated from the values recorded by the CCD camera with respect to the levels measured by the Chromameter. This suggests that, in the experimental study, the illuminance received at the position where the photometric instruments were located could have potentially been higher than the illuminance received at the test subject's eyes when the venetian blinds were no longer set at their default 'cut-off' position.

To examine in more detail the relationship between the illuminance values calculated from the images captured by the CCD camera and the levels measured by the Chromameter, both the MAD and RMSE were calculated by making use of both sets of illuminance values according to the following formulas (equation 1E and 2E) (Chai & Draxler, 2014; Willmott & Matsuura, 2005):

$$MAD = \frac{1}{N} \sum_{i=1}^{n} |E_{CCD} - E_{CM}| \qquad (1E)$$

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (E_{CCD} - E_{CM})} \qquad (2E)$$

Where E_{CCD} is the illuminance calculated from the HDRI images captured by the CCD camera, and E_{CM} is the illuminance measured by the Chromameter.

Table E1 displays, for each shading setting, the mean and standard deviation (SD) of the vertical illuminance levels calculated from the CCD camera images ($Mean_{CCD}$) and the values measured by the Chromameter ($Mean_{CM}$), the MAD, and the RMSE.

Table E1 Descriptive analysis between calculated (CCD) and measured (CM) illuminance

Shading Setting	$Mean_{CCD}$ (SD)	$Mean_{CM}$ (SD)	MAD	RMSE
Default Shading	683.49 (468.80)	682.92 (470.66)	17.39	26.65
Open Shading	1,926.74 (467.31)	1,915.96 (454.72)	39.26	57.76

The results indicate that, for both the MAD and the RMSE, the average errors that exist between the calculated and measured illuminance values are lower under the default shading setting. This finding is consistent with the visual inspection of graphical displays, which suggested that the difference in illuminance values recorded by the instruments was lower when the venetian blinds were set at their default ‘cut-off’ position.

In interpreting the results of Table E1, it should be noted that both the MAD and RMSE – which are expressed in the units of the variable of interest (lux) – present low values. In fact, the errors observed in this analysis are lower than those obtained within the experimental study, when a comparison was made between the illuminance calculated from the CCD camera images and the illuminance measured by the Chromameter with the two instruments being mounted adjacent to each other (see Section 4.1.2.1.4, Figure 4-12).

Consequently, the results from this analysis provide both statistical (MAD and RMSE) and graphical (linear regression) evidence to suggest that the difference between the illuminance calculated from the images captured by the CCD camera and the illuminance measured by the Chromameter at the approximate position of the test subject's head is not substantive.

Therefore, this analysis verified that, in the experimental study, the effect of displacement of the photometric instruments relative to the test subject's head position is not likely to have caused any differences in illuminance reading. Since, according to the literature, the DGP – used as the primary performance parameter to detect the effect of time of day on glare response from daylight – is mainly dependent on the vertical illuminance received at the eye (Wienold & Christoffersen, 2006), this analysis has demonstrated that it is unlikely that any relevant experimental bias caused by the displacement of the CCD camera relative to the head position of the test subject may have influenced the findings of this investigation.

Appendix F

Results for Unified Glare Rating

This Appendix presents the results of the Discomfort Glare and Time of Day laboratory study (Section 3.1), using the Unified Glare Rating (UGR) index as evaluation parameter.

Table F1 Wilcoxon matched-pairs tests and effect sizes for ‘Just Perceptible’

Time of Day	M_{dn} (IQR)	M_{dn} (IQR)	ΔM_{dn}^{NHST}	Positive	Negative	Ties	Effect Size (r)
Mid. vs. Morn.	9.80 (1.55)	9.06 (2.18)	0.74 n.s.	21	8	1	0.27
Aft. vs. Morn.	10.06 (2.54)	9.06 (2.18)	1.00*	23	6	1	0.43
Even. vs. Morn.	10.60 (2.62)	9.06 (2.18)	1.54**	26	4	0	0.56
Aft. vs. Mid.	10.06 (2.54)	9.80 (1.55)	0.26*	18	8	4	0.44
Even. vs. Mid.	10.60 (2.62)	9.80 (1.55)	0.80***	24	5	1	0.66
Even. vs. Aft.	10.60 (2.62)	10.06 (2.54)	0.54 n.s.	18	10	2	0.33

With Bonferroni Correction: * 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

Table F2 Wilcoxon matched-pairs tests and effect sizesfor ‘Just Noticeable’

Time of Day	M_{dn} (IQR)	M_{dn} (IQR)	ΔM_{dn}^{NHST}	Positive	Negative	Ties	Effect Size (r)
Mid. vs. Morn.	12.48 (3.36)	10.88 (3.51)	1.60*	24	5	1	0.48
Aft. vs. Morn.	13.40 (5.07)	10.88 (3.51)	2.52***	25	5	0	0.68
Even. vs. Morn.	13.61 (3.55)	10.88 (3.51)	2.72***	25	5	0	0.69
Aft. vs. Mid.	13.40 (5.07)	12.48 (3.36)	0.92*	21	9	0	0.53
Even. vs. Mid.	13.61 (3.55)	12.48 (3.36)	1.13*	23	6	1	0.52
Even. vs. Aft.	13.61 (3.55)	13.40 (5.07)	0.21 n.s.	19	9	2	0.23

With Bonferroni Correction: * 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

Table F3 Wilcoxon matched-pairs tests and effect sizes for ‘Just Uncomfortable’

Time of Day	M_{dn} (IQR)	M_{dn} (IQR)	ΔM_{dn}^{NHST}	Positive	Negative	Ties	Effect Size (r)
Mid. vs. Morn.	16.86 (4.20)	13.61 (4.88)	3.25***	25	5	0	0.66
Aft. vs. Morn.	17.41 (6.55)	13.61 (4.88)	3.80***	26	4	0	0.75
Even. vs. Morn.	17.86 (6.50)	13.61 (4.88)	4.25***	26	4	0	0.74
Aft. vs. Mid.	17.41 (6.55)	16.86 (4.20)	0.55*	21	9	0	0.46
Even. vs. Mid.	17.86 (6.50)	16.86 (4.20)	1.00 n.s.	18	10	2	0.35
Even. vs. Aft.	17.86 (6.50)	17.41 (6.55)	0.45 n.s.	17	13	0	0.08

With Bonferroni Correction: * 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

Table F4 Wilcoxon matched-pairs tests and effect sizes for ‘Just Intolerable’

Time of Day	M_{dn} (IQR)	M_{dn} (IQR)	ΔM_{dn}^{NHST}	Positive	Negative	Ties	Effect Size (r)
Mid. vs. Morn.	22.21 (5.74)	17.95 (6.04)	4.26***	23	7	0	0.62
Aft. vs. Morn.	22.70 (6.56)	17.95 (6.04)	4.75***	24	6	0	0.75
Even. vs. Morn.	22.28 (6.05)	17.95 (6.04)	4.33***	24	6	0	0.72
Aft. vs. Mid.	22.70 (6.56)	22.21 (5.74)	0.49*	19	11	0	0.46
Even. vs. Mid.	22.28 (6.05)	22.21 (5.74)	0.07 n.s.	16	13	1	0.27
Even. vs. Aft.	22.28 (6.05)	22.70 (6.56)	-0.42 n.s.	12	16	2	0.14

With Bonferroni Correction: * 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