MASTER

The Relation of Uniformity and Room Brightness and the Role of the Mean and Median Luminance in Brightness Prediction

Versteden, Julia A.E.

Award date:
2023

The Relation of Uniformity and Room Brightness and the Role of the Mean and Median Luminance in Brightness Prediction

J.A.E. Versteden

1259083

in partial fulfillment of the requirements for the degree of

Master of Science

In Human-Technology Interaction

July 2023

Graduation Committee:

prof. dr. ir. Y.A.W de Kort	Department of Industrial Engineering and Innovation Sciences
dr. A. de Vries	Department of Lighting Applications – Signify Research
dr. K. Chamilothori	Department of Industrial Engineering and Innovation Sciences
dr. ir. K.C.H.J. Smolders	Department of Industrial Engineering and Innovation Sciences

---

Eindhoven University of Technology
Department of Industrial Engineering and Innovation Sciences

Abstract

The lighting distribution within a room is essential for lighting design. Previous research has highlighted the importance of the uniformity and brightness of lighting distributions for the appraisal of a room. Several studies found that uniform lighting distributions appear brighter while others suggest the opposite. To understand these contrasting results, the current research investigated the effect of uniformity of a lighting distribution (represented by the number of spots used and the beam width of these spots) on brightness, in a two-alternative forced choice paired comparisons experiment in virtual reality. Results show that there does not seem to be an overarching effect of uniformity on brightness, but that brightness appears to be influenced by multiple separate factors within the lighting distribution. Moreover, the results suggest that the median luminance is a good predictor for brightness, performing better than the mean luminance. This work highlights the importance of systematically testing different aspects in the lighting distribution in order to gain a better understanding of brightness perception.

Keywords: brightness, perception, lighting distribution, uniformity, VR

Table of Contents

1. Introduction

Understanding brightness is essential for lighting design. Obtaining a good understanding of brightness can guide lighting designers to provide sufficient light while also considering other factors, such as energy-savings, cost-efficiency, well-being, or performance in lighting design. An example application where well-being and performance are important parameters to consider is in an office environment. A good understanding of the interrelation of lighting parameters can help the lighting design for offices. For instance, research into environmental lighting has indicated that brightness and uniformity are essential parameters influencing the atmosphere of a room (Stokkermans et al., 2018).

Although research on brightness goes all the way back to the 1800's with Weber and later Fechner (1860), it is still a poorly quantified concept. Multiple researchers have tried to predict the brightness of objects or scenes, but it is a complicated phenomenon as it is influenced by multiple mechanisms in human visual perception. Specifically for room brightness, Loe et al. (1994) found that the mean luminance within a 40 degrees horizontal band was a good predictor of brightness of a scene, whereas De Vries et al. (2022) found that the median luminance within the same and a slightly wider band was a better predictor. However, all of these studies utilized a relatively uniform luminance distribution, so it is unsure how these models predict brightness in less uniformly lit environments.

The brightness of a surface depends on more than solely the absolute luminance that reaches the eye. Instead, brightness perception is biased by, among others, lightness, color, shapes, and associations (Kingdom, 2011; Meier et al., 2007). Simple models – such as those proposed by Loe et al. (1994) or Cuttle (2009) – do not incorporate the luminance variation that could be present within a surface. Several researchers have tried to understand the relationship between uniformity and brightness, but results are contradictory. Tiller and Veitch (1995) found that non-uniform lighting distributions appeared brighter, whereas Kirsch (2015) found the opposite. In order to gain a better understanding of brightness, the current research investigates how uniformity in a lighting distribution relates to brightness and explores whether the mean or median luminance provides a better basis for comparison to understand the effects of brightness.

1.1. What is brightness?

1.1.1. Physics of Light

Light is the visual part of the electromagnetic spectrum. For the human eye this is between 380 and 780 nm. The photoreceptors in the human eye transduce radiation and send signals to the visual cortex which enables us to see the environment around us at different light levels and distinguish colors (Mather, 2016).

Light can be measured in radiometric and photometric units. The difference between the two is that the former is based on the entire optical spectrum and the latter is weighted by the sensitivity of the human eye, defined in $V(\lambda)$ and $V'(\lambda)$ for photopic and scotopic vision, respectively (CIE, 1926; CIE, 1951). The different aspects of light in this report are referred to in photometric units (Mather, 2016):

- *Luminous flux* (in lumen) is the amount of energy emitted by a light source.

- *Luminous intensity* (in candela) is the amount of energy emitted per unit solid angle (cone).
- *Illuminance* (in lux (lumen/meter^2)) is the amount of light falling onto a surface.
- *Luminance* (in candela/meter^2) is the amount of light being reflected off the surface per unit area in the measurement direction.

1.1.2. Luminance-Brightness relationship

In the 1800's, Weber and later Fechner (1860) introduced psychophysics. This is the domain where the psychological response to a physical stimulus is measured. Fechner discovered that an increase of stimulus does not result in the same increase of sensation. In terms of brightness, a doubling of luminance is not equal to a doubling of brightness. The Weber-Fechner law describes that the sensation intensity is a logarithmic function of stimulus intensity. Later, Stevens (1960) discovered that the sensation intensity is a power function of stimulus intensity, with a relationship that differs for different sensations. His power law relationship between brightness and luminance is described as follows: $B = k * L^n$, where *k* and *n* are constants. He found that *n* is 0.33-0.5 for brightness, 1.2 for lightness and 0.6 for loudness.

However, Stevens (1960) made use of a small target in a dimmed surround, therefore it was uncertain how this relationship would hold for more complex stimuli. Marsden (1969) found that this exponential was dependent on the previous stimuli, the surrounding space, target size, the lightness of the surface and the hue of the surface. Bodmann and La Toison (1994) subsequently created a model to predict brightness-luminance relationships over a range of different background luminance values. They found that the higher the luminance contrast in the room, the lower the brightness of darker surfaces. Increasing the luminance of bright surfaces does not have a considerable impact on the brightness of the room. In order to increase the overall brightness in a room, the luminance of the darker surfaces should be increased. This finding indicates that uniform lighting appears brighter than less uniform lighting.

While these models will be helpful to understand the basic relationship between luminance and brightness, they should be used cautiously in more complex interiors. Marsden (1970) did an experiment in a furnished interior space where he had participants rate the brightness of different surfaces in a room based on a ratio scale. He found that the brightness of a single surface increases with a power law index of 0.35, but with multiple surfaces it would increase with a power law index of 0.6. These relationships can predict the change in brightness of surfaces in a room where the brightness of the brightest surface in the room can then be described as $B_{max} = L_{max}^{0.35}$ and the brightness of the other surfaces as $\mathrm{B} = a * L^{0.6}$, with $a = \frac{B_{max}}{L_{max}^{0.6}} = L_{max}^{-0.25}$.

Although these relationships can predict the brightness in certain situations, its applicability will be limited when scenes become more complex. Note that Marsden's relationship is based on the maximum luminance values within the surface. The surfaces Marsden used in his study contained relatively uniform distributions. In such cases the values of maximum, minimum, mean, and median luminance are relatively close to each other, while these parameters can greatly vary in non-uniform light distributions. Hence, his model does not consider the luminance variation that could be present on the surface.

1.1.3. Light Spectrum

Harrington (1954) found that brightness was dependent on the correlated color temperature (CCT) of a light source. According to his study a higher CCT has a higher brightness and vice versa. Alman (1977) found that this effect was due to a shift on the blue-yellow axis over the Planckian locus in the CIELAB space. Furthermore, highly saturated colors appear brighter than less saturated colors with the same luminance. This is known as the Helmholtz-Kohlrausch effect (Nayatani, 1998). On the other hand, translucent surfaces appear less bright (Marsden, 1970). However, later it was found that both CCT and CRI were not sufficient predictors of brightness (Boyce, 1977; Van de Perre et al., 2023).

In 1990, Berman et al. discovered that not only the fovea and photopic vision assess brightness, but that also scotopic vision outside of the fovea contributes to brightness perception. In their study they compared different light sources with varying correlated color temperatures, spectra and ratios of scotopic/photopic (S/P) illuminance. In an earlier study Berman et al. (1987) found that two light sources with equal photopic illuminance, but differing spectra evoked a different pupil size. They concluded that the pupil controls the amount of scotopic light that falls into the eye. This indicates that the brightness is dependent on the whole spectrum of a light source and not solely the correlated color temperature. On the other hand, Fotios et al. (2015b) found that the S/P ratio was insufficient to predict the brightness. They argue that both the S/P ratio and gamut area metrics are necessary for predicting the brightness under different spectral power distributions.

The fairly new discovery of the intrinsically photosensitive retinal ganglion cells (ipRGCs) has shed new light on the influence of light on our body. These cells have been found to be contributing to short-term, but also long-term lighting effects. More specifically, the ipRGC's are found to contribute to non-visual effects of environmental lighting. Stimulation to these cells is found to affect alertness (Lok et al., 2018), mood (Bedrosian, & Nelson (2013) and the circadian rhythm (Legates et al., 2014). Moreover, there has been found evidence that these cells contribute to brightness judgements as well (Brown et al., 2012; Yamakawa et al. 2019). Recent literature suggests the ipRGCs may be the main driver for pupil dilation (Sandoval Salinas et al., 2020) in response to brightness perceptions.

These theories suggest that brightness is a very complex phenomenon and dependent on much more than just the amount of light that is being received by the eye. The next section will cover more theories behind brightness perception.

1.2. Mechanisms behind brightness perception

In order to apply brightness theories, it is important to have a fundamental understanding of how brightness is perceived. An interesting take is that of Purves et al. (2004), who suggest that what is perceived is determined by the probability distributions of possible light sources, and is, therefore, related to past experiences rather than physical attributes of the stimulus and surfaces. Although perception will sometimes be affected by associations and interpretation, other theories explain that brightness perception already occurs in early visual processing (e.g., the retina (Shapley & Enroth-Cugell, 1984) or the primary visual cortex (Rossi et al., 1996)). Multiple researchers have tried to understand the mechanisms behind brightness perception, but it is a complex field. This section explores some mechanisms in visual processing that seek to explain different effects in brightness perception.

1.2.1. Brightness and Lightness

Brightness is not to be confused with lightness. Brightness is described by Boyce (2014) as "*an attribute of light, related to whether more or less light is seen to be emitted*", whereas lightness is "*an attribute of the surface, related to whether more or less light is seen to be reflected*". Due to lightness constancy (Kingdom, 2011), the lightness of an object will be constant across different lighting conditions, but the brightness will change. This can be illustrated by Figure 1 (Adelson, 1993). Patch A appears brighter than patch B, while B appears lighter than A, whereas the luminance of the two is the same (Perdreau & Cavanagh, 2011). In any lighting condition, the visual system adapts to perceive the environment as constant. This top-down process influences our perception of brightness. Thus, the visual system compensates for the perceived difference in illumination since patch B appears to be in the shadow.

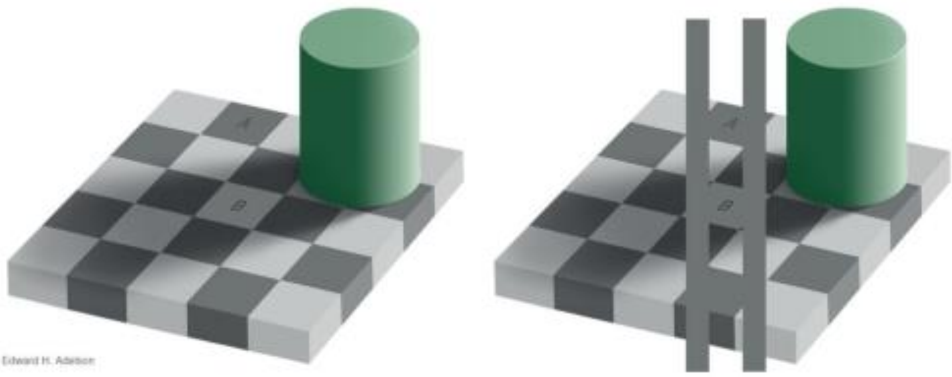


Figure 1. Example of lightness constancy. Patch B appears to be lighter than patch A although they have the same luminance (Adelson, 1993).

Furthermore, Barraza and Martín (2020) found that the texture of a surface has a significant influence on brightness as well. In their study they conclude that with equal luminance, a smooth wall is perceived as brighter than a textured wall.

1.2.2. Adaptation

The ability of the human eye to detect absolute light levels is relatively poor because the eye can respond well to spatial and temporal changes. When the light levels change, the eye adapts accordingly. This process involves three stages: mechanical adaptation (pupillary response), neural adaptation and photochemical adaptation (Ferwerda et al., 1996). The different layers of cells in the retina enable us to see in multiple different lighting conditions, while maintaining a stable perception of brightness. For instance, with one bright light source, the brightness will vary across a room and create sharp gradients and dark corners. The eye will adapt while looking around to try to create an even brightness within the whole room.

Furthermore, the time required for the visual system to adapt differs for different stimuli. For example, dark-to-light adaptation is faster than light-to-dark adaptation (Mather, 2016). Furthermore, adaptation to visually unfamiliar and more complex scenes takes longer (Webster, 2015; Stokkermans & Heynderickx, 2014).

1.2.3. Anchoring Theory

Anchoring theory suggests that the visual system uses an anchor to judge the luminance and lightness of surrounding surfaces. For example, the surface with the highest luminance could appear white, the darkest could appear black or the average could be mid-gray (Gilchrist et al., 1999). Subsequently, this anchor is used to judge the brightness and/or lightness of the surrounding surfaces.

1.2.4. Lateral Inhibition

The brightness of a surface depends on the contrast with its surrounding surfaces. For instance, a gray patch appears brighter (or lighter) when placed next to a black patch than to a white patch. This phenomenon is known as simultaneous brightness contrast (Frisby & Clatworthy, 1975) and is argued to be due to anchoring theory (Economou et al., 2007) or lateral inhibition. The latter refers to the ability of the excited neurons to inhibit the activity of the neighboring cells. Consequently, only the most and least stimulated neurons will respond. This mechanism enhances contrast and reduces noise in the visual signal, therefore sharpens the edges of objects (Mather, 2016). The cells responsible for this phenomenon are the off-center and on-center ganglion cells (Purves et al., 2004; Baxant et al., 2016).

Lateral inhibition is also responsible for the Chevreul illusion (Ratliff, 1972). When two different gray luminance patches touch, the contrast around the edge increases. The edge along the darker area appears even darker and the lighter area even lighter (Figure 2, left). This effect does not occur when the patches do not touch directly (Figure 2, right).

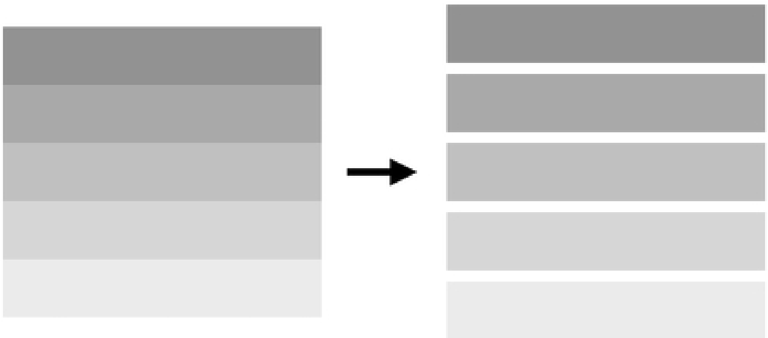


Figure 2. Example of the Chevreul illusion. When two different luminance patches touch (left) there seems to be an increase of contrast along the edge. This is not the case when they do not touch (right) (Weerakkody et al., 2009).

This can be exaggerated (Figure 3, bottom) by adding a smooth luminance gradient that progresses in the same direction as the Chevreul staircase, also known as a ramp. On the contrary, this illusion is inhibited (Figure 3, top) when the ramp goes in the opposite direction (Lu & Sperling, 1996; Geier & Hudak, 2011).

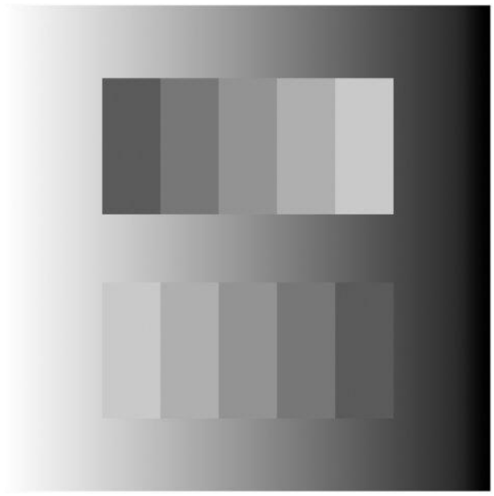


Figure 3. Addition of a gradient to the Chevreul illusion, the illusion is more extreme when the gradient goes in the same direction as the ramp (bottom) but is inhibited when it goes in the opposite direction (top) (Geier & Hudak, 2011).

1.2.5. Filling-In

The brightness of a luminant surface is dependent on the luminance at the edge. One illusion that can illustrate this phenomenon is the Craik-O'Brien-Cornsweet effect (Figure 4) (Todorović, 1987). The addition of a luminance gradient at the edge affects how the whole patch is perceived. The patch on the right has a dark gradient at the border, while the left has a light gradient at the border. The two patches have a different brightness, while the luminance away from the border is the same.

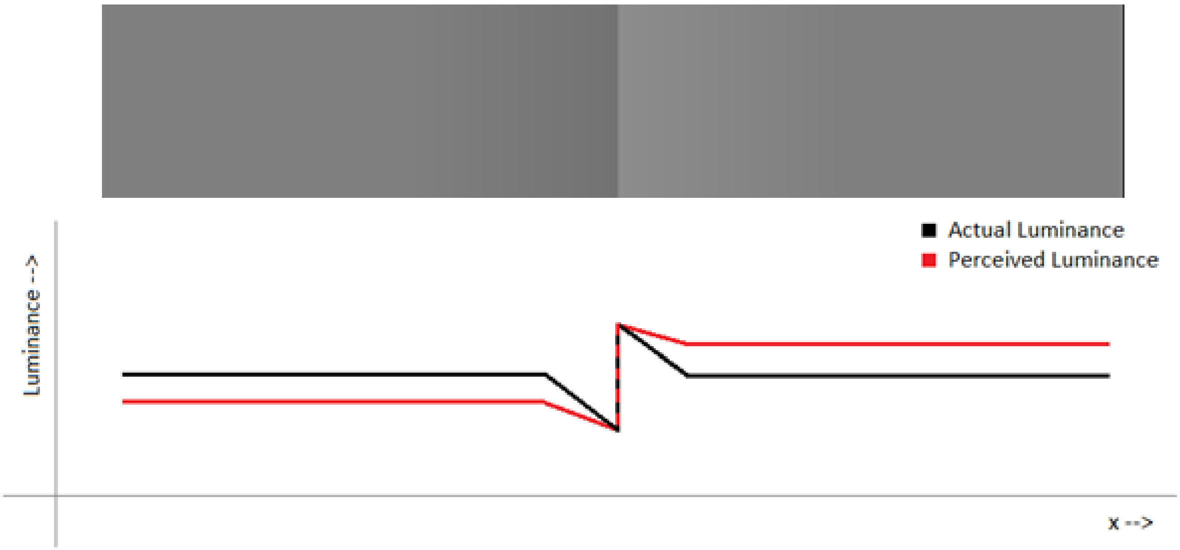


Figure 4. Craik-O'Brien-Cornsweet effect. Due to the luminance gradient at the border, the luminance of the two patches seems to differ although it is the same shade (Adapted from Masuda et al., 2014).

This phenomenon is argued to be explained by the filling-in theory (Gerrits & Vendrik, 1970; Grossberg & Todorovic, 1988; Paradiso & Nakayama, 1991; Davey et al., 1998), but others have suggested that this is not the case, and that the brightness is determined by solely edge-integration effects (Land & McCann, 1971; Cornelissen, 2006) or more complicated (Masuda et al., 2011).

1.2.6. Glare

Another phenomenon that influences brightness perception is glare, related to an extreme brightness non-uniformity (Boyce, 2014). Glare can take two forms: disability glare and discomfort glare, whereas the former is related to a reduction in visual performance, the latter is related to a feeling of discomfort. Glare can also be created artificially. The addition of glare or a halo will vary the brightness, with equal luminance (Figure 5). The addition of a luminance gradient towards the center will give the illusion of glare and increase the brightness whereas a luminance gradient towards the outer border will reduce the brightness and give the illusion of a halo. The center on the left appears brighter than the center on the right, even though they have equal luminance (Kinzuka et al., 2021).

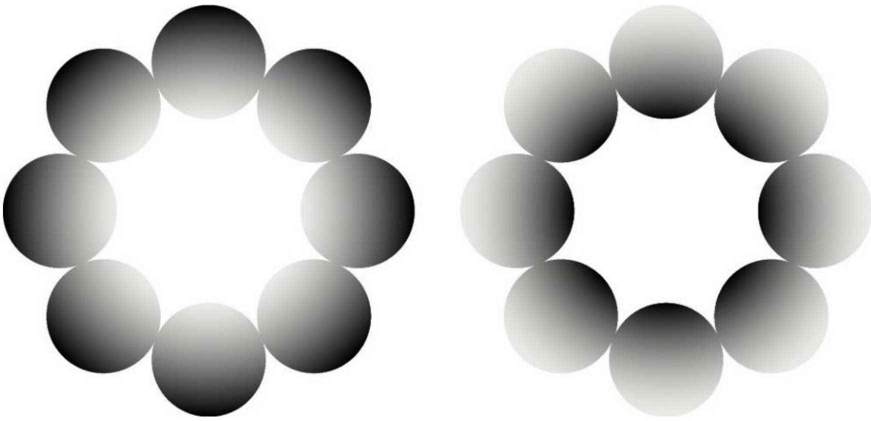


Figure 5. Glare versus halo illusion. The addition of a gradient at the border influences brightness. A gradient towards the center will give the illusion of glare (left), while a gradient away from the center will give the illusion of a halo (right). The center on the left appears brighter than the center on the right, even though they are equal in luminance (Kinzuka et al., 2021).

Furthermore, it was found that brief exposure to glare in the peripheral field of view reduces the brightness of a test patch in the foveal field of view (Colombo et al., 2000). Moreover, previous research has shown that the perceived intensity of glare is dependent on the location of the source within the visual field (Iwata & Tokura, 1997; Kim & Kim, 2011)

All of these effects are examples to understand how the visual system perceives luminance. Due to the complexity of the field, the literature is still inconclusive. Nevertheless, this section illustrated that brightness is affected by many internal and external influences, such as properties of the surface, the luminance of the surrounding surfaces, adaptation of the visual system, edge effects and glaring sources. Thus, the exact brightness of the environment is dependent on more than solely the absolute luminance of the surface. These mechanisms can have a significant influence on the perception of the surrounding environment. Therefore, it is important to understand these mechanisms so they can guide lighting designers.

1.3. Environmental lighting

For humans, the most important sense for perceiving the surrounding environment is the eye. Fundamental knowledge of visual perception is essential for lighting design. In general, the primary goal of lighting design for offices is to enhance performance by creating a safe and comfortable working environment (Yilmaz, n.d.). In recent years, the term “Integrative Lighting”, described by ISO/CIE (2022) as *“lighting integrating both visual and non-visual effects, and producing physiological and/or psychological benefits upon humans*” has gained more attention. The term is frequently used in research focused on studying human perception mechanisms in response to a wide variety of lighting aspects. Obtaining a good understanding of the interrelation of different lighting aspects creates the opportunity to design lighting conditions that fulfill lighting requirements while also considering relevant design goals such as energy-savings, well-being, performance, or cost-efficiency (Kozusznik et al., 2019).

Multiple environmental lighting designs can fulfill lighting requirements but might differ in subjective impressions of a room (Flynn et al. 1973). The lighting distribution in a space has been found to affect both the appearance and the appraisal of a room. Different aspects include the perceived spaciousness (Lindh & Billger, 2021), atmosphere (Vogels, 2008; Stokkermans et al., 2018), comfort and interest (Abboushi et al., 2020).

Although traditionally an office is uniformly lit, Hawkes et al. (1979) conclude that an office space with bright and interesting lighting distribution is preferred. The former is related to brightness, whereas the latter is related to the perceived uniformity. Stokkermans et al. (2018) did a study on the influence of attributes of light – brightness and perceived uniformity – on the perceived atmosphere in a room. Vogels (2008) found that the atmosphere can be described in terms of coziness, liveliness, tenseness, and detachment. Stokkermans et al. found that rooms with medium perceived uniformity and brightness were perceived as most cozy. Liveliness can be increased by increasing the brightness of a room. Additionally increasing the perceived uniformity will give the room a more detached feeling, while lowering brightness and uniformity makes a room appear tense. Moreover, brightness was found to be well correlated with the mean luminance within the 40 degrees horizontal band, while uniformity was found to be affected by the type of luminaire and spatial distribution of the luminaires. Their study indicates that the lighting distribution is important for the perception of a room.

Houser et al. (2002) did an experiment where the lighting distribution was varied to be either indirect, direct or a mixture between the two while keeping the horizontal illuminance on the work plane fixed. Their results show that, in contrast to the floor and desk, the ceiling and wall brightness did significantly influence the overall room brightness. Furthermore, with increasing the direct component, the wall appeared less uniform due to the appearance of scallops on the wall. Finally, the room appeared more spacious and was preferred with an increase in the contribution of the indirect component.

Brightness is an important component of room appraisal. Generally, room appraisal can be improved by increasing brightness, although this is not always the case (De Vries et al., 2021). Furthermore, increasing wall luminance, and consequently brightness, has been found to be stimulating (Van Ooyen et al., 1987) and increasing alertness (De Vries et al., 2018).

1.4. Brightness of a room

The relationship between the lighting distribution and the overall brightness of a room has been investigated previously. This section will explore different brightness prediction metrics and will discuss different parameters within the room that affect the perceived brightness.

1.4.1. Brightness Prediction Metrics

Multiple attempts have been made to predict the brightness of a room. Loe et al. (1994) found that the average luminance over a 40 degrees horizontal band was a good predictor for the visual lightness of a room. On the other hand, De Vries et al. (2022) found that both the logarithm of the median luminance over a 60 degrees horizontal band and the ratio of the logarithm of the 95% percentile/median luminance over the 40 degrees horizontal band were good predictors for the brightness of a room. Furthermore, Kato and Sekiguchi (2005) found that brightness increases with an increase in directional diffusivity, related to the incident angle of light. Furthermore, Cuttle (2009) introduced Mean Room Surface Exitance, which predicts the brightness based on the reflectance within a room. More recently, Hu et al. (2023) found that the indirect corneal illuminance is a good predictor for brightness. The indirect corneal illuminance is directly related to the amount of light that falls onto the eye.

What many of these metrics have in common is that they describe the amount of light that falls onto the eye. However, they do not consider what we actually perceive. Brightness is not simply related to the absolute level of light that enters the eye (Jacobsen & Gilchrist, 1988).

Instead, perception is biased; we focus on certain areas, are biased by lightness and shapes, and make associations (Kingdom, 2011; Meier et al., 2007). The average luminance within two rooms might be the same, but if the lighting distributions differ, there might be a difference in brightness between these two rooms. As discussed in section 1.2., this could be due to these two rooms inducing a different anchor or accounting for adaptation. The way light is distributed within a room is therefore an important aspect of brightness perception.

1.4.2. Uniformity

An important parameter of the lighting distribution is uniformity. Uniformity is generally described as the variation of lighter and darker areas of a surface. Perfect uniformity would require no luminance variation on the surface. In that case, the minimum, maximum and mean luminance would be equal. Frequently used uniformity metrics are formulated as:

- A ratio of the minimum to the average luminance, as $U = \frac{L_{min}}{L_{avg}}$ (CIE, 2011).
- A ratio of the minimum to the maximum luminance, as $U = \frac{L_{min}}{L_{max}}$ (Hsieh & Li, 2013).
- A statistical measure coefficient of variation, ratio of the standard deviation to the average luminance, as $CV = \frac{\sigma}{L_{avg}}$ (Armstrong, 1990).
- Weighted to the human visual system, as $U_{HVS} = \frac{100}{1+k*CV^{\alpha}*CV^{\beta}{}_{HVS}}$ in %, where k is a constant and powers α and β are adjusted based on their relative importance (for details see Moreno, 2010).

The ratio-based metrics do not incorporate the spatial luminance variation within a distribution. With equal mean luminance, a distribution with a single large area of peak luminance has an equal uniformity as a distribution with multiple smaller areas of the same peak luminance, while the perceived uniformity might be different. The ratio metrics are dependent on extreme values. This can be helpful in applications where glare is important but is less useful in the current study because this study specifically focusses on the luminance pattern of the lighting distribution.

The statistical measure is not sensitive to extremes but does not limit the influence of extremes. For instance, one high peak luminance will increase the standard deviation while decreasing the uniformity substantially, even though visually the lighting distribution appears relatively uniform with one small glaring source. Thus, this metric does not incorporate lighting patterns. When evaluating lighting uniformity, it is important to consider the impact of extremes and assess their significance in relation to the specific context of the lighting application.

The uniformity weighted by the human visual system (Moreno, 2010), considers spatial frequencies with the contrast sensitivity function. It incorporates pattern detection and the contrast sensitivity of the human visual system. Yao et al. (2017) tested the performance of these three metrics by correlating uniformity rankings in a perceptual experiment. Note that their data were rankings and not ratings, reducing information about the absolute difference between the conditions. Results show that U_{HVS} outperforms the other two metrics, but its performance is lacking in relatively uniform lighting distributions. Furthermore, this metric is adjusted dependent on the subjective interpretation of the importance of different factors. As the current research seeks to understand the relation of uniformity and brightness objectively, this metric does not suit the current research goal.

None of these metrics has a true correlation with the perceived uniformity. Therefore, in the current study the term uniformity is described with the following written definition rather than a metric: "*Lighting distribution providing homogeneous wall coverage with minimal luminance variation from the mean".* Perfect uniformity would imply no luminance variation. Uniformity will thus increase when the lighting assures higher wall coverage, and when extremes are lowered and approach the mean. Uniformity is then influenced by the type of luminaire and light source, the number of luminaires, the location of the luminaires within the space, the intensity of the light, the surface reflectance and more.

1.4.3. Uniformity of the Lighting Distribution

Even if the light falling onto the work plane is sufficient, the way in which the light is distributed can greatly affect the overall brightness of the room. If all light is being concentrated in one single area, then the surrounding areas may appear relatively dim due to an increase of contrast within the room (Jay, 1971).

It is, however, still unclear how (the uniformity of) the lighting distribution of a room affects its perceived brightness. Several attempts have been made to answer this question but resulted in contrasting findings. A number of studies indicate that more uniformly lit spaces appear brighter (Kobayashi et al., 1998; Kato & Sekiguchi, 2005; Hsieh, 2012; Kirsch, 2015), while others suggest the opposite (Tiller & Veitch, 1995; Newsham et al., 2004; Chraibi et al., 2013, Sullivan & Donn, 2016; Sullivan & Donn, 2018).

A frequently referenced paper is Tiller and Veitch (1995), who conducted a pilot study comparing the two lighting conditions with either a uniform distribution or a non-uniform distribution. One scene contained a luminaire equipped with a plastic prismatic diffuser and the other scene was equipped with a deep-cell parabolic louvre. This made the lighting in two scenes differ in gradient, contrast, and the location on the wall. They concluded that – with equal average luminance – a non-uniformly lit space requires 5-10% less lighting than a uniformly lit space. However, in their study they only compared two conditions that were not manipulated systematically, making it hard to draw a conclusion on why this difference was found. Moreover, they did not give a quantification of the level of uniformity and luminance values that were used, making it impossible to compare them to the results of others. In a pilot study, Sullivan and Donn (2016; 2018) compared several different scenes with varying light distributions on the walls, ceiling and floor, by manipulating the reflectance of these surfaces. They found that with equal average luminance, an increased uniformity led to a lower perceived brightness. However, they conclude that this finding is not necessarily due to a difference in uniformity – as metrics of uniformity often do not include the pattern – but might be due to a missing essential parameter in the lighting distribution.

In contrast, Kirsch (2015) found evidence that suggests the opposite effect. In his research he compared a uniform distribution to a gradient and step-gradient, and varied the direction of the gradient (top-to-bottom, away from or towards the observer). He found that the uniform luminance distribution had a slightly higher visual lightness than all the other (non-uniform) lighting distributions. Furthermore, the two step-gradient luminance distribution had the lowest perceived visual lightness. Despite his focus being on the effect of luminance distribution on brightness in relation to the location of the lighting within the field of view, these outcomes suggest that uniform lighting distributions appear brighter than non-uniform lighting distributions. Moreover, his research indicates that the pattern is an important factor in the assessment of brightness of a space.

As suggested by Kirsch (and by Kato & Sekiguchi, 2005), not only the distribution of lighting is important in assessing the overall brightness of a room, but also the location of the lighting within the field of view is an important contributor. This was also concluded in research by Loe et al. (1994) and De Vries et al. (2022), who found that a horizontal band of 40 to 60 degrees is the most important for assessing the visual lightness of a room.

Ishida and Ogiuchi (2002) found that brightness was highly correlated with the amount of light and not the intensity of light. In other words, brightness was related to how much of the light fills the space, more than the intensity of the light source. This finding also indicates that uniform lighting would be perceived as brighter.

To summarize, different models have been proposed to predict the brightness of a room, but they often omit the luminance variation present within a room. While the majority of authors suggest that uniform luminance distributions appear brighter, others suggest the opposite. What all of them have in common is that the luminance variation within a space has an influence on the brightness perception of a space. Their contrasting results can be possible due to missing an essential parameter or simply due to measurement errors. Therefore, the next section will explore different methods for brightness assessment that have been used in the past.

1.5. Methods for Brightness Assessment

Multiple different methods have been used in the literature to present lighting stimuli and assess brightness judgments. This section will briefly explain methods used by others and discuss the most important advantages and disadvantages of these methods.

1.5.1. Test Scenes

In order to make brightness judgements, scenes with different lighting settings have to be created. In previous research this is done by using existing rooms (Van Ooyen et al., 1987), full-scale mock-ups of a space (Loe et al., 1994), scaled mock-ups (Ishida & Ogiuchi, 2002; Hsieh, 2012), 2D computer simulations (Stokkermans et al., 2018; Pracki & Krupiński, 2021) and virtual reality (VR) (Jin et al., 2022). The advantage of using existing rooms is that it is directly applicable, but the biggest disadvantage is that it is inflexible and prone to non-controllable influences. Mock-ups can be used to investigate different light settings without the need for a whole existing room. Moreover, they allow for more control (Bellazzi et al., 2022). Both full-scale and scaled mock-ups can be useful, depending on the research question, although scaled mock-ups might have difficulties reproducing real-life lighting effects. However, both these methods can be costly and time consuming. Therefore, computer simulations can be helpful (Newsham, Richardson, Blanchet & Veitch, 2005). VR is more suited than 2D computer simulations due to its immersiveness (Natephra, Motamedi, Fukuda & Yabuki, 2017). Moreover, VR is an emerging tool in perception experiments in the lighting field. The application of VR in lighting research will be discussed in more detail in section 1.6.

1.5.2. Methods for Assessment

Different methods for brightness assessments have been done in the past. These include both absolute and relative evaluations, ranging from brightness matching (Tiller & Veitch, 1995; Sullivan & Donn, 2016; Fotios & Cheal, 2011) and questionnaires (De Vries et al., 2022) to comparisons (Kobayashi, Nakamura & Inui, 1998; Kato & Sekuguchi, 2005) and forced choice procedures (Royer & Houser, 2002; Hu et al, 2023). In the brightness matching procedures, participants are instructed to directly adjust the lighting settings to match the brightness of the reference. The advantage of this method is that the result is absolute and directly applicable, but this method can often be difficult to realize. In questionnaires, either verbal or written, participants are asked to rate the brightness of a scene relative to the other scenes. The biggest advantage of this method is that it is relatively quick, but results can be hard to generalize due to the relativity of the measurements. With comparisons, participants are located in a space with certain lighting characteristics after which they will be located in a second space with different lighting characteristics. Consequently, they are asked to indicate how the brightness of this condition relates to the previous condition. The biggest disadvantage of this method is that it is quite time-consuming and is prone to adaptation effects. On the other hand, the addition of a forced-choice procedure allows for implementing more comparisons that can be done in a shorter time frame, but the number of conditions that can be used will be limited when comparing all possible combinations (Fotios & Houser, 2013). However, there are certain biases in forced choice procedures that need to be addressed (Fotios & Houser, 2013). Firstly, interval bias, where there is a consistent asymmetry in the response due to temporal position of the stimuli. The previous scene will influence what is perceived at the next scene. This occurs when scenes are presented sequentially instead of simultaneously. Secondly, position bias, where there is a consistent asymmetry in the response due to the spatial position of the stimuli. This occurs when a specific scene is always presented in the same position (e.g. left). This might influence the judgment of the participant. Thirdly, centering bias, related to the responses centering around the midpoint of the range of stimuli. Participants possibly may assume that the stimuli are centered around a midpoint, while this is not necessarily the case. This can result in participants being biased towards judging about half of the stimuli as brighter and about half of the stimuli as dimmer.

1.6. Virtual Reality

VR has been widely used in perception research in a variety of fields thanks to its versatility, immersiveness and its ability to allow for quick assessments. Unlike other representation methods, with stimuli integrated in the virtual environment, VR is comparable to the real-world in many aspects and can be used to test several stimuli that appear in real-world environments (Moscoso et al., 2021). VR has gained popularity in lighting research over the past few years, both for daylighting and electric lighting applications (Scorpio et al., 2021).

1.6.1. Advantages and limitations of VR

One of the main advantages of VR as a tool is that it allows for a completely controlled environment which limits external influences. It enables the researcher to easily isolate, add or remove stimuli, depending on what is of interest to the researcher (Bellazzi et al., 2022). For instance, it gives the possibility to have control over daylighting, which can be challenging in real-world scenes. Moreover, VR can show photo-realistic immersive scenes with the possibility for the user to move around in, which is a great advantage compared to other visual representation methods (Scorpio et al., 2021). Another advantage of VR is that it is very flexible. The manipulations are not limited to the physical environment and existing instruments, but VR creates the opportunity to show a big variety of stimuli without the expense of purchasing or developing new apparatus. Additionally, VR gives the possibility to test multiple variations without boundaries of time and physical space (Heydarian & Becerik-Gerber, 2016).

Although VR has great opportunities, the realism is limited due to imperfections in software and hardware. Even highly realistic renderings still lack the amount of detail to appear indistinguishable from real scenes. For instance, the luminance range of a display is limited, making it challenging to reproduce real-life lighting stimuli. Furthermore, the tasks participants can perform in VR are limited to what is physically possible in a VR environment. Only tasks that can be done with the sensors in the headset are possible. For example, it can be challenging to touch or pick up items in VR.

1.6.2. Accuracy of brightness judgements in VR

The quality and accessibility of VR as a research tool has increased, thanks to recent developments in VR headsets. However, understanding what can and cannot be translated between virtual and physical environments is important to evaluate the opportunities and limitations of their application. Several studies have compared the perceptual accuracy of virtual scenes to real-world scenarios and have shown that judgements of brightness between real and VR-scenes are largely accurately comparable (Abd-Alhamid et al., 2019; Chen et al., 2019; Hegazy et al., 2020; Rockcastle et al., 2021; Jin et al., 2022). Nevertheless, Rockcastle et al. (2021) found a larger spread of the perceived brightness in VR scenes than in real scenes. This indicates that there are some individual differences of brightness perception in VR. Furthermore, the results show that in some cases – mainly in dimly lit or highly contrasted scenes – the perceived brightness was significantly higher in the VR scene than in the real scene. This finding could have been due to inaccurate tone-mapping. On the other hand, Jin et al. (2022) compared perceptions of brightness, glare, spaciousness, and visual acuity between a real-world environment and both a photographed VR-environment and a rendered VR-environment. The results showed no significant difference of brightness and glare for a rendered-VR scene, while the photographed VR-scene was perceived as significantly brighter and more glaring than the real-world scene, but they were unable to explain these differences.

However, the perception of brightness in VR has its limitations. Firstly, the luminance range that can be displayed is limited due to restrictions within the hardware of the head-mounted display (Moscoso et al., 2021). Secondly, the eye adapts to luminance variation depending on the viewing direction within the scene, but the current unavailability of dynamic tone-mapping algorithms constrains the ability to mimic the adaptability of the human eye (Moscoso et al., 2021). In order to carefully generalize results to real-world scenarios, additional testing is necessary to establish the accuracy of brightness perception in VR.

1.6.3. Methods in VR

Indoor lighting has been investigated in VR scenes with daylighting, electrical lighting and combinations of the two. In previous studies, photo-realistic visualizations have been created by either photographing an existing space (Rockcastle et al., 2021) or rendering (Chamilothori, et al., 2019). The advantage of rendering is no dependence on an existing space that fulfills the necessary aspects for the study.

Assessment of different scenes in VR has been done in various ways, including questionnaires provided in VR (Hegazy et al., 2021), verbal questionnaires (Chamilothori et al., 2019; Abd-Alhamid et al., 2019; Rockcastle et al., 2021), physiological data (Chamilothori et al., 2022), adjusting settings (Wong et al., 2019) or forced-choice discrimination (Fotios & Houser, 2013).

For a review of different methods used in VR, see Bellazzi et al. (2022). The advantage of rating individual scenes with verbal questionnaires is that participants are not restricted to push any buttons or move within the VR environment. However, this method can induce observer bias, where the answers a participant gives may be influenced by possible expectations and hopes of the experimenter. The method of adjustment has the great advantage of absolute results that can be directly applicable, but this method can be hard to implement in a VR environment as it requires a dynamic scene instead of a static scene. To directly apply the results obtained from an adjustment method, it first needs to be investigated to what extent the results are generalizable to the real-world. Another useful method is forced-choice discrimination. As discussed in section 1.5.2., the advantage of this method is the quick assessment which allows comparing multiple different scenes to each other in a short time frame. Furthermore, this method does not require a dynamic scene and can utilize a static scene.

1.7. Purpose

Previous research has indicated that the degree of uniformity in the lighting distribution has an influence on the brightness of a room. However, the literature is contradictory and inconclusive. Although previous studies have tried to predict the brightness of a space, there are no true metrics of brightness and uniformity (Abboushi et al., 2022). This makes it difficult to compare results of different studies. They can only inform us about a general trend. Contrasting results in the literature can be due to flaws in the measurement of brightness and/or uniformity or due to a missing critical parameter that has not been investigated previously. As discussed in section 1.2., brightness sensation is a highly complex topic that can be influenced by multiple parameters, therefore this study limited the factors and only focused on the lighting distribution on one wall specifically. Different parameters in the lighting distribution have not been investigated sufficiently and independently. Therefore, the current research investigated how different lighting patterns, resulting in differences in uniformity contribute to judgements of brightness, leading to the following research question:

RQ: "*How does the uniformity of the lighting distribution on a wall, created by electrical light sources, affect the brightness of a room?*"

The uniformity is dependent on several different lighting parameters in a room, including the number, spatial distribution, and type of luminaire (specifically the beam shape of the light). It is hypothesized that these parameters are the most influential on the brightness of a room, based on results from previous research (Tiller & Veitch, 1995; De Vries et al., 2022; Kirsch, 2015; Stokkermans et al., 2018; Hsieh, 2012). The lighting distribution in an office is typically symmetrical with relatively uniform and downward directed spots. Manipulating the spatial distribution of these spots includes changing the number of spots used and the beam width of the spots, giving the following sub questions:

1) "*How does the number of light sources on the wall influence brightness?*"
2) "*How does the beam width of the light source influence brightness?*"

The majority of the existing literature suggests that uniformly lit spaces appear brighter than non-uniformly lit spaces. Therefore, it is expected that more uniform lighting distributions appear brighter than non-uniform lighting distributions. Thus, brightness is expected to increase with adding light sources and widening the beam of the light source.

Research by Loe et al. (1994) found that the average luminance within the 40 degrees horizontal band was a good predictor of brightness. On the other hand, research by De Vries, Heynderickx and De Kort (2022) found that both the logarithm of the median over a 60 degrees horizontal band and ratio of the logarithm of the 95th percentile/median over a 40 degrees horizontal band were good predictors of brightness and better than the average, as suggested by Loe et al. (1994). However, the lighting distributions in the studies of Loe (1994) and De Vries (2022) were relatively uniform, partially due to the usage of an entire room. In such conditions, the mean and median luminance are quite close to each other. It is unsure how these metrics predict brightness in less uniform illuminated conditions where the mean and median luminance are further apart.

The current research is interested in which of these two metrics predicts brightness best and therefore varies these as a third factor, leading to a third sub question:

3) "*Does the mean or median luminance provide a better prediction of brightness?*"

It was hypothesized that the median luminance is a better predictor for brightness, due to the changes in the mean, median and maximum luminance when comparing a non-uniform to a uniform lighting distribution. This is easiest illustrated with an example.

Considering the example of Figure 6, comprised of a uniform and a non-uniform lighting distribution with equal mean luminance of 49 cd/m^2 (Table 1, top two rows), the non-uniform lighting distribution will have a higher maximum luminance than the uniform lighting distribution, but a lower median luminance than the uniform lighting distribution.

Table 1.
Two sets of hypothetical lighting distributions (uniform vs non-uniform) containing either equal mean or median luminance, with corresponding mean, median and maximum luminance values in cd/m^2.

		Mean luminance	Median luminance	Maximum luminance
Equal mean luminance	Uniform	49.0	50.1	102.4
	Non-uniform	49.0	31.0	175.2
Equal median luminance	Uniform	37.9	38.5	86.2
	Non-uniform	57.4	38.5	178.0

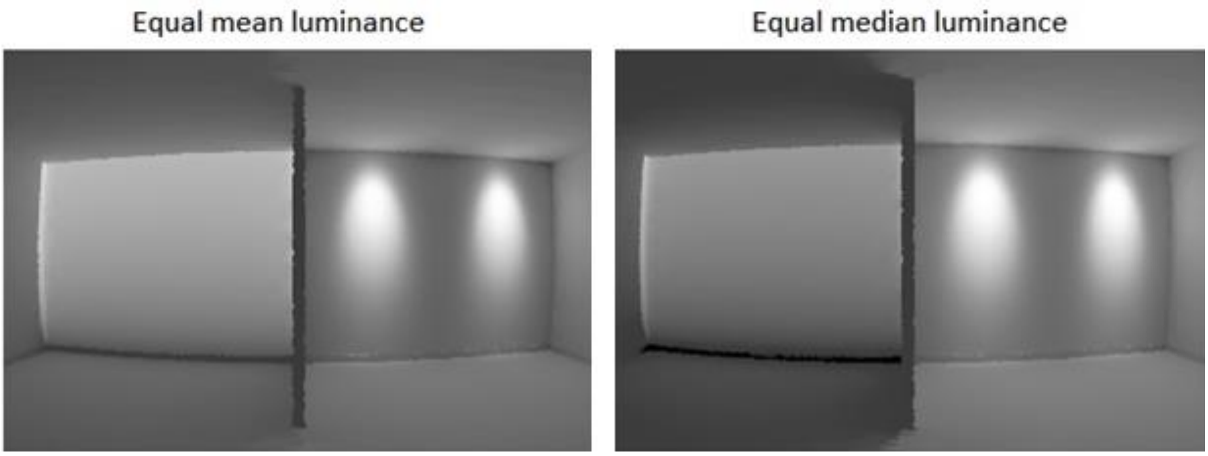


Figure 6. Visualization of two sets of hypothetical uniform (left on both sides) and non-uniform (right on both sides) lighting distributions with an equal mean luminance (left) and equal median luminance (right) condition.

Assuming the uniform lighting distribution would appear brighter, in the equal mean luminance situation (Table 1, top two rows), the increase in the median (from non-uniform to uniform) would support the assumption that uniform is perceived as brighter. However, if the example were to be based on the equal median luminance situation (Table 1, bottom two rows), the decrease in the mean (from non-uniform to uniform) would not support the assumption, leading to our hypothesis that the median is a better predictor than the mean.

If the two lighting distributions were equal in median luminance (Figure 6, right; Table 1, bottom two rows) and the uniform distribution would still be considered brighter, the hypothesis of a positive relation between uniformity and brightness would be confirmed.

To answer these research questions the current study employed a two-alternative forced choice procedure in virtual reality with computer rendered scenes. As discussed in section 1.6., VR is a great tool in research as it can accurately reproduce real-world stimuli. Moreover, VR is flexible and allows for a large variety of stimuli. The two-alternative forced choice (2AFC) paired-comparisons procedure is chosen due to its advantage of quickly assessing different scenes.

2. Method

2.1. Design

The study followed a two-alternative forced choice (2AFC) within-subject design. Two independent sets of stimuli were created where in one the mean luminance and in the other the median luminance was kept constant. Both sets were manipulated on two three-level factors: number of light sources (two vs. three vs. four – see details in section 2.4.3.) and beam width (narrow vs. medium vs. wide). Additionally, a control condition of a linear uniform gradient was added. This resulted in a total of 20 conditions.

The experiment investigated the influence of the lighting distribution on the brightness of a scene. In a virtual reality environment, participants were asked to make a forced choice as to which of two adjacent walls, divided by a small divider, appeared brighter. To account for position bias, each scene was presented twice, where the position of the conditions was swapped between left and right. This resulted in a total of 180 comparisons. The study protocol was approved by the Ethical Review Board if the HTI group at TU/e and the Signify Internal Studies with Human Subjects assessment.

2.2. Participants

A total of 28 participants (20 males and 8 females; age range 18 – 64) were internally recruited at Signify. A comparable number of participants is used in studies with a similar method (Ratliff et al., 2019; Kinzuka et al., 2021).

Participants were interns or scientists in an unrelated area and are therefore assumed not to be influential on the results. All participants had normal or corrected-to-normal vision. A test with a short version of the Ishihara showed that none of the participants exhibited a color-vision deficiency. Finally, epilepsy and sensitivity to nausea and/or vertigo were exclusion criteria, because of potential risks due to the usage of virtual reality (Nichols & Patel, 2002).

2.3. Apparatus

2.3.1. VR headset

The VR headset used in the experiment was the Oculus Quest 2. This VR headset has a resolution of 1832x1920 pixels for both eyes (3664x1920 pixels in total), a 2.2 gamma, refresh rate up to 120 Hz, field of view of 89°, peak brightness of 100 nits (equal to 100 cd/m^2) and xy-standard color space primaries: R (0.640, 0.330), G (0.292, 0.586), B (0.156, 0.058) and W (D65 (0.313, 0.329)) (Oculus Developers, n.d.). A luminance measurement (as discussed in section 2.4.7.) in the headset showed a maximum luminance of around 78 cd/m^2 in the current study.

2.3.2. Application

For the experimental procedure, a web application was created with WebXR Layers API using openGL (Cabanier, 2023). This application (discussed in more detail in section 2.4.6.) showed the 180 scenes one by one to the participant. The order of these images was completely randomized. Which scene was shown (including which condition appeared on which side), the choice of the participant (left or right) and the response time was logged in a JSON datafile.

When participants made a choice, by clicking the side button on the left or right controller, the next scene was presented and the timer was reset. To account for interval bias, a short fade-to-black (0.5 seconds fade out, 0.5 seconds pause, 0.5 seconds fade in) was added in between the scenes, in order to limit the influence of the previous scene on the judgments of the next scene.

2.4. Stimuli

To control for the different mechanics in brightness perception, discussed in section 1.2, the scenes were designed to prevent brightness perception effects as much as possible. For example, lightness constancy is attempted to be assured by keeping the color and reflectivity constant across all of the walls and throughout all of the scenes. Moreover, the appearance of the surrounding environment and the baseboards were identical in all of the scenes, giving the appearance of a constant lightness within the room. Moreover, no sharp lines and strong contrasts are used in the lighting distributions. Furthermore, glaring sources are excluded and the differences in luminance between scenes are limited to prevent adaptation effects.

The stimuli consist of grayscale stereoscopic HDR renderings, which were created with the software Radiance 5.4a (2023-01-19) 4a16c6c (Ward, 1994). Radiance utilizes raytracing to create visually realistic lighting scenes, allowing to create simulations and do calculations.

2.4.1. Test Room

A gray room with dimensions 8.1x9x3m and reflectivity of 60% was created with functions *genbox* and *xform*. Secondly, to divide the room into two separate rooms, a small divider was placed with dimensions 0.1x4x3m resulting in two test rooms with dimensions 4x3. A floor plan of the room can be found in Figure 7. The red square indicates the location of the observer.

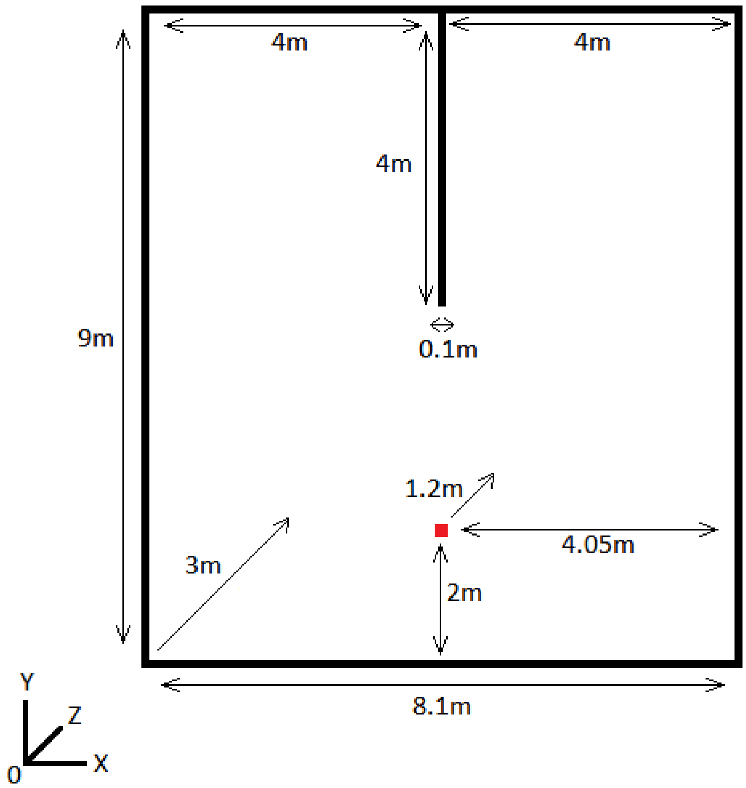


Figure 7. Floor plan of the room with dimensions 8.1x9x3m. A 4x0.1x3m long dividing wall was placed to separate the room into two rooms. The location of the observer is marked with a red square at x = 4.05m, y = 2m and z = 1.2m.

Thirdly, to enhance the realism of the room, 0.1m high baseboards with a reflectivity of 50% were added along the walls. An example low quality scene with two conditions (two narrow and three narrow spots) can be found in Figure 8, created with Radiance function *rpict*.

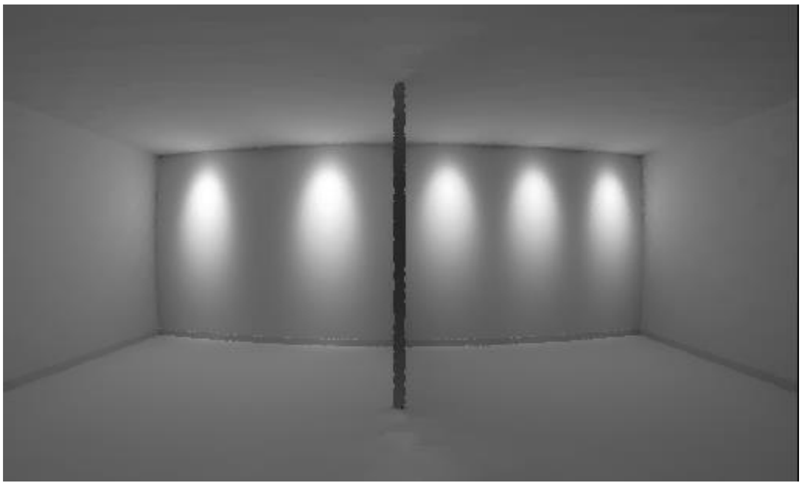


Figure 8. Example low quality scene, created with Radiance function *rpict*, showing two walls with differing lighting distribution, divided by a wall. These are the conditions two narrow spots and three narrow spots with equal median luminance.

2.4.2. Background luminance

As can be seen in Figure 8, the lighting from the different spots reflects onto the surrounding walls, ceiling and floor. The amount of light falling on the surrounding surfaces can influence how the test wall itself is perceived. To control for these reflections and variations within the room, a mask was created by placing a total of 20 (4x5) luminaires on the ceiling, as well as on the floor to create a uniform and symmetrical background lighting in the room. Results from function *evalglare* show a mean and median luminance of around 27 cd/m^2 on the surrounding walls. Note that no distinction was made with regards to the RGB values of the lighting as correlated color temperature was not a factor in this experiment, resulting in all light being set to white. The luminaire used for the general lighting was the Philips PowerBalance gen2 recessed luminaire RC461B PSD W60L60 with light source 1xLED34S/940. Tone-mapping (discussed in section 2.4.5.) of the masking was based on the same histogram as the rest of the scenes, however, no gamma correction was applied to the masking image. This mask was applied over the renderings to cover everything in the room (including plinths) except the test walls in Adobe Photoshop 2018 CC (version 19.1.2). This resulted in the background space being constant throughout the whole experiment, such that only the test walls varied.

2.4.3. Conditions

Two sets of 10 conditions have been created with three types of luminaires and three different numbers of spots. The luminaire used for the narrow spot was the Philips TrueFashion projector ST714T FPO18 with light source 1xLED27S/PC930. This luminaire has a beam width of 18° (FWHM). Next, the luminaire used for the medium spot was the Philips TrueFashion projector ST714T FPO36 with the light source 1xLED27S/PC930 and a beam width of 36° (FWHM). Furthermore, for the wide spot the Philips CoreLine Downlight DN140B PSED-E IP54 D162 WR with light source 1xLED10S/830 was used, containing a beam width of 84° (FWHM). Finally, the uniform gradient was manually created such that it linearly decreased in luminance going from top to bottom, while maintaining approximately an equal mean and median luminance.

The narrow and medium spots were placed 0.5m from the wall at 3m height. The wide spot was placed at 2.7m height to match the location of the beam on the wall with the other two This resulted in two sets of 10 conditions, which can be seen in Figures 9 and 10 for the equal mean and equal median luminance sets before masking, respectively.

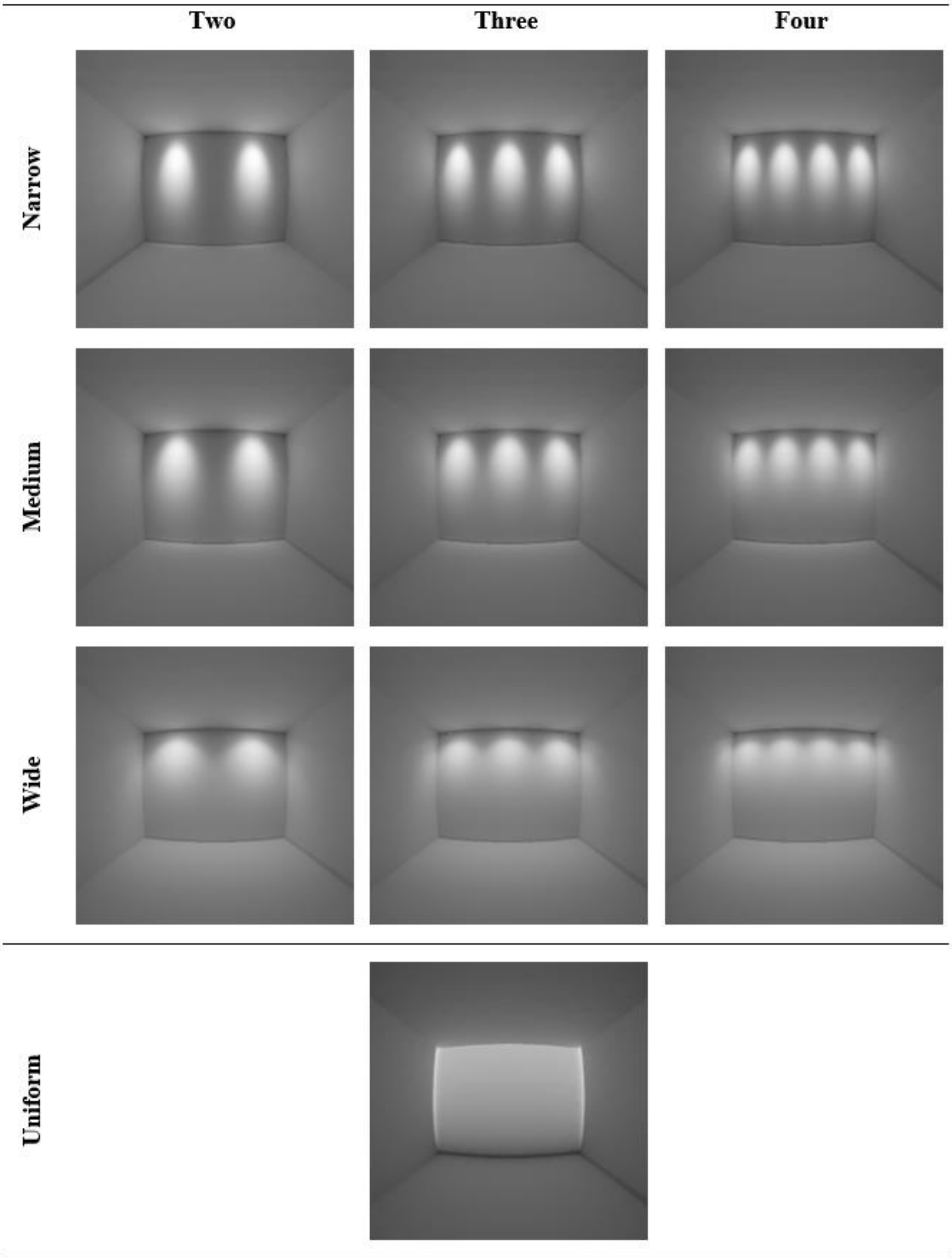


Figure 9. The 10 conditions used in the equal mean luminance set of conditions before masking. Created with Radiance function *rpict*

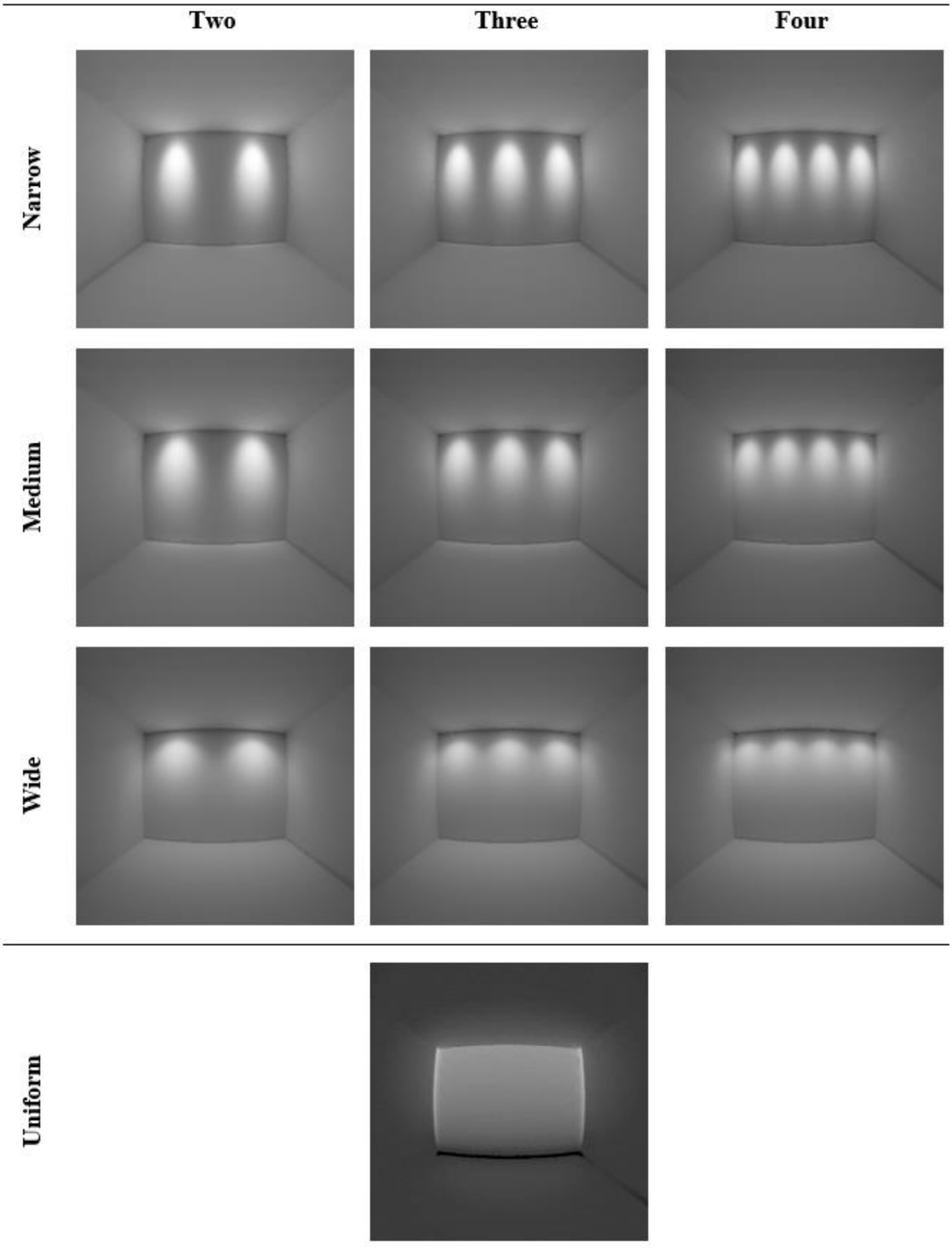


Figure 10. The 10 conditions used in the equal median luminance set of conditions before masking. Created with Radiance function *rpict.*

The image sets with equal mean and equal median luminance were created by manually dimming the luminaires in the different conditions. From the tone-mapped HDR images (discussed in section 2.4.5) the mean, median, minimum and maximum luminance of each of the two walls was calculated with function *evalglare*. Based on the results from evalglare, the dimming factors of the luminaires in the equal-median set were adapted until the median was equal in all conditions. These final dimming factors (-*m*) and their corresponding evalglare output can be found in Table 2. Remarkably, the most non-uniform lighting distribution (two narrow spots) has the highest minimum luminance of the set. This could be due to the increased scaling of this condition. As can be seen in Figure 12, the lighting in this condition reaches further down the wall than the lighting does in the other conditions, therefore more of the lighting reaches the corners. Hence, less of the walls is covered in this conditions, giving this condition a relatively large dark area, but has a much higher peak than the rest of the conditions.

Table 2.
Dimming factors (-*m*) – that were calculated resulting from Radiance function *evalglare* output for the equal median set of conditions – and their corresponding mean, median and maximum luminance values (in cd/m^2) for each condition.

Number of spots	Beam width	*-m*	Mean luminance	Median luminance	Minimum luminance	Maximum luminance
Two	**Narrow**	3.855	57.4	38.5	12.2	178.0
Three	**Narrow**	1.945	55.9	38.5	9.9	172.4
Four	**Narrow**	1.13	52.3	38.5	8.3	154.9
Two	**Medium**	2.325	53.6	38.5	8.1	160.5
Three	**Medium**	1.13	49.0	38.5	8.1	139.5
Four	**Medium**	0.75	47.5	38.5	8.1	124.8
Two	**Wide**	3.99	47.1	38.5	7.8	173.8
Three	**Wide**	2.46	46.3	38.5	8.1	150.7
Four	**Wide**	1.815	46.3	38.5	8.1	156.3
Uniform		1.47	37.9	38.5	10.5	86.2

To enable linking the equal median and equal mean sets, one of the conditions was used as a starting point for the other set of conditions. To limit the luminance range for preventing adaptation effects, the average mean luminance in the equal median set was calculated (49 cd/m^2). The condition in the set with a mean luminance value closest to 49 cd/m^2 was the condition with 3 medium spots. Therefore, the mean from the condition with 3 medium spots is used as the target mean for the equal mean luminance set of conditions.

In the same way as in the equal-median set, the dimming factors in the equal-mean set were adapted until the output from evalglare showed an equal mean in all conditions. These final dimming factors (-*m*) and their corresponding evalglare output can be found in Table 3.

Table 3.
Dimming factors (-*m*) – that were calculated resulting from Radiance function *evalglare* output for the equal mean set of conditions – and their corresponding mean, median, minimum and maximum luminance values (in cd/m^2) for each of the different conditions.

Number of spots	Beam width	*-m*	Mean luminance	Median luminance	Minimum luminance	Maximum luminance
Two	**Narrow**	2.793	49.0	31.0	8.1	175.2
Three	**Narrow**	1.507	49.0	32.3	8.1	163.3
Four	**Narrow**	0.997	49.0	35.4	8.1	152.1
Two	**Medium**	1.963	49.0	34.2	8.1	156.3
Three	**Medium**	1.13	49.0	38.5	8.1	139.5
Four	**Medium**	0.791	49.0	39.9	8.1	124.8
Two	**Wide**	4.279	49.0	39.9	8.1	173.8
Three	**Wide**	2.702	49.0	40.6	8.1	153.5
Four	**Wide**	1.993	49.0	40.6	8.1	157.7
Uniform		2.195	49.0	50.1	16.9	102.4

These dimming factors were used to tune the output of the different luminaires to create all of the 20 different conditions. Note that within the scope of this research, no distinction was made with regard to the brightness related to the span of the horizontal band.

In terms of uniformity, increasing the number of luminaires and widening the beams results in higher wall coverage and lower extremes, and is therefore assumed to increase uniformity. Thus, the two narrow beams condition is assumed to be the least uniform and the uniform control condition is assumed to be the most uniform of the 10 distributions.

False color renderings were created to check the luminance range and the change of peak luminance values. Figures 11 and 12 show the renderings for the equal mean and median luminance set of conditions, respectively. As can be seen, the maximum luminance decreases as more luminaires are added or when the beam widens.

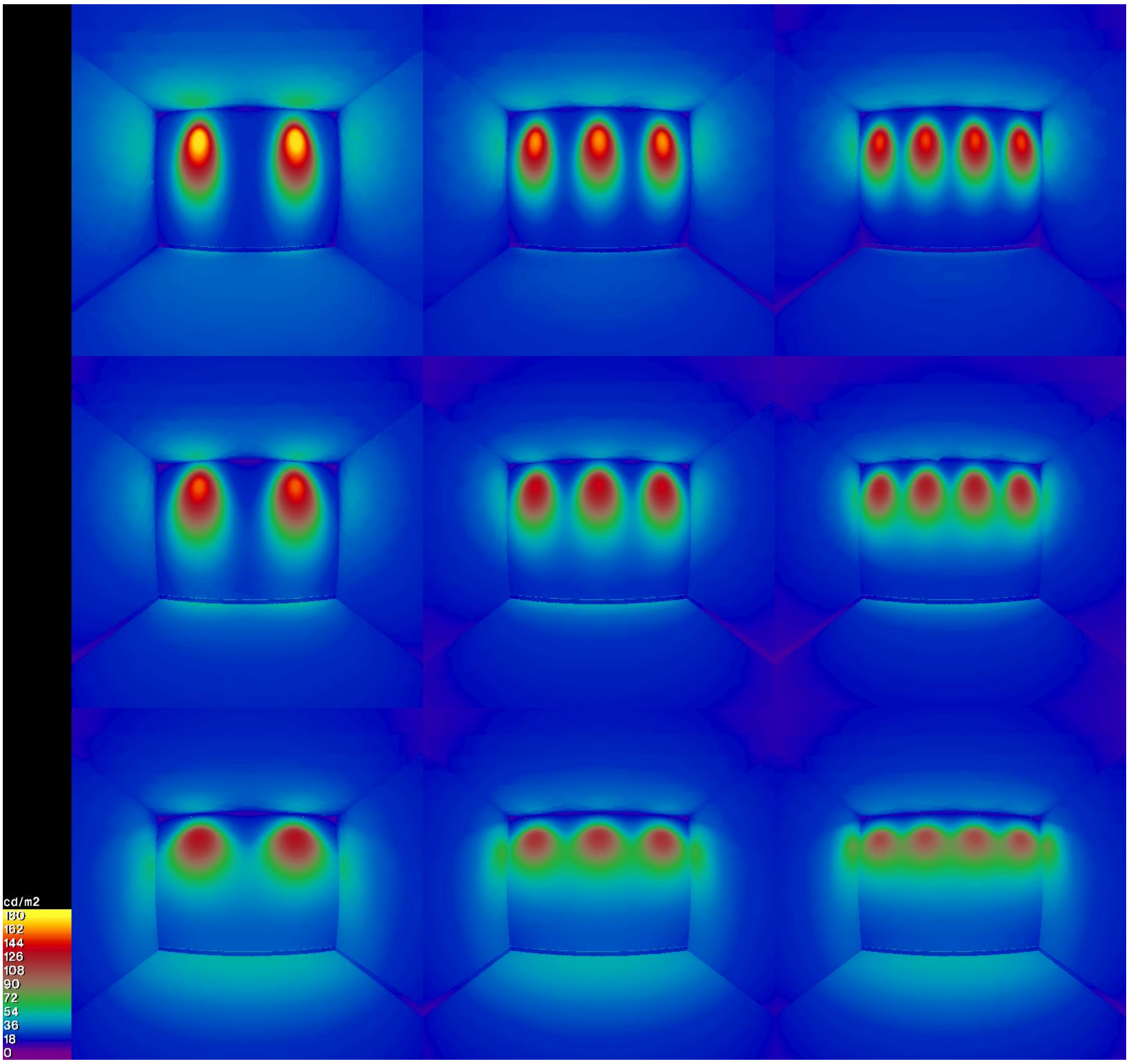


Figure 11. False color renderings for the equal mean luminance set of conditions before masking, showing that the peak luminance varies over the image sequence and an overall lumiance range of 0 to 180 cd/m^2. Created with Radiance function *falsecolor*.

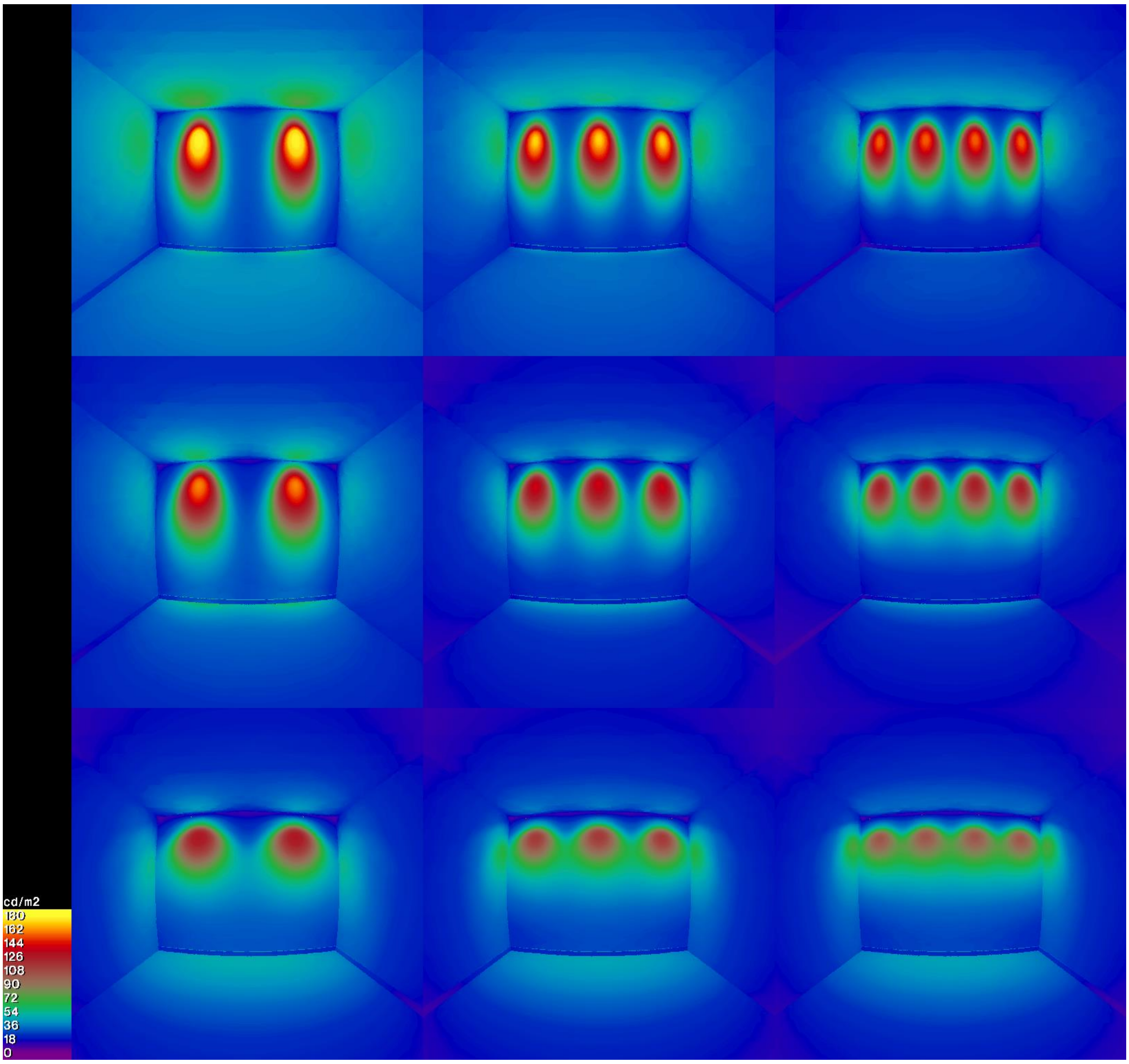


Figure 12. False color renderings for the equal median luminance set of conditions before masking, showing that the peak luminance varies over the image sequence and an overall luminance range of 0 to 180 cd/m^2. Created with Radiance function *falsecolor.*

2.4.4. Renderings

A total of 180 4K (4096x4096) 360° stereoscopic HDR images with equirectangular projection were rendered with the Radiance software using the function *rtrace*, with the script *stereo_sphere.cal* (based on Google Inc., n.d.) and rendering parameters set to default, except -*ab* = 6, -*ad* = 4096, -*as* = 2048, -*aa* = 0.15 and -*ar* = 1024. Renderings were created with a viewpoint in the middle of the room (x = 4.05m), 7 meters away from the test walls (y = 2m) and at sitting height (z = 1.2m). These renderings included the information for both the left and right eye with an inter-pupillary distance of 63mm. These two equirectangular projections were stacked on top of each other in one rendering. An example output rendering can be found in the top left of Figure 14. Section 2.4.6. discusses how ow these renderings are projected in the VR environment.

2.4.5. Tone-Mapping

A high-dynamic range image (HDR) has a larger amount of brightness information than a regular image, allowing for more realistic representations of scenes. Most displays have a limited dynamic range and cannot accurately reproduce the full HDR content. With tone-mapping the dynamic range of the HDR image is mapped to the limited range of the display. It is important to consider the human-visual system in this process to make the scene appear realistic. The total tone-mapping process for the renderings in this study is visualized in a flowchart in Figure 13.

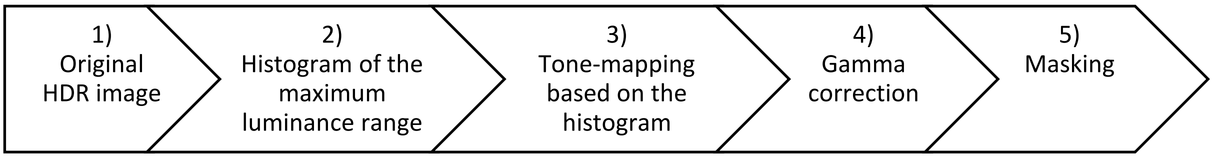


Figure 13. Tone-mapping process visualized in a flowchart. 1) Starting with the original HDR image, created with function *rpict*, a 2) histogram of the maximum luminance range was identified with function *phisto*. 3) This histogram was used to tone-map the scene with function *pcond* and 4) to increase the overall brightness of the room a gamma correction of 4 was applied with function *ra_tiff*. 5) Finally, a mask was applied to cover everything except the test walls.

1) With Radiance function *rpict* HDR images of each condition were created. These HDR images were used to check the initial luminance measurements of the walls for the different conditions. As can be seen (Figure 14, top left) the display is unable to show the full luminance range, therefore each luminance value higher than the maximum of the display is set to maximum.

2) From the original HDR images the maximum luminance range was identified. The image with the highest peak luminance was chosen to create a histogram with function *phisto*. This was the pair of the conditions with two and three narrow spots in the equal median set. This image had the widest range of luminance values (7 to 6263 cd/m^2).

3) Based on the histogram of this first image, all the images were tone-mapped with function *pcond* (Figure 14, top right). This function was found to represent brightness with an acceptable accuracy (Chamilothori, 2019). The option *-h* was used to mimic the human visual response to the scene. However, as now only the peak luminance is set to the maximum, the tone-mapping process caused the rest of the luminance in the scene to be scaled down making the images appear relatively dark.

4) To increase the overall brightness of the image, it was chosen to increase the gamma correction in the images. This was accomplished with Radiance function *ra_tiff* with a gamma correction of 4 (Figure 14, bottom left). A gamma correction of 4 was chosen based on visual inspection, giving the room a visually bright appearance corresponding to an office room. Note that a gamma correction is a non-linear adjustment and may have changed the intended output of the tone-mapping process and may have no longer resembled the human visual system accurately. With *ra_tiff* these TIF files were converted back to HDR in order to calculate the final luminance values with function *evalglare*.

These HDR images were then converted into a PNG image with IrfanView software (version 4.62, Skiljan, 2012). No tone-mapping or correction was applied in this conversion.

5) Finally, as mentioned in section 2.4.2, a mask of background lighting was placed on the images using Adobe Photoshop CC 2018 (version 19.1.2). The final result can be seen on the bottom right of Figure 14. A full-size figure of the final example rendering can be found in Appendix I.

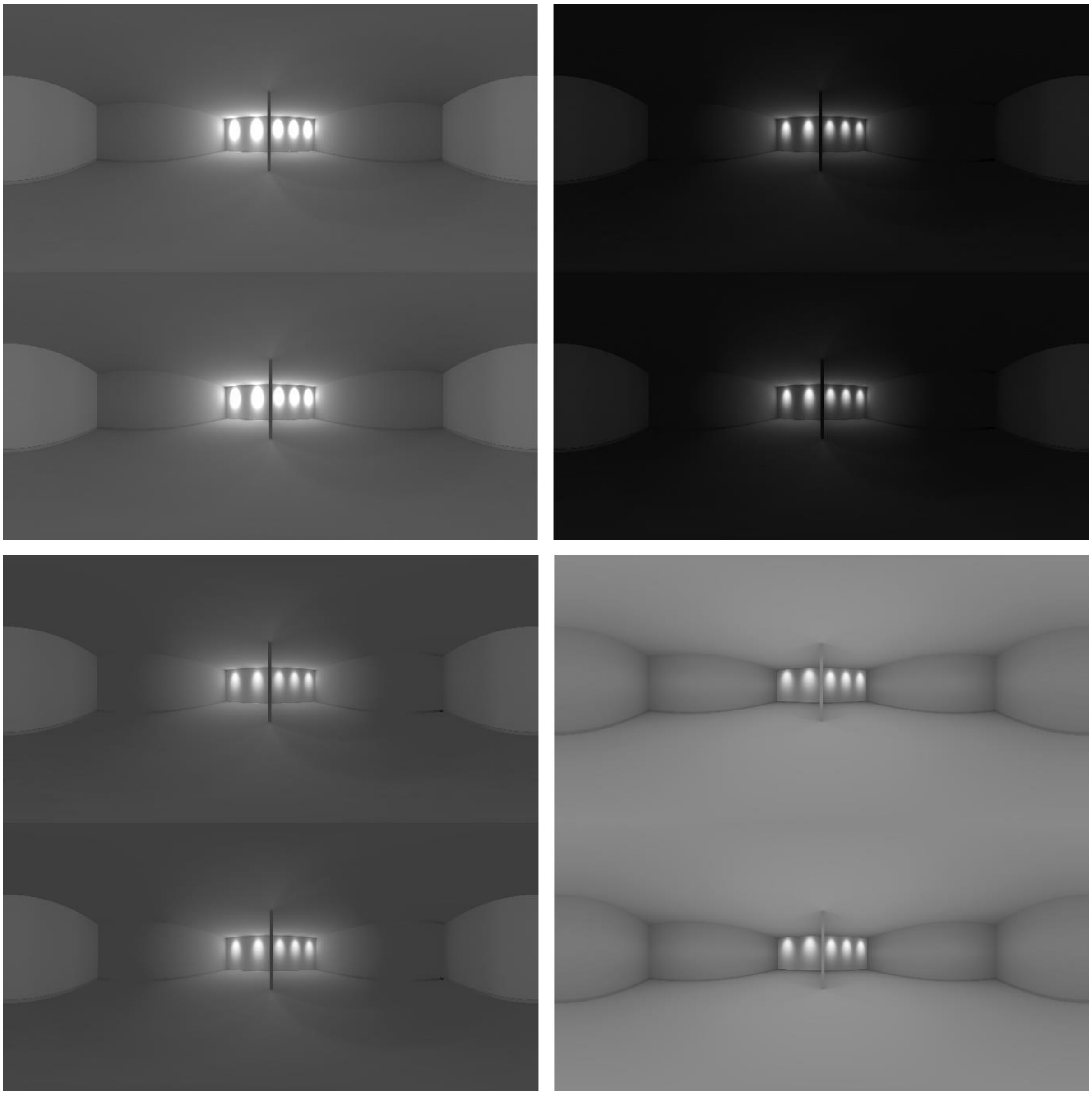


Figure 14. Stacked 360° Stereoscopic HDR renderings of an example condition (two narrow beams vs three narrow beams with equal median luminance). Top left: rendering directly from Radiance before tone-mapping. Top right: rendering after tone-mapping with *pcond*. Bottom left: rendering after gamma correction. Bottom right: rendering with mask applied.

2.4.6. Projecting in VR

The scenes are projected in VR with an application developed with the WebXR Layer API using openGL (Cabanier, 2023). Upon initialization, an XREquirectLayer was created with the resolution of the images, which is reused for the rest of the run-time. This layer mapped the equirectangular projection into the inside of a sphere. The renderings consisted of two equirectangular projections stacked on top of each other. In order to be projected in VR, these are split into two textures and mapped into two spheres, one for each eye. After a new equirectangular texture was loaded, it was sent to the GPU directly. There were no color-space conversions as there was no processing on the textures.

2.4.7. Luminance Check

The actual luminance that was being perceived by the observer when the stimuli were projected in the VR headset was measured with a luminance camera (LMK5 Color, using a fisheye lens), centering each condition in the camera's field of view. An example luminance photograph of the condition with 3 medium spots can be found in Figure 15. The luminance photographs for all the conditions can be found in Appendix II. A rectangular measurement grid was placed on top of the wall to calculate the mean, median and maximum values from each condition. The results showed the maximum luminance was around 78 cd/m^2. Moreover, the results indicated that the mean was around 24 cd/m^2 in the equal mean set and the median was around 20 cd/m^2 in the equal median set. The actual measurements for all conditions can be found in Table 4. Note, in the equal median set, the median has 1.6 cd/m^2 error. This can be due to measurement flaws, or the translation between screens. However, this error is assumed to be negligible as it is less than 10%.

The luminance differences within the sets were a little smaller than expected. For instance, the ratio of maximum to mean ratio in the equal mean set was 3.57 in the simulation and 3.21 from the measurement, indicating that the luminance differences in the scenes are about 11% smaller than expected. However, this is assumed to be negligible as differences between scenes can still be seen when visually inspecting the conditions in the VR headset.

Table 4.
Luminance measurements from the luminance camera for the equal mean (left) and equal median (right) sets of conditions in cd/m^2.

Number of spots	Beam width	Equal Mean Luminance				Equal Median Luminance			
		Mean	Median	Min.	Max.	Mean	Median	Min.	Max.
Two	**Narrow**	24.1	16.7	7.0	77.3	28.1	20.7	9.1	78.1
Three	**Narrow**	23.7	17.2	7.3	70.8	27.2	20.3	9.0	75.8
Four	**Narrow**	24.0	18.7	7.5	67.8	25.6	20.2	8.2	69.2
Two	**Medium**	24.2	18.3	5.9	68.4	26.0	19.9	6.8	70.9
Three	**Medium**	24.1	19.9	6.9	62.6	24.1	19.9	7.1	63.0
Four	**Medium**	23.8	20.4	7.9	56.3	23.4	19.7	7.7	56.7
Two	**Wide**	24.1	20.7	6.0	58.1	23.4	19.9	5.8	57.7
Three	**Wide**	24.1	20.9	6.8	53.7	22.8	19.6	6.2	52.6
Four	**Wide**	24.1	20.9	7.0	49.3	22.9	19.6	6.6	47.9
Uniform		24.0	24.8	12.1	37.7	18.5	19.1	9.5	31.1

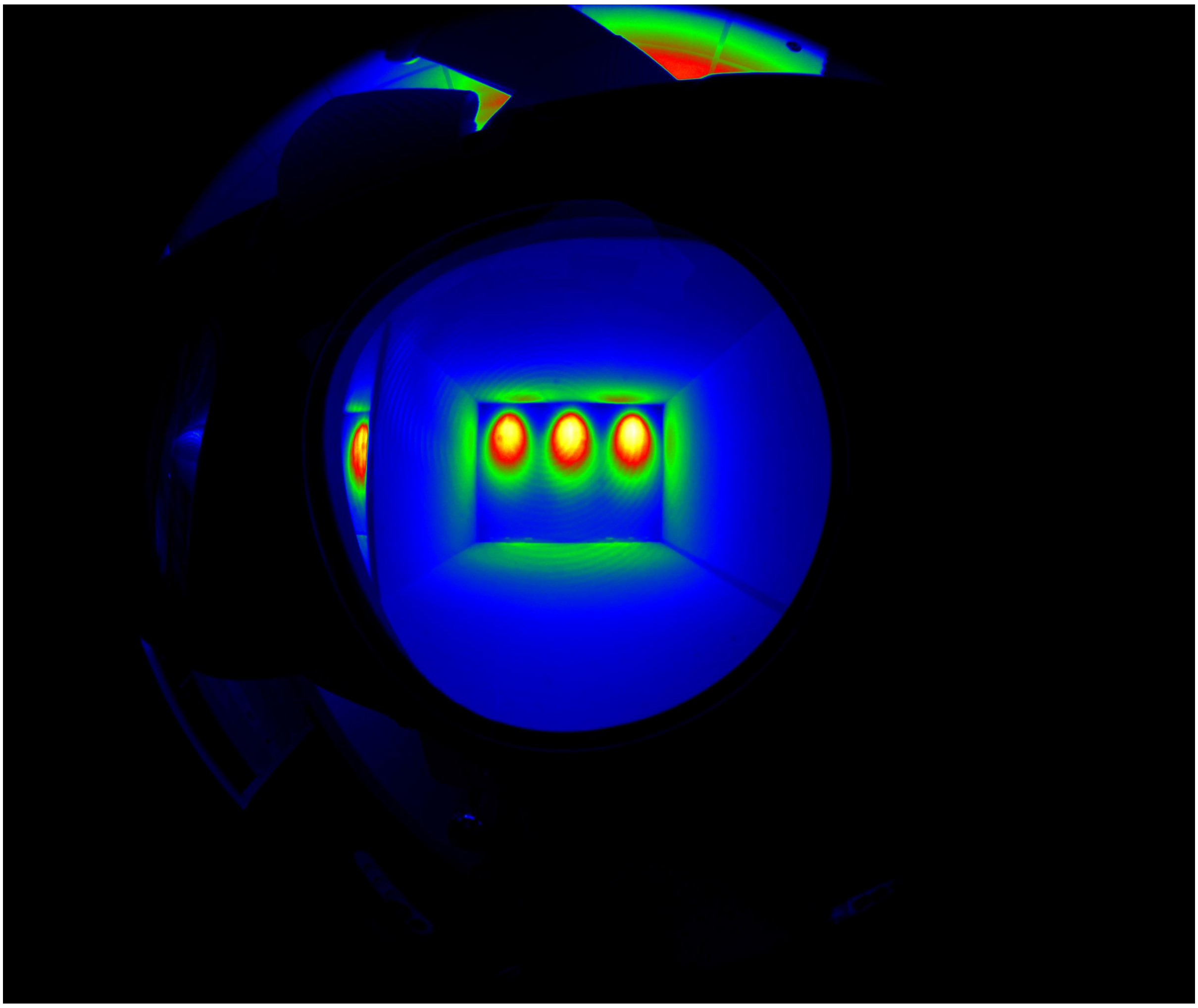


Figure 15. Example luminance photograph of the condition with 3 medium spots.

2.5. Measurements

In the two-alternative forced choice paired comparisons experiment, participants were presented with two walls with differing lighting distributions. Which scene was presented, and whether a condition appeared left or right was completely randomized. Participants were forced to choose which of the two lighting distributions appeared brighter by answering the question "*Does the left or right wall appear brighter?*". Participants could answer this question by pressing the side button on either the left or right controller, depending on their choice. As it was a forced choice, there was no option to indicate that the two conditions appeared equally bright. Furthermore, participants were instructed to answer based on their first impression and not to think about it for too long. The response times were recorded to enable examining the duration participants looked at each scene.

2.6. Procedure

After participants entered the room, they were given a short explanation of the experiment, a short version of the Ishihara test and the opportunity to ask questions. After informed consent was given, participants were seated on a rotating chair, given the VR headset, and given an instruction on how to wear and use it. The experiment started with a short test round of eight trials to adapt to the environment and to check whether the task was understood. These test stimuli were not from the stimulus set, instead they were created to show a more pronounced brightness difference. These test scenes were always in the same order. After this test trial participants were given the opportunity to ask any remaining questions before the actual experiment started. Throughout the experiment, participants were given the opportunity to take a small break if this was desired. The experiment, including instructions, had a total duration of around 30 minutes. After the participants were finished, they were given a last opportunity to ask questions and were thanked for their participation. Participants were not compensated for their participation.

2.7. Statistical Analysis

The analysis of the data was conducted with the many-facet Rasch model (MFRM) in Facets (Linacre, 2023) and additional tests were done in RStudio (Posit team, 2023). The MFRM uses the natural logarithm of the probability (P) of choosing one condition (X_1) over the other (X_2) by the difference in estimated brightness of the first (δ_1) and the second (δ_2) condition, which is calculated as follows (Linacre, 1997):

$$\ln \left(\frac{P(X_1 > X_2)}{1 - P(X_1 > X_2)}\right) = \delta_1 - \delta_2$$

The probability of choosing one condition over the other is calculated from the effect sizes, measured in log odd units (logits), which can be derived from the previous formula:

$$P(X_1 > X_2) = \frac{e^{\delta_1 - \delta_2}}{1 + (e^{\delta_1 - \delta_2})}$$

The model was used to find information about the expected probability of participants indicating one wall as brighter than the other, where one characteristic changed while the others remained constant. The model fits the objects on an interval scale by giving meaning to the distance between the conditions. The conditions themselves are nominal as there is no true order within the conditions. As there is no true origin, the data cannot be fitted onto a ratio scale. As a result, the data are relative, rather than absolute. The final measurement will indicate which condition is more likely than another to be judged brightest and by how much, but cannot tell something about the absolute information that can be directly applicable (Wright & Linacre, 1989), in this research being luminance information.

In order to compare the estimates of one characteristic between conditions, the estimates are required to have a common zero point. This is achieved by anchoring each characteristic, in this case, the number of spots and beam width. The 3 medium spots condition was used as an anchor, because this condition was present in both sets of conditions, allowing to plot the estimates from both sets on the same scale. This condition is anchored to a brightness score of zero and referred to as the reference condition from now on.

The MFRM method was chosen, because it gives detailed information about the difficulty of choosing between scenes. Another one of the advantages of the Rasch model is the possibility to gain insight into the performance of the participants, in order to understand how well the data from each participant fits into the model. As this is a perception test of physical attributes of the space, it is assumed that the results of the test should not differ between people. Thus, the person doing the test should not have an impact on the score. Therefore, the participant ID will be anchored to zero, allowing to identify when a participant had an influence on the estimation (Linacre, n.d.).

The Rasch model will transform the observations into linear measures. Fit statistics can help to understand to what extent the linear measurement confirmed unidimensionality and other necessary specifications (Wright & Linacre, 1989). When these are confirmed, the brightness measurements are measured on a true interval scale. The two different fit methods, Infit Mean Square (IMS) and Outfit Mean Square (OMS), were used to weigh the inliers and outliers, respectively. These fit statistics show when a person or item falls outside the majority of responses. (Wright & Linacre, 1996). A mean-square (MS) of 1.0 indicates a perfect fit, between 0.7 and 1.3 is a good fit, between 0.5 and 1.5 is reasonable and below 0.5 or above 1.5 is not meaningful. Furthermore, a mean-square above 2.0 is considered degrading (Linacre, 2002). Mean-squares below 1.0 indicate an overfit, suggesting the data are more predictable than the model expects, while mean-squares above 1.0 indicate a misfit, suggesting the data are less predictable than the model expects. A mean-square of 1.5 would indicate that there is 50% more randomness than expected. In other words, this could indicate random guessing or insensitivity (Wright, 1994).

Two separate Rasch analyses were done. One on condition level, to enable direct comparisons between the separate conditions, including the uniform control condition. On the other hand, an analysis on factor-level was included, in order to investigate the contribution of the two factors (number of spots and beam width) to the brightness judgements. A post-hoc z-test was done to do a pairwise comparison between levels of factors (Altman & Bland, 2003). A statistical significance level of $\alpha = 0.05$ was used. Note that the two estimates should be independent and therefore not be obtained from the same population. Although this study contains a within-subjects design and estimates are obtained from the same population, with a good data-to-model fit, scores from individuals can be assumed to not have influenced the estimates.

3. Results

3.1. Participant Fit Statistics

To interpret whether there were influential inter-personal differences within the data, the participant ID was added as a factor in the data analysis. The participant IDs were anchored to zero, as they were not allowed to influence the estimates. In the results from the equal mean luminance set of conditions only, there are three participants that exceed the 1.5 threshold for the IMS and OMS fit statistics. This suggests that these three participants do not fit the model very well and possibly have a differing opinion, used a different assessment method or did not understand the task completely. However, as these values do not exceed 2.0, they are not considered degrading. These results can be found in Appendix III.

3.2. Condition level

Before the factors were analyzed individually, brightness estimates were made on condition level. The results for this can be found in Table 5, and the full output Table can be found in Appendix IV. All items had very good fit statistics with MS-values ≤ 1.10. Figure 16 presents the vertical ruler map for participants' brightness perceptions, with the equal mean luminance set of conditions ($set_{mean)}$ on the left and the equal median luminance set of conditions (set_{median}) on the right. The acronyms 2, 3 and 4 represent two, three and four spots, and N, M and W represent narrow, medium and wide, respectively. This Figure presents an overview of the log odd unit (logit) scores for the individual conditions. Note, the condition of 3 medium spots was used as an anchor, because this condition was identical in both sets. Therefore, the other conditions are relative to this reference condition. A positive value indicates the condition was perceived as brighter than the reference condition and a negative value indicates the condition was perceived as dimmer than the reference condition. The larger the distance between two conditions, the higher the perceived difference in brightness between these two conditions.

Table 5.
Perceived brightness estimates for the two sets of 10 conditions (in logits) on condition level, standard errors of the estimate (SE) and fit statistics IMS and OMS (Infit Mean-Square and Outfit Mean-Square), with equal mean luminance on the left and equal median luminance on the right.

	Equal Mean Luminance				**Equal Median Luminance**			
	Brightness	SE	IMS	OMS	Brightness	SE	IMS	OMS
2 Narrow	-0.61	0.10	1.06	1.07	0.19	0.09	1.00	1.00
3 Narrow	-0.64	0.10	1.01	1.02	0.22	0.09	1.00	1.00
4 Narrow	-0.95	0.10	0.99	1.01	-0.38	0.09	0.98	0.97
2 Medium	0.05	0.09	0.95	0.95	0.64	0.09	0.98	0.97
3 Medium	0.00	0.09	0.98	0.98	0.00	0.09	1.01	1.01
4 Medium	-0.03	0.09	0.98	0.98	-0.27	0.09	1.00	1.00
2 Wide	0.27	0.09	0.95	0.95	0.27	0.09	1.00	1.00
3 Wide	0.27	0.09	1.01	1.01	0.27	0.09	1.02	1.02
4 Wide	0.43	0.09	1.00	1.01	0.21	0.09	1.01	1.01
Uniform	0.77	0.10	1.06	1.10	-0.78	0.10	1.01	1.01

To give an indication of the effect size, in the equal mean luminance set of conditions the condition with the biggest difference in brightness relative to the reference condition is the four narrow spots condition. The four narrow spots condition was chosen over the three medium spots condition 27.9% of the time ($P = \frac{e^{-0.95}}{1+(e^{-0.95})}$ = 0.279). On the other hand, the uniform control condition was chosen over the four narrow spots condition 84.8% of the time ($P = \frac{e^{1.72}}{1+(e^{1.72})}$ = 0.848). No post-hoc pairwise comparisons were made at condition level.

Conditions of equal mean luminance	*Measure*	**Conditions of equal median luminance**
	+ 1	
	+ 0.9	
Uniform	+ 0.8	
	+ 0.7	
	+ 0.6	2M
	+ 0.5	
4W	+ 0.4	
2W 3W	+ 0.3	2W 3W
	+ 0.2	4W 2N 3N
	+ 0.1	
3M 2M 4M	0	3M
	- 0.1	
	- 0.2	
	- 0.3	4M
	- 0.4	4N
	- 0.5	
2N 3N	- 0.6	
	- 0.7	
	- 0.8	Uniform
	- 0.9	
4N	-1	
Conditions of equal mean luminance	*Measure*	**Conditions of equal median luminance**

Figure 16. Vertical ruler map for the brightness estimates on condition level, separated per set of conditions, with equal mean luminance on the left and equal median luminance on the right. A positive estimate (in logits) indicates a higher brightness and a negative estimate indicates a lower brightness compared to the reference.

3.3. Factor level

The brightness estimates, separated by the factors number of spots and beam width, can be found in Table 6. The corresponding vertical ruler map can be found in Figure 17, for set_{mean} and set_{median}, respectively. All items had very good fit statistics with MS-values ≤ 1.04.

Table 6.
Brightness estimates for the two sets of 10 conditions (in logits) on factor level, standard errors of the estimate (SE) and fit statistics IMS and OMS (Infit Mean-Square and Outfit Mean-Square) for the factors: side, number of spots and beam width, with equal mean luminance on the left and equal median luminance on the right.

		Equal Mean Luminance				Equal Median Luminance			
		Brightness	SE	IMS	OMS	Brightness	SE	IMS	OMS
Side	Left	0.00	0.05	1.00	1.00	0.00	0.05	1.00	1.00
	Right	0.00	0.05	1.00	1.00	-0.06	0.05	1.00	1.00
Number of spots	Two	0.05	0.07	1.00	1.00	0.18	0.06	1.00	1.00
	Three	0.00	0.07	1.00	1.01	0.00	0.06	1.01	1.01
	Four	-0.03	0.07	0.98	0.98	-0.33	0.06	0.99	0.99
Beam width	Narrow	-0.77	0.07	1.01	1.02	-0.14	0.06	1.00	1.00
	Medium	0.00	0.07	0.95	0.95	0.00	0.06	0.98	0.98
	Wide	0.38	0.07	1.03	1.04	0.13	0.06	1.02	1.02

In Figure 17, statistically significant differences from the post-hoc z-tests for pairwise comparisons between pairs are marked with * for p < .05, with ** for p <.01 and with *** for p <.001.

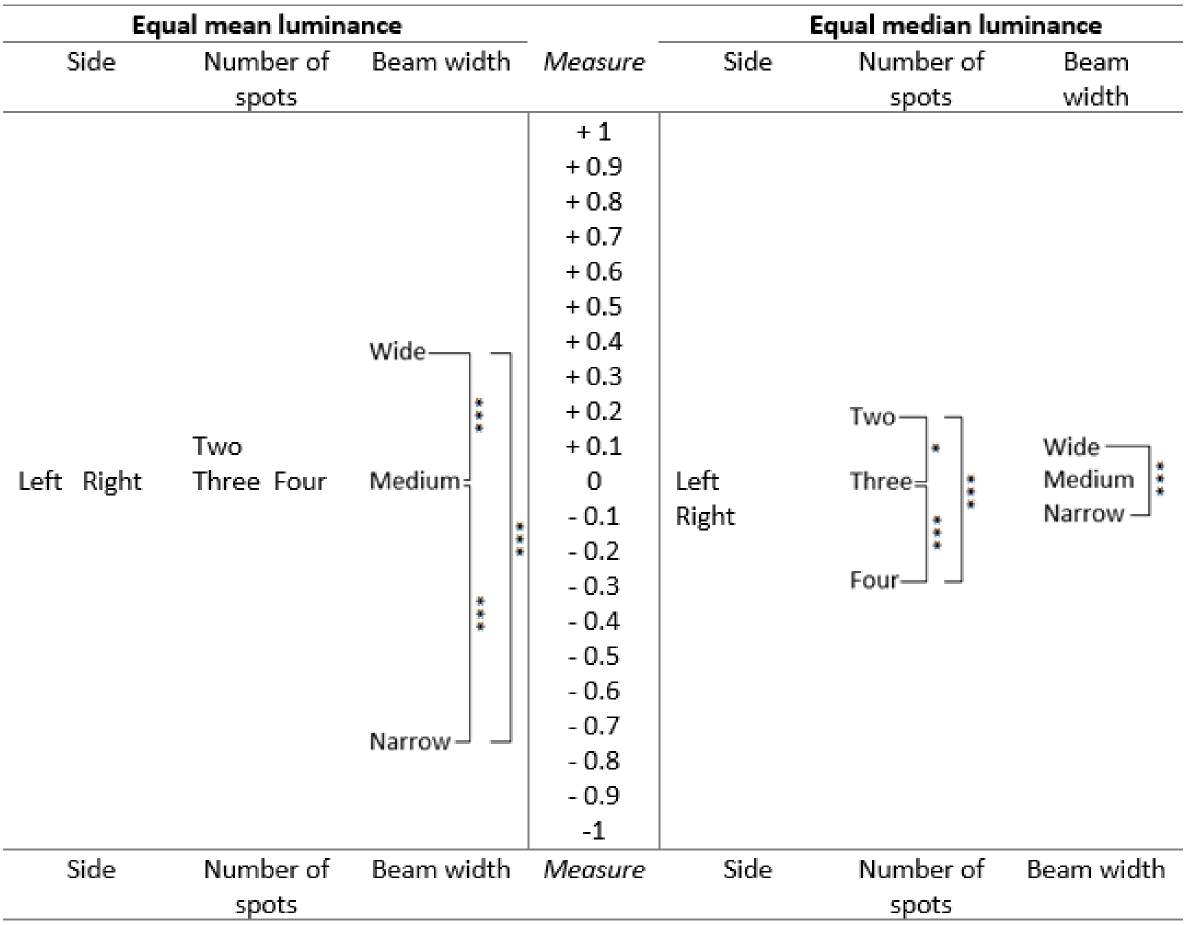


Figure 17. Vertical ruler map for the estimates on factor level, separated per set of conditions, with equal mean luminance on the left and equal median luminance on the right. The estimation of brightness is estimated by the factors: side, number of spots and beam width. A positive estimate (in logits) indicates a higher brightness, and a negative estimate indicates a lower brightness compared to the reference. Pairs with statistically significant differences are marked as follows:
* *p* < .05, ** *p* < .01, *** *p* < .001.

3.2.1. Side

To check for position bias, the side on which the condition was presented was included in the model. The left side was anchored to zero. For set_{mean}, the right side has an estimate of 0.00 logits (*SE* = 0.05), indicating that the right side was chosen over the left side 50% of the time. On the other hand, for set_{median}, the right side has an estimate of -0.06 logits (*SE* = 0.05), indicating that the right side was chosen over the left side 48.5% of the time. However, this effect was not found to be statistically significant with $p = .396$. Therefore, it can be concluded that the position on which the condition was presented had no influence on the probability of participants choosing the condition.

3.2.2. Number of spots

One of the factors that was manipulated was the number of spots used to create the lighting distribution. It was hypothesized that an increase in the number of spots would make the lighting distribution more uniform, and therefore brighter. The three spots were anchored to zero, as the reference condition contained three medium spots.

For set_{mean}, the estimate score for two spots was 0.05 logits (*SE* = 0.07), for three spots was 0.00 logits (*SE* = 0.07) and for four spots was -0.03 logits (*SE* = 0.07) with p-values $\geq .419$. Thus, the hypothesis that adding more spots will increase brightness is not supported for the set of conditions with equal mean luminance.

For set_{median}, the estimate score for two spots was 0.18 logits (*SE* = 0.06), for three spots was 0 logits (*SE* = 0.06) and for four spots was -0.33 logits (*SE* = 0.06) with $p = .034$ for two versus three, $p < .001$ for three versus four and $p < .001$ for two versus four. This indicates that participants chose two over three 54.5% of the time, three over four 58.2% of the time and two over four 62.5% of the time. Although the effects are relatively small, the hypothesis that uniform appears brighter is not supported based on these data. Conversely, these results suggest that non-uniform appeared brighter with equal median luminance.

3.2.3. Beam width

Another factor that was manipulated was the beam width of the spots used to create the lighting distribution. It was hypothesized that widening the beam width would increase the uniformity, and therefore the lighting distribution would appear brighter. The medium spots were anchored to zero, as the reference condition contained three medium spots.

For set_{mean}, the estimate score for narrow beams was -0.77 logits (*SE* = 0.07), for medium beams was 0 logits (*SE* = 0.07) and for wide beams was 0.38 logits (*SE* = 0.07). A pairwise z-test between the three conditions reveals that there was a statistically significant difference between all three combinations of beam widths, with p-values $< .001$. This indicates that participants chose medium over narrow 68.4% of the time, wide over medium 59.4% of the time and wide over narrow 76.0% of the time. Thus, the hypothesis that a wider beam width was perceived as brighter is supported by the data for the set of conditions with equal mean luminance.

For set_{median}, the estimate score for narrow beams was -0.14 logits (*SE* = 0.06), for medium beams was 0 logits (*SE* = 0.06) and for wide beams was 0.13 (*SE* = 0.06). A pairwise z-test between the three conditions reveals only partially statistically significant differences between the three combinations of the factor, with $p = .099$ for narrow versus medium, $p = .126$ for wide versus medium and $p = .001$ for narrow versus wide. This indicates that participants were not able to differentiate between narrow and medium, and between medium and wide, but chose wide over narrow 56.7% of the time. Although this effect is statistically significant, the effect is very small and therefore does not directly indicate that a wider beam width results in a higher perceived brightness. Thus, the hypothesis that a wider beam width results in a higher perceived brightness is only partially supported in the equal median luminance set of conditions.

3.4. Mean versus Median

To check whether the mean or median was a better predictor for brightness a Pearson correlation test was conducted with the brightness estimates. In set_{mean}, the brightness estimates were correlated with the corresponding medians. The results show a very strong correlation between the median and brightness with $r = .84$, and $p = .002$. This suggests that the median luminance was a good predictor of the brightness within set_{mean}.

Similarly, in set_{median}, the brightness estimates were correlated with the corresponding means. The results show a moderate correlation between the mean and brightness with $r = .58$, and $p = .078$. This correlation is not statistically significant, suggesting that the mean luminance was a predictor of brightness to a lesser extent.

An explorative correlation test was conducted to investigate the contribution of the maximum luminance to the brightness sensation. The minimum luminance was not used as a factor, as it did not deviate as much across conditions. A very strong negative correlation was found between the maximum luminance and the corresponding median in set_{mean}, with $r = -.97$, and $p < .001$. On the other hand, a very strong positive correlation was found between the maximum luminance and the corresponding mean in set_{median}, with $r = .98$, and $p < .001$. Suggesting that the maximum had a big influence on the median in set_{mean} and mean in $set_{median.}$ This indicates that the maximum can substitute the median in set_{mean} and the mean in set_{median} to predict brightness.

3.5. Uniformity

The overall data does not seem to provide strong evidence for an overall effect of uniformity on brightness. Within set_{mean}, brightness increases significantly with a wider beam width, which is considered to increase uniformity, but there does not seem to be an effect of the number of spots on brightness.

In contrast, within set_{median}, there seems to be an indication that decreasing the number of spots increases brightness, which is considered to decrease uniformity. On the other hand, a wider beam width appears to increase brightness, but although the effect of beam width is significant (only between the wide and narrow beams), the effect is very small. These results fail to accept the hypothesis that there is one overall effect of uniformity on brightness.

3.6. Participant Exclusion

However, as mentioned in section 3.1., there were three participants that exceeded the 1.5 fit statistics threshold in set_{mean} for both the analysis on condition level and factor level. When removing these three participants and repeating the analysis, the effect of number of spots does not change. However, the effect of beam width on brightness becomes more pronounced (see Table 7; Figure 18). The changes to the analysis outcomes are discussed in more detail below.

In set_{mean}, the new estimates for the narrow beams become 0.62 logits (SE = 0.08), for medium beams become 0 logits (SE = 0.07) and for wide beams become -1.00 logits (SE = 0.08), with p-values <0.001 from the pairwise z-tests. This indicates that participants chose medium over narrow 65.0% of the time, wide over medium 73.1% of the time and wide over narrow 83.5% of the time. The effect sizes have increased a little in the equal mean luminance set of conditions.

Moreover, in set_{median}, the new estimates for the narrow beams become 0.28 logits (SE = 0.07), for medium beams become 0 logits (SE = 0.07) and for wide beams become -0.28 logits (SE = 0.07). The pairwise z-test results show p = 0.005 for narrow versus medium, p = 0.005 for medium versus wide and p < 0.001 for narrow versus medium. This indicates that participants chose medium over narrow 57.0% of the time, wide over medium 57.0% of the time and wide over narrow 63.6% of the time. In other words, widening of the beam width results in a small increase in the perceived brightness.

These results indicate that there may have been participants in the sample population that had an opposite opinion on the effect of beam width on brightness.

Table 7.
Brightness estimates for the two sets of 10 conditions (in logits) on factor level after removing three participants that exceed the fit statistics threshold, standard errors of the estimate (SE) and fit statistics IMS and OMS (Infit Mean-Square and Outfit Mean-Square) for the factors: side, number of spots and beam width, with equal mean luminance on the left and equal median luminance on the right.

		Equal Mean Luminance				**Equal Median Luminance**			
		Brightness	SE	IMS	OMS	Brightness	SE	IMS	OMS
Side	Left	0.00	0.05	0.99	1.01	0.00	0.05	1.00	1.00
	Right	0.00	0.05	0.99	1.01	-0.07	0.05	1.00	1.00
Number of spots	Two	-0.05	0.07	1.00	1.02	0.20	0.07	1.00	1.00
	Three	0.00	0.07	1.00	1.03	0.00	0.07	1.02	1.02
	Four	-0.07	0.07	0.97	0.97	-0.33	0.07	0.98	0.98
Beam width	Narrow	-1.00	0.08	1.03	1.06	-0.28	0.07	1.01	1.01
	Medium	0.00	0.07	0.92	0.92	0.00	0.07	0.96	0.96
	Wide	0.62	0.08	1.05	1.08	0.28	0.07	1.02	1.02

In Figure 18, statistically significant differences from the post-hoc z-tests for pairwise comparisons between pairs are marked with * for p < .05, with ** for p <.01 and with *** for p <.001.

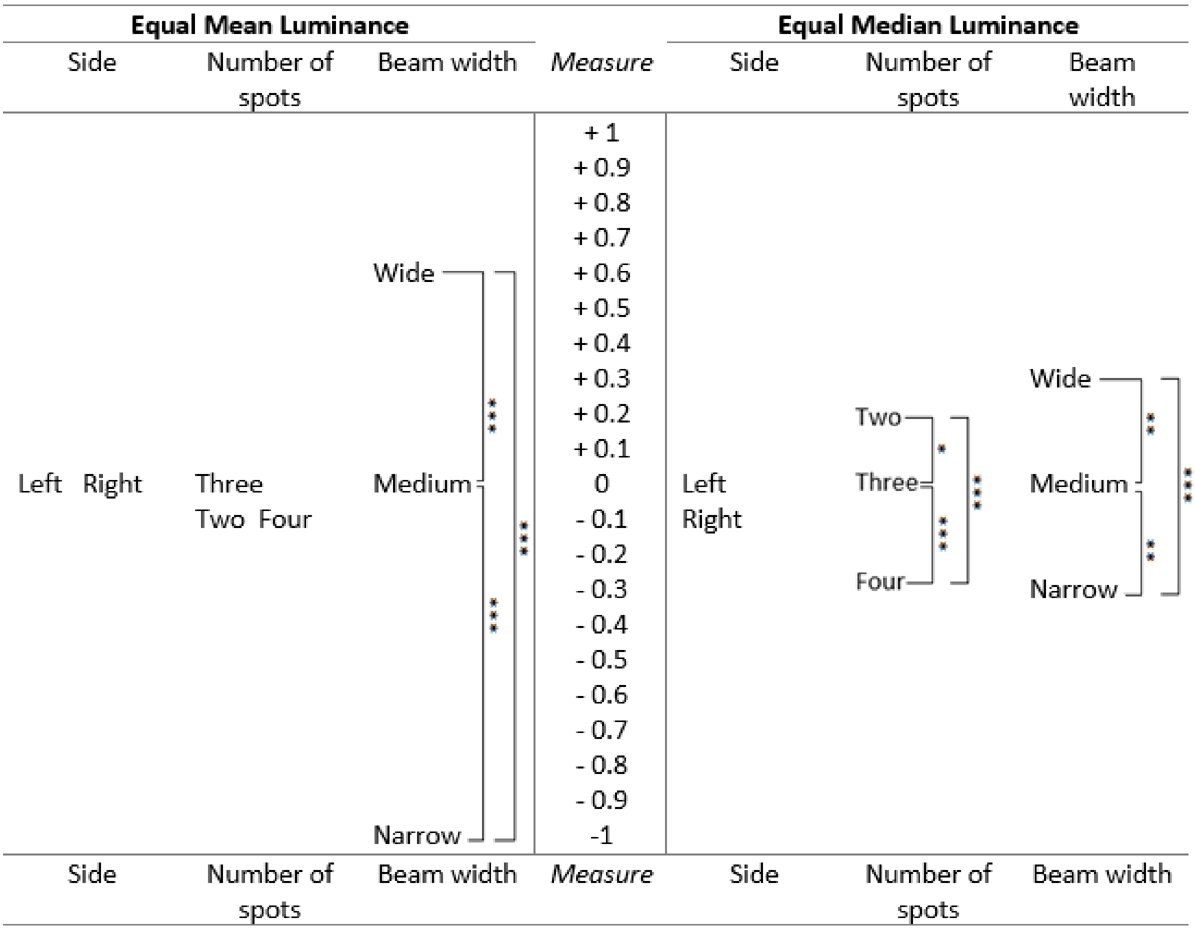


Figure 18. Vertical ruler map for the estimates on factor level after removing three participants that exceed the fit statistics threshold, separated per set of conditions, with equal mean luminance on the left and equal median luminance on the right. The estimation of brightness is estimated by the factors: side, number of spots and beam width. A positive measure (in logits) indicates a higher brightness, and a negative measure indicates a lower brightness compared to the reference. Pairs with statistically significant differences are marked as follows:
* p < .05, ** p < .01, *** p < .001.

4. Discussion

The aim of the current research was to investigate how uniformity relates to brightness and whether the mean or median luminance predicts brightness better to be used as basis for comparison between scenes. It was hypothesized that an increase in the number of spots and widening of the beam width would increase the uniformity of the lighting distribution and would therefore increase the perceived brightness. This hypothesis, however, cannot be firmly confirmed with the current results, as discussed in more detail later. Nevertheless, the results confirm that the way lighting is distributed within a space has an influence on the perceived brightness.

To allow easy comparison between the conditions, the two sets of conditions can be found in Figure 19. With set_{mean} on the left and set_{median} on the right. This is a repetition of Figures 9 and 10.

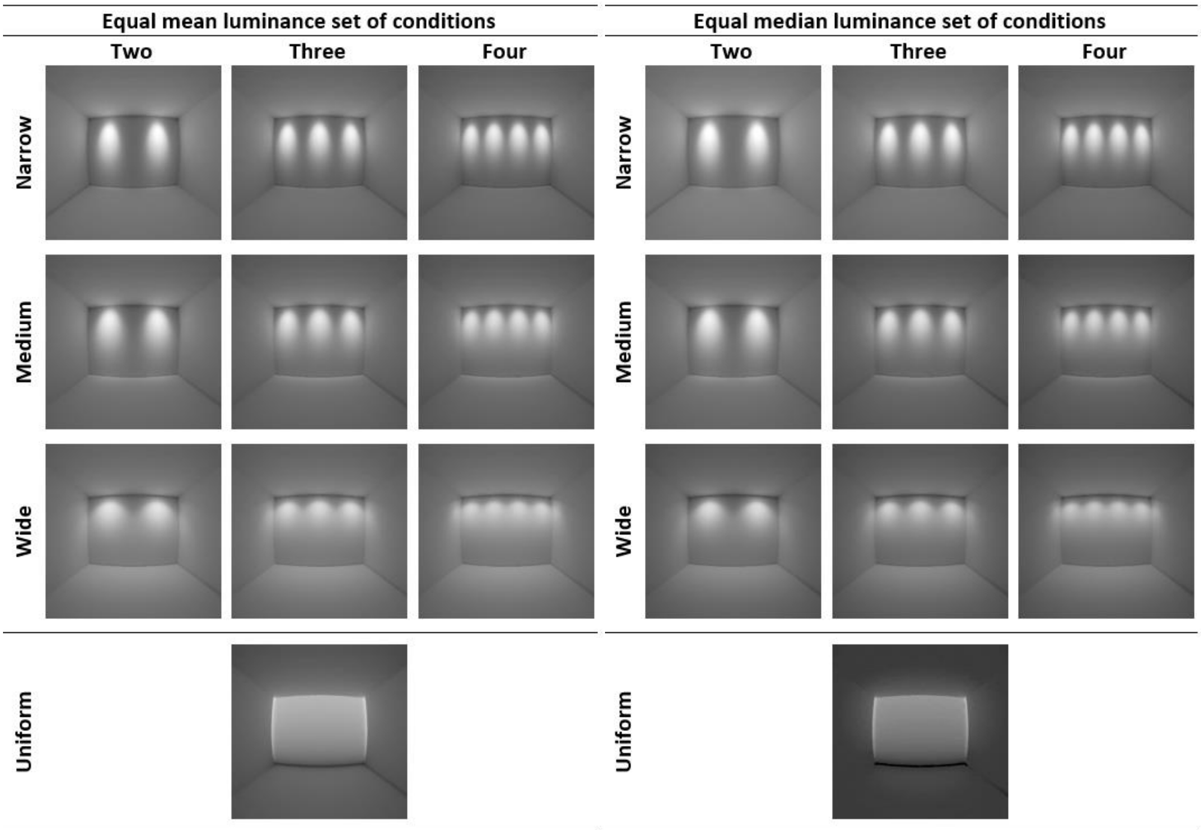


Figure 19. The two sets of conditions side by side, with set_{mean} on the left and set_{median} on the right. This is a repetition of Figures 9 and 10.

When visually observing and comparing the two sets of ten conditions, it can be seen that the conditions differ slightly in brightness between the equal mean and median luminance sets, but not considerably. In terms of luminance difference, the scenes of two narrow spots have a difference in mean luminance of 4 cd/m^2 and the two uniform control conditions differ with 5.5 cd/m^2. These differences seem minor, but when closely observing the two sets of scenes, some mild changes can be seen. Nevertheless, significant differences in brightness were found between the scenes. This indicates that this mild change is very important for brightness perception and that the human visual system is very sensitive to luminance differences. On the other hand, this difference might be exaggerated due to the use of virtual reality. The luminance range of a display is limited, therefore 5 cd/m^2 is 5% of the luminance range of the display. The power law of Stevens (1960) indicates that the human visual system is more sensitive to luminance differences at low luminance levels than at high luminance levels. The relatively low luminance level of the VR display in the current study, as opposed to a real scene, might have affected the outcome. Furthermore, as a result of this limited range, the visual adaptation of the eye may have been quicker than would have been in a real-world scene.

4.1. Condition versus Factor Level

The results of the equal mean luminance set of conditions (set_{mean}) on condition level, with varying median luminance, show that the scene with the uniform control condition is perceived as the brightest and chosen over the dimmest condition (four narrow spots) 84.8% of the time. The uniform control condition was perceived as the brightest, followed by the three conditions with wide beams, the three conditions with medium beams, two of the conditions with narrow beams and lastly the condition with four narrow beams. Separated on factor level, the results of set_{mean} show that a wider beam width was chosen over a narrower beam 59.4% to 76.0% of the time, indicating that more uniform beam widths are perceived as brighter. Unlike beam width, changing the number of spots had no influence on the perceived brightness. These results suggest that uniform (particularly related to wider beams) lighting distributions appear brighter than non-uniform lighting distributions with equal mean luminance.

These results are in line with results from Kirsch (2015), Kobayashi et al. (1998), Kato & Sekiguchi (2005) and Hsieh (2012), who found that with equal mean luminance, uniform scenes appeared brighter. On the contrary, the results differ from the results from Tiller and Veitch (1995), Newsham et al. (2004), Chraibi et al. (2013), and Sullivan and Donn (2016, 2018) who found that non-uniform lighting distributions were perceived as brighter than uniform lighting distributions with equal mean luminance. A possibility for the disagreement of the current results with the results of Tiller and Veitch is the difference in manipulation of the stimuli. In the study of Tiller and Veitch, the lighting distribution differed in beam width, contrast and location on the wall. The location of the lighting on the wall may have been an important factor in the assessment of brightness. This was found by Kirsch (2015), showing that the location of the lighting relative to the observer was an important factor for the assessment of visual lightness. On the other hand, the method of brightness matching in real spaces was used by Tiller and Veitch and by Sullivan and Donn, while the current study used a 2AFC paired comparisons method in VR. Participants in the study of Tiller and Veitch reported that the brightness assessment was based on the contrast within the room. Possibly, their participants used anchoring theory and used the lightest area of the lighting distribution for their assessment. Non-uniform lighting distributions may appear brighter when directly adjusted due to anchoring but may still appear relatively dim due to the large presence of shadows when compared as a whole. Furthermore, brightness perception in real scenes may differ from VR scenes.

In contrast, the results of the equal median luminance set of conditions (set_{median}), with varying mean luminance, show that the two medium spots condition was perceived as the brightest and chosen over the dimmest condition (the uniform control condition) 80.5% of the time. Two spots are relatively high on the ruler, while four spots are relatively low on the ruler (Figure 16), suggesting that in general less uniform lighting distributions appear brighter than more uniform lighting distributions with equal median luminance. However, the order of the conditions is less structured in terms of uniformity than in the set_{mean}. Unlike set_{mean}, the effect of beam width is not found to be prominent in set_{median}. A statistically significant effect of widening the beam width from narrow to wide was found, where a wide beam width was chosen over a narrow beam 56.7% of the time. On the other hand, decreasing the number of spots resulted in a small increase in brightness, with less spots being chosen over more spots 54.4% to 62.5% of the time. This is surprising, because decreasing the number of spots was assumed to decrease the uniformity of the lighting distribution. Thus, in one aspect – beam width – an increased uniformity increases the perceived brightness, whereas in the other aspect – number of spots – a decreased uniformity increases the perceived brightness. Therefore, these results seem quite arbitrary and cannot directly answer the main research question.

When looking at the results on condition level (Figure 16), the effect sizes in set_{mean} and set_{median} are about equal in magnitude, while the effects are less pronounced on factor level (Figure 17). The reason for this is that on factor level, the effects of the separate walls are stacked, especially in set_{median}. The condition with two medium spots is perceived as the brightest wall, but there is no indication that two spots or medium beams are the determinative factor in this assessment as the other walls with two spots or medium beams are not structurally distributed over the ruler. Furthermore, the four wide spots are perceived as significantly brighter than four narrow and four medium spots. Therefore, the effect of four spots is not clearly present on factor level. On the other hand, the conditions (apart from the uniform control condition) assumed to have the lowest and highest uniformity, respectively the two narrow spots and four wide spots, have the same brightness estimates. This indicates that these two conditions are perceived as equally bright. This suggests that the number of spots and beam width are not accurate predictors for assessing brightness when the median luminance is kept constant.

4.2. Mean versus Median

A possible reason for an effect of beam width, and not of number of spots, to be prevalent in set_{mean} is the correlation of the brightness estimates with the corresponding median luminance. When the mean luminance is kept constant, the median has a strong correlation with brightness. Hence, the median luminance is a good predictor for the perceived brightness. Increasing the number of spots while maintaining the beam width barely increased the median luminance (with 0.2 cd/m^2 to 2 cd/m^2), whereas widening the beam width while maintaining the number of spots increased the median luminance with 2.2 cd/m^2 to 4 cd/m^2 (see Table 4). Likewise, as can be seen in Figures 11 and 12, the peak luminance decreases substantially more between the factor beam width than between the factor number of spots. It is not surprising that the effect of beam width is more pronounced than the effect of number of spots, because beam width seems to influence uniformity more than the number of spots in a lighting distribution.

On the other hand, if the median luminance is kept constant, the mean has only a moderate correlation with brightness, suggesting the mean luminance is a predictor for brightness to a lesser degree. This is in line with results by De Vries et al. (2022), who found that the median luminance was a better predictor for brightness than the mean luminance.

A possible explanation for this finding is the logarithmic nature of our perception. As discussed in section 1.1.2, a doubling of luminance is not equal to a doubling in brightness (Stevens, 1960), but results in less than a doubling in brightness. Therefore, the perceived peak brightness in non-uniform scenes is proportionally lower than the measured peak luminance. An increase in the peak luminance has a significant effect on the mean, but only a slight effect on the median, making the median more robust to different lighting patterns. In order to distinguish between a uniform and non-uniform scene, participants likely tried to average the two scenes. Due to lateral inhibition or anchoring theory and possibly the highest-luminance-as-white rule (Gilchrist et al., 1999), participants may have perceived the peak luminance as white and scaled the rest of the room accordingly. Then, as the peak brightness is perceived to be proportionally lower than the actual peak luminance, the rest of the scene would also be scaled down, resulting in the perceived average brightness to be lower than the mean luminance and might therefore have been closer to the median luminance. This may have occurred since the median luminance was lower than the mean luminance in all conditions, except the uniform control condition. This was not designed intentionally but may have been an artifact of the study. On the other hand, if this was an artifact, this finding should not have occurred in the uniform control condition, as here the median was slightly higher than the mean. Excluding this condition from the correlation should make the correlation stronger, however, the correlation becomes weaker (from $r = .84$ to $r = .78$) when this condition is excluded. Based on these results it is difficult to conclude whether this was an artifact and future research is required.

4.3. Uniformity

As discussed above, the median luminance was found to be a better predictor for brightness than the mean luminance. With this in mind, it is reasonable to expect that effects independent from the median emerge when the median luminance remains constant. For instance, it was expected that uniformity would be isolated when the median was kept constant. However, this does not directly appear to be the case. When looking at the results from set_{median} on factor level, it can be seen that an increase in the number of spots results in a lower perceived brightness, while increasing the beam width slightly increased the perceived brightness. However, when excluding participants that exceed the fit statistics threshold, the effect of beam width becomes more prominent, indicating there might be large inter-personal differences. These differences may be due to a different assessment method. Some participants indicated that they focused on the wall as a whole while others mentioned that they decided based on the contrast in the lighting distribution. For instance, when only focusing on the contrast within the room, participants may have used an anchor to base their judgement on, such as the brightest-as-white rule (Gilchrist et al., 1999). On the other hand, when looking at the lighting within the room as a whole, the decision may have been based on the brightness of the room, rather than a specific area in the lighting distribution. Alternatively, the difference between participants may have been based on a difference in sensation. De Vries et al. (2022) found two distinct groups of people, whereas one group responded evenly to brightness differences, while the other group exhibited great sensitivity to brightness differences. This may have occurred in the current study as well. Participants less sensitive to brightness differences may then have decided based on other factors. For example, it could be possible that the judgments in set_{median} were mainly based on aesthetic preferences when there was no clear factor for brightness. Either way, these are only speculations and require further investigation.

Nevertheless, from these results, it is not apparent that there is one straightforward answer for the effect of uniformity on brightness. Instead, it is possibly an interplay of multiple factors. For instance, it is possible that the results of previous studies (Kobayashi et al., 1998; Kato &

Sekiguchi, 2005; Hsieh, 2012; Kirsch, 2015; Tiller & Veitch, 1995; Newsham et al., 2004; Chraibi et al., 2013, Sullivan & Donn, 2016; Sullivan & Donn, 2018) were mediated by the median luminance. All of these studies used a constant mean luminance in their setups and did not consider the median luminance. These studies reported an effect of uniformity, while it may have been a mechanism of the median luminance. For example, in results from Tiller and Veitch (1995), the median in the non-uniform lighting distribution may have been higher than the median in the uniform lighting distribution. However, the luminance values in their report were not reported. A possible explanation for the contrasting results between studies is the use of relatively uniform lighting distributions. In such situations, the median is relatively close to the mean luminance.

On the other hand, for example, the results from research by Sullivan and Donn (2016; 2018) found that non-uniform lighting distributions appeared brighter than uniform lighting distributions. In their study, they manipulated the reflectivity of the walls, ceiling and floor in the room in which the luminance of a surface is uniformly increased or decreased. This has a large influence on the median luminance, which could have influenced their results. However, the median luminance was not reported in their report. Alternatively, as Sullivan and Donn manipulated the lighting of the walls, ceiling and floor, it is possible that their results show the relative importance of the different surfaces. Given that the effect of glare varies depending on the location of the light source in the field of view (Iwata & Tokura, 1997; Kim & Kim, 2011), the possibility exists that this also applies for brightness perception. Thus, uniformity might be influenced by more parameters that have not yet been investigated in detail, such as the location of the lighting on the wall and in the room, contrast within the space or in more detail the gradient of the beam.

4.5. Limitations

Throughout this study, there were multiple limitations. The equal mean and median luminance scenes were presented randomly mixed. Participants were unaware that they were presented with two different sets of conditions, where some thought there were multiple repetitions. It could have occurred that participants' judgments for the equal median luminance scenes were influenced by the equal mean luminance scenes and vice versa. If the wide beam was perceived as significantly brighter than the other beam widths in set_{mean}, participants could have used this as a bias in set_{median} as well. Then, if the sets were not mixed and participants would have been unable to use this bias in set_{median}, it is possible that participants would choose narrower beam widths more frequently and a negative relation of uniformity on brightness could emerge more prominently. If the conditions were presented in two sets separately, the effect of uniformity could have been investigated in isolation.

The spatial distribution of the lighting in the room was manipulated by changing the number of spots used in the lighting distribution. However, as suggested by previous research, the location of the lighting on the wall may have been a more important factor in perceived uniformity than the number of spots. The location of the lighting has more influence on the subjective than the objective assessment of uniformity as it does not directly influence the luminance values. Perhaps, this is an essential factor in the assessment of brightness as well.

Furthermore, some participants asked whether the color or paint of the walls, referring to the lightness, was constant across the conditions. This suggests that the lightness of the walls was perceived differently across the conditions. This conception could be caused by the masking image that was placed on top of the surrounding environment to cover reflections, possibly making the room appear non-realistic and potentially influencing the perceived brightness of the walls. For instance, less of the wall is covered by light in non-uniform lighting distributions than in uniform lighting distributions and could therefore be perceived as darker. This could also have occurred due to the use of VR instead of a real room, where one would be able to see that the lightness of the wall does not change.

The median luminance was in all, except one condition, lower than the mean luminance. This could be an error in the study design and follow-up studies should include conditions where the median luminance is higher than the mean luminance in order to conclude whether the median luminance is still a good predictor in those scenes.

Next, the distance the participant is positioned from the wall is quite small. It is possible to see both walls in one field of view, but the walls are relatively close to the observer. If the distance had been a little further, there would have been a better overview of the room. Due to the close proximity to the wall, participants might have focused on one specific area of the wall instead of the wall as being part of a room. When a wall is lit by narrow spots, the room itself might appear relatively dark, while a wall lit by wide beams spreads the lighting more diffuse throughout the room and may not give this impression. It is important to have a good overview of the rooms in order to make a distinction between the brightness of the rooms and not necessarily of one specific area of the lighting.

The measurements of the luminance distribution directly in the VR headset with the luminance camera showed that the median luminance in the equal median luminance set of conditions is not completely constant. However, it is possible this difference is not visible to the human eye, but for future research it would be important to keep the median luminance perfectly equal.

Last, as the study was in VR and not in a real-world scenario, a replication study in the real world would be necessary verify the validation of the method.

5. Conclusion and Recommendations

Previous research has shown that the lighting distribution of a room is essential for the way the room is perceived. Brightness has been found to be an important factor in this. However, it was not clear how brightness is affected by the lighting distribution. The majority of studies reported that uniform lighting distributions appear brighter, while others suggested the opposite.

This study aimed to answer the main research question: "*How does the uniformity of the lighting distribution on a wall, created by electrical light sources, affect the brightness of a room?*" and therefore investigated the effect of uniformity on the brightness of scenes with varying lighting distributions. Uniformity was manipulated through the factors number of spots and the beam width of these spots. Furthermore, the study stimuli consisted of two sets of conditions where in one the mean luminance and in the other the median luminance was kept constant, to investigate the role of both. A total of 28 participants participated in a two-alternative forced choice paired comparisons VR study. The main hypothesis of a positive relation between uniformity and brightness cannot directly be confirmed based on the results of the current study. The results show that when the mean luminance is constant, uniform scenes are perceived as brighter than non-uniform scenes, however, this effect was possibly mediated by the median as there was a high correlation between brightness and the median in this set.

On the other hand, when the median luminance is kept constant, the results show a significant increase in brightness when decreasing the number of spots, but a small significant increase in brightness with widening the beams of lighting used in the lighting distribution. From a numeric perspective, beam width seems to have a bigger influence on the uniformity than the number of spots. From these results, it seems that either uniformity is not an important factor for brightness, or that there is no straightforward answer of the influence of uniformity on brightness, instead, uniformity cannot be described in one term and is most likely an interplay of multiple factors.

This work highlights the importance of considering different parameters in the lighting distribution to gain an understanding of brightness in order to guide lighting designers to achieve effective and efficient lighting design. For future research it would be interesting to add more different factors to the study. One interesting factor would be the location of the spots on the wall or within the space, since these factors are possibly important for brightness assessment as suggested by previous studies. On the other hand, more elaborate manipulations of beam width and shape can be interesting to investigate to allow for a more extensive conclusion on the effect of the beam shape and gradient of a light source on brightness. Furthermore, the reflectance of the walls could be manipulated or contrast in relation to brightness could be used as a factor to investigate the effect of uniformity in more detail.

In the current study, both the mean and median luminance were kept equal in separate sets, but due to the correlation test this was not necessarily a requirement to investigate the role of both. This allows future research to include more factors in the model.

Finally, it is important to repeat the procedure in the real-world to check the validity of the results from the VR study.

Acknowledgements

I would like to express my gratitude to my supervisors Yvonne de Kort and Kynthia Chamilothori from the HTI department at the TU/e for their advice, support, and weekly meetings throughout this project. Secondly, I would like to thank Signify and in particular Adrie de Vries for his professional advice, guidance and cooperation, and for giving me the opportunity to do this project within Signify. It was a very interesting project in which I learned a whole lot!

I would like to express my special thanks to Sander Meessen for his incredible help on creating the VR application for the experiment. His help played a crucial role in the project.

Finally, I would like to thank all the volunteers from Signify for participating in my study and everyone else who helped and supported me throughout this project.

References

Abboushi, B., Elzeyadi, I., Wymelenberg, K. Van Den, Sereno, M., & Jacobsen, G. (2020). Assessing the Visual Comfort, Visual Interest of Sunlight Patterns , and View Quality under Different Window Conditions in an Open-Plan Office. *The Journal of the Illuminating Engineering Society*, *00*(00), 1–17. https://doi.org/10.1080/15502724.2020.1785309

Abboushi, B., Irvin, L., Rodriguez-Feo Bermudez, E., & Royer, M. (2022). Evaluating Luminance Uniformity Metrics Using Online Experiments. *LEUKOS - Journal of Illuminating Engineering Society of North America*, *00*(00), 1–16. https://doi.org/10.1080/15502724.2022.2133964

Abd-Alhamid, F., Kent, M., Bennett, C., Calautit, J., & Wu, Y. (2019). Developing an Innovative Method for Visual Perception Evaluation in a Physical-Based Virtual Environment. *Building and Environment*, *162*(July), 106278. https://doi.org/10.1016/j.buildenv.2019.106278

Adelson, E. H. (1993). *Perceptual Organization and the Judgment of Brightness*. *262*(December), 2042–2044.

Altman, D. G., & Bland, J. M. (2003). Interaction revisited: the difference between two estimates. *BMJ*, *326*(7382), 219. doi:10.1136/bmj.326.7382.219

Armstrong, J. D. (1990). A New Measure of Uniformity for Lighting Installations. *Journal of the Illuminating Engineering Society*, *19*(2), 84–89. doi:10.1080/00994480.1990.10747967

Barraza, J. F., & Martín, A. (2020). The Effect of Texture on Brightness Perception in Simulated Scenes. *LEUKOS - Journal of Illuminating Engineering Society of North America*, *16*(4), 279–287. https://doi.org/10.1080/15502724.2019.1674661

Baxant, P., Ing, A. P., & Ph, D. (2016). *Contrast Analysis in Lighting Technology*.

Bedrosian, T. A., & Nelson, R. J. (2013). Influence of the modern light environment on mood. *Molecular Psychiatry*, *April*, 751–757. https://doi.org/10.1038/mp.2013.70

Bellazzi, A., Bellia, L., Chinazzo, G., Corbisiero, F., D'Agostino, P., Devitofrancesco, A., Fragliasso, F., Ghellere, M., Megale, V., & Salamone, F. (2022). Virtual reality for assessing visual quality and lighting perception: A systematic review. *Building and Environment*, *209*(December 2021), 108674. https://doi.org/10.1016/j.buildenv.2021.108674

Berman, S. M., Jewett, D. L., Fein, G., Saika, G., & Ashford, F. (1990). Photopic luminance does not always predict perceived room brightness. *Lighting Research and Technology*, *22*(1), 37–41.

Bodmann, H. W., & La Toison, M. (1994). Predicted brightness-luminance phenomena. *October*, *26*(5), 135–143.

Boyce, P. R. (1977). Investigations of the subjective balance between illuminance and lamp colour properties. *Lighting Research & Technology*, *9*(1), 11–24. https://doi.org/10.1177/096032717700900102

Boyce, P. R. (2014). *Human factors in lighting* (3rd ed.). CRC Press. https://doi.org/10.1201/b16707

Brown, T. M., Tsujimura, S. I., Allen, A. E., Wynne, J., Bedford, R., Vickery, G., Vugler, A., & Lucas, R. J. (2012). Melanopsin-based brightness discrimination in mice and humans. *Current Biology*, *22*(12), 1134–1141. https://doi.org/10.1016/j.cub.2012.04.039

Cabanier, R (2023) WebXR Layers API Level 1. *W3C Working Draft*. URL https://www.w3.org/TR/webxrlayers-1/

Chamilothori, K. (2019). Perceptual effects of daylight patterns in architecture. *Lausanne: Ecole Polytechnique Fédérale de Lausanne.* https://doi.org/10.5075/epfl-thesis-9553 [PhD Dissertation]

Chamilothori, K., Wienold, J., & Andersen, M. (2019). Adequacy of Immersive Virtual Reality for the Perception of Daylit Spaces: Comparison of Real and Virtual Environments. LEUKOS - *Journal of Illuminating Engineering Society of North America*, 15(2–3), 203–226. https://doi.org/10.1080/15502724.2017.1404918

Chamilothori, K., Wienold, J., Moscoso, C., Matusiak, B., & Andersen, M. (2022). Subjective and physiological responses towards daylit spaces with contemporary façade patterns in virtual reality: Influence of sky type, space function, and latitude. *Journal of Environmental Psychology*, 82(June), 101839. https://doi.org/10.1016/j.jenvp.2022.101839

Chen, Y., Cui, Z., & Hao, L. (2019). Virtual reality in lighting research: Comparing physical and virtual lighting environments. *Lighting Research and Technology*, *51*(6), 820–837. https://doi.org/10.1177/1477153518825387

Chraibi, S., Crommentuijn, L., Loenen, E. van, & Rosemann, A. (2017). Influence of wall luminance and uniformity on preferred task illuminance. *Building and Environment*, *117*, 24–35. https://doi.org/10.1016/j.buildenv.2017.02.026

CIE, Commision Internationale de l'Éclairage. (1926). *Commission Internationale de l'Éclairage Proceedings*, 1924. Cambridge: Cambridge University Press.

CIE, Commission Internationale de l'Éclairage. (1951). *Proceedings* Vol. 1, Sec 4; Vol 3, p. 37. Paris: Bureau Central de la CIE

CIE, Commission Internationale de l'Éclairage. (2011). 017/E: 2011 2011 ILV: International Lighting Vocabulary.

Colombo, E., Barraza, J., & Issolio, L. (2000). Effect of brief exposure to glare on brightness perception in the scotopic-mesopic range. *Lighting Research & Technology*, *32*(2), 65–69.

Cornelissen, F. W., Wade, A. R., Vladusich, T., Dougherty, R. F., & Wandell, B. A. (2006). No functional magnetic resonance imaging evidence for brightness and color filling-in in early human visual cortex. *Journal of Neuroscience*, *26*(14), 3634–3641. https://doi.org/10.1523/JNEUROSCI.4382-05.2006

Cuttle, C. (2009). Towards the third stage of the lighting profession. *Lighting Research and Technology*, *42*(1), 73–93. https://doi.org/10.1177/1477153509104013

Davey, M. P., Maddess, T., & Srinivasan, M. V. (1998). The spatiotemporal properties of the Craik-O'Brien-Cornsweet effect are consistent with "filling-in." *Vision Research*, *38*(13), 2037–2046. https://doi.org/10.1016/S0042-6989(97)00329-5

De Vries, A., Heynderickx, I., & de Kort, Y. A. W. (2022). From luminance to brightness: A data-driven approach to support brightness assessments in open plan offices. *Lighting Research and Technology*, 798–818. https://doi.org/10.1177/147715352211117365

De Vries, A., Heynderickx, I., Souman, J., & de Kort, Y. (2021). Putting the ceiling center stage – The impact of direct/indirect lighting on room appraisal. *Building and Environment*, *201*, 107989. https://doi.org/10.1016/j.buildenv.2021.107989

De Vries, A., Souman, J. L., de Ruyter, B., Heynderickx, I., & de Kort, Y. A. W. (2018). Lighting up the office: The effect of wall luminance on room appraisal, office workers' performance, and subjective alertness. *Building and Environment*, *142*, 534–543. https://doi.org/10.1016/j.buildenv.2018.06.046

Economou, E., Zdravkovic, S., & Gilchrist, A. (2007). Anchoring versus spatial filtering accounts of simultaneous lightness contrast. *Journal of vision*, 7(12), 2-2. https://doi.org/10.1167/7.12.2

Fechner, G. T. (1860). *Elemente der Psychophysik*. Leipzig: Breitkopf und Härtel.

Ferwerda, J. A., Pattanaik, S. N., Shirley, P., & Greenberg, D. P. (1996). A model of visual adaptation for realistic image synthesis. *Proceedings of the 23rd Annual Conference on Computer Graphics and Interactive Techniques*, SIGGRAPH 1996, 249–258. https://doi.org/10.1145/237170.237262

Flynn, J. E., Spencer, T. J., Martyniuk, O., & Hendrick, C. (1973). Interim study of procedures for investigating the effect of light on impression and behavior. *Journal of the Illuminating Engineering Society*, *3*(1), 87–94. https://doi.org/10.1080/00994480.1973.10732231

Fotios, S. A., Atli, D., Cheal, C., & Hara, N. (2015). Lamp spectrum and spatial brightness at photopic levels: Investigating prediction using S/P ratio and gamut area. *Lighting Research and Technology*, *47*(5), 595–612. https://doi.org/10.1177/1477153514542295

Fotios, S. A., & Cheal, C. (2011). Brightness matching with visual fields of different types. *Lighting Research and Technology*, *43*(1), 73–85. https://doi.org/10.1177/1477153510369478

Fotios, S. A., & Houser, K. W. (2013). Using Forced Choice Discrimination to Measure the Perceptual Response to Light of Different Characteristics. *LEUKOS*. 245–259. https://doi.org/10.1582/LEUKOS.2013.09.04.002

Frisby, J. P., & Clatworthy, J. L. (1975). Illusory contours: curious cases of simultaneous brightness contrast? *Perception*, *4*(3), 349–357. https://doi.org/10.1068/p040349

Geier, J., & Hudák, M. (2011). Changing the chevreul illusion by a background luminance ramp: Lateral inhibition fails at its traditional stronghold - s psychophysical refutation. *PLoS ONE*, *6*(10). https://doi.org/10.1371/journal.pone.0026062

Gerrits, H. J. M., & Vendrik, A. J. H. (1970). Simultaneous contrast, filling-in process and information processing in man's visual system. *Experimental Brain Research*, *11*(4), 411–430. https://doi.org/10.1007/BF00237914

Google Inc. (n.d.). *Rendering Omni-directional Stereo Content*. Derived from https://developers.google.com/vr/jump/rendering-ods-content.pdf

Hawkes, R. J., Loe, D. L., & Rowlands, E. (1979). A Note Towards the Understanding of Lighting Quality. *Journal of the Illuminating Engineering Society*, *8*(2), 111–120. https://doi.org/10.1080/00994480.1979.10748578

Hegazy, M., Ichiriyama, K., Yasufuku, K., & Abe, H. (2021). Comparing daylight brightness perception in real and immersive virtual environments using perceptual light maps. *Automation in Construction*, *131*(August 2020), 103898. https://doi.org/10.1016/j.autcon.2021.103898

Heydarian, A., & Becerik-Gerber, B. (2017). Use of immersive virtual environments for occupant behaviour monitoring and data collection. *Journal of Building Performance Simulation*, *10*(5–6), 484–498. https://doi.org/10.1080/19401493.2016.1267801

Houser, K. W., Boyce, P. R., Zeitzer, J. M., & Herf, M. (2020). Human-centric lighting: Myth, magic or metaphor? *Lighting Research and Technology*, *53*(2), 97–118. https://doi.org/10.1177/1477153520958448

Houser, K. W., Tiller, D. K., Bernecker, C. A., & Mistrick, R. G. (2002). The subjective response to linear fluorescent direct/indirect lighting systems. *Lighting Research & Technology*, *34*(3), 243–260. https://doi.org/10.1191/1365782802li039oa

Hsieh, M. (2012). The energy-saving effect and prediction method under various illuminance distribution types. *Building and Environment*, *58*, 145–151. https://doi.org/10.1016/j.buildenv.2012.07.001

Hsieh, C., & Li, Y. (2013). Optical Microstructure Design Optimization for Display Backlighting. 2013(November), 202–207. http://dx.doi.org/10.4236/mme.2013.34027

Hu, Z., Zhang, P., Wei, B., Ding, W., & Dai, Q. (2023). *Assessment of spatial brightness for a visual field in interior spaces based on indirect corneal illuminance*. *31*(2), 997–1013. https://doi.org/10.1364/OE.477637

Ishida, T., & Ogiuchi, Y. (2002). Psychological determinants of brightness of a space perceived strength of light source and amount of light in the space-. *Journal of Light and Visual Environment* (Vol. 26, Issue 2, pp. 29–35). https://doi.org/10.2150/jlve.26.2_29

ISO/CIE TR 21783:2022 | ISO/CIE TR 21783 (2022). *Light and lighting - Integrative lighting - Non-visual effects*.

Iwata, T., & Tokura, M. (1997). Position Index for a glare source located below the line of vision. *Lighting Research and Technology*, *29*(3), 172–178. doi:10.1177/147715359702900308 01

Jacobsen, A., & Gilchrist, A. (1988). *The ratio principle holds over a million-to-one range of illumination*.

Jay, P. A. (1971). Lighting and visual perception. *Lighting Research & Technology*, *3*(2), 133–146. https://doi.org/10.1177/096032717100300207

Jin, X., Meneely, J., & Park, N. K. (2022). Virtual Reality Versus Real-World Space: Comparing Perceptions of Brightness, Glare, Spaciousness, and Visual Acuity. *Journal of Interior Design*, *47*(2), 31–50. https://doi.org/10.1111/joid.12209

KATO, M., & SEKIGUCHI, K. (2005). "Impression of Brightness of a Space" Judged by Information from the Entire Space. In *Journal of Light & Visual Environment*, *29*(3), 123–134). https://doi.org/10.2150/jlve.29.123

Kim, W., & Kim, J. T. (2011). A position index formula for evaluation of glare source in the visual field. *Indoor and built environment*, *20*(1), 47-53.

Kingdom, F. A. A. (2011). Lightness , brightness and transparency : A quarter century of new ideas , captivating demonstrations and unrelenting controversy. *Vision Research*, *51*(7), 652–673. https://doi.org/10.1016/j.visres.2010.09.012

Kinzuka, Y., Sato, F., Minami, T., & Nakauchi, S. (2021). Effect of glare illusion-induced perceptual brightness on temporal perception. *Psychophysiology*, *58*(9), 1–14. https://doi.org/10.1111/psyp.13851

Kirsch, R. (2015). *Lighting Quality and Energy Efficiency in Office Spaces*. urn:nbn:de:kobv:83-opus4-65760

Kobayashi, S., Nakamura, Y., & Inui, M. (1998). Impressions of Overall Brightness in a Non-Uniformiy Illuminated Space. In *Journal of Light and Visual Environment* (Vol. 22, Issue 1, pp. 34–41). https://doi.org/10.2150/jlve.22.1_34

Kozusznik, M. W., Maricutoiu, L. P., Peiró, J. M., Vîrgă, D. M., Soriano, A., & Mateo-Cecilia, C. (2019). Decoupling office energy efficiency from employees' well-being and performance: A systematic review. *Frontiers in Psychology*, *10*(FEB). https://doi.org/10.3389/fpsyg.2019.00293

Land, E. H., & McCann, J. J. (1971). Lightness and Retinex Theory. *Journal of the Optical Society of America*, 61(1), 1–11.

Legates, T. A., Fernandez, D. C., & Hattar, S. (2014). Light as a central modulator of circadian rhythms , sleep and affect. *Nature Publishing Group, June*. https://doi.org/10.1038/nrn3743

Linacre, J.M. (n.d.). Paired comparisons of objects. Winsteps. https://www.winsteps.com/facetman/pairedcomparisons.htm

Linacre J.M. (1997) Paired comparisons with standard Rasch software. *RMT*, 11:3, 584-5. www.rasch.org/rmt/rmt113o.htm

Linacre, J. M. (2002). What do infit and outfit, mean-square and standardized mean. Rasch measurement transactions, 16(2), 878.

Linacre, J. M. (2023). Facets computer program for many-facet Rasch measurement, version 3.85.1. Beaverton, Oregon: Winsteps.com

Lindh, U. W., & Billger, M. (2021). Light distribution and perceived spaciousness: Light patterns in scale models. *Sustainability (Switzerland)*, *13*(22), 1–21. https://doi.org/10.3390/su132212424

Loe, L., Mansfield, K. P., & Rowlands, E. (1994). Appearance of lit environment and its relevance in lighting design: Experimental study. *Lighting Research & Technology*, *26*(3), 119–133. https://doi.org/10.1177/096032719402600301

Lok, R., Smolders, K. C. H. J., Beersma, D. G. M., & Kort, Y. A. W. De. (2018). *Light , Alertness , and Alerting Effects of White Light : A Literature Overview*. 589–601. https://doi.org/10.1177/0748730418796443

Lu, Z. L., & Sperling, G. (1996). Second-order illusions: Mach bands, Chevreul, and Craik-O'Brien-Cornsweet. *Vision Research*, 36(4), 559–572. https://doi.org/10.1016/0042-6989(95)00139-5

Marsden, A. M. (1969). Brightness - a review of current knowledge. *Lighting Research and Technology*, *1*, 3.

Marsden, A. M. (1970). Brightness-luminance relationsihps in an interior. *Lighting Research and Technology*, *2*(1), 10–16.

Masuda, A., Watanabe, J., Terao, M., Watanabe, M., Yagi, A., & Maruya, K. (2011). Awareness of central luminance edge is crucial for the Craik-O'Brien-Cornsweet effect. *Frontiers in Human Neuroscience*, *5*(OCTOBER), 1–9. https://doi.org/10.3389/fnhum.2011.00125

Masuda, A., Watanabe, J., Terao, M., Yagi, A., & Maruya, K. (2014). A temporal window for estimating surface brightness in the Craik-O'Brien-Cornsweet effect. *Frontiers in Human Neuroscience*, 8. doi:10.3389/fnhum.2014.00855

Meier, B. P., Robinson, M. D., Crawford, L. E., & Ahlvers, W. J. (2007). When "light" and "dark" thoughts become light and dark responses: Affect biases brightness judgments. *Emotion*, *7*(2), 366–376. https://doi.org/10.1037/1528-3542.7.2.366

Moscoso, C., Chamilothori, K., Wienold, J., Andersen, M., & Matusiak, B. (2021). Window Size Effects on Subjective Impressions of Daylit Spaces: Indoor Studies at High Latitudes Using Virtual Reality. *LEUKOS - Journal of Illuminating Engineering Society of North America*, *17*(3), 242–264. https://doi.org/10.1080/15502724.2020.1726183

Natephra, W., Motamedi, A., Fukuda, T., & Yabuki, N. (2017). Integrating building information modeling and virtual reality development engines for building indoor lighting design. *Visualization in Engineering*, 5(1). https://doi.org/10.1186/s40327-017-0058-x

Nayatani, Y. (1998). A colorimetric explanation of the Helmholtz-Kohlrausch effect. *Color Research and Application*, 23(6), 374–378. https://doi.org/10.1002/(SICI)1520-6378(199812)23:6<374::AID-COL5>3.0.CO;2-W

Newsham, G. R., Marchand, R. G., & Veitch, J. A. (2004). Preferred surface luminances in offices, by evolution. *Journal of the Illuminating Engineering Society*, *33*(1), 14–29. https://doi.org/10.1080/00994480.2004.10748423

Newsham G, Richardson C, Blanchet C, Veitch J. Lighting quality research using rendered images of offices. *Lighting Research & Technology*. 2005;37(2):93-112. doi:10.1191/1365782805li132oa

Nichols, S., & Patel, H. (2002). Health and safety implications of virtual reality: a review of empirical evidence. *Applied Ergonomics*, 33(3), 251–271. doi:10.1016/s0003-6870(02)00020-0

Oculus Developers (n.d.). *Set Specific Color Space in Unity*. Meta Quest Developer Center. URL https://developer.oculus.com/documentation/unity/unity-color-space/

Paradiso, M. A., & Nakayama, K. (1990). Brightness perception and Filling-In. *Vision Research*, *31*(7), 1221–1236. https://doi.org/10.1016/B978-012370880-9.00296-6

Perdreau, F., & Cavanagh, P. (2011). *Do Artists See Their Retinas ? Do artists see their retinas ? May 2014*. https://doi.org/10.3389/fnhum.2011.00171

Posit team (2023). RStudio: Integrated Development Environment for R. Posit Software, PBC, Boston, MA. URL http://www.posit.co/

Pracki, P., & Krupiński, R. (2021). Brightness and uniformity perception of virtual corridor with artificial lighting systems. *Energies*, 14(2). https://doi.org/10.3390/en14020412

Purves, D., Williams, S. M., Nundy, S., & Lotto, R. B. (2004). Perceiving the Intensity of Light. *Psychological Review*, 111(1), 142–158. https://doi.org/10.1037/0033-295X.111.1.142

Ratliff, F. (1972). Contour and Contrast. *Scientific American*, *226*(6), 90–103.

Rea, M. S., Mou, X., & Bullough, J. D. (2016). Scene brightness of illuminated interiors. *Lighting Research and Technology*, *48*(7), 823–831. https://doi.org/10.1177/1477153515581412

Rockcastle, S., Danell, M., Calabrese, E., Sollom-Brotherton, G., Mahic, A., Van Den Wymelenberg, K., & Davis, R. (2021). Comparing perceptions of a dimmable LED lighting system between a real space and a virtual reality display. *Lighting Research and Technology*, *53*(8), 701–725. https://doi.org/10.1177/1477153521990039

Rossi, A. F., Rittenhouse, C. D., & Paradiso, M. A. (1996). The representation of brightness in primary visual cortex. *Science*, 273(5278), 1104–1107. https://doi.org/10.1126/science.273.5278.1104

Sandoval Salinas, C., Hermans, S., Sandoval, J., Smet, K. A. G., Hanselaer, P., & Colombo, E. (2020). Relationship between pupillary size, brightness, and photoreceptor responses for unrelated self-luminous stimuli at low photopic light levels. *Color Research and Application*, *45*(6), 977–991. https://doi.org/10.1002/col.22546

Scorpio, M., Laffi, R., Teimoorzadeh, A., & Sibilio, S. (2021). Immersive virtual reality as a tool for lighting design: Applications and opportunities. *Journal of Physics: Conference Series*, *2042*(1). https://doi.org/10.1088/1742-6596/2042/1/012125

Shapley, R., & Enroth-Cugell, C. (1984). Visual Adaptation and Retinal Gain Controls. *Engineering Sciences*, 3, 263–346.

Skiljan, I. (2012). IrfanView [computer software]. URL http://irfanview.tuwien.ac.at/

Stevens, S. S. (1960). *The psychophysics of sensory function - into the nature of sensory communication*. *48*(2), 226–253.

Stokkermans, M. G. M., & Heynderickx, I. E. J. (2014). Temporal dark adaptation to spatially complex backgrounds: effect of an additional light source. *JOSA A*, *31*(7), 1485-1494. doi:10.1364/JOSAA.31.001485

Stokkermans, M., Vogels, I., de Kort, Y., & Heynderickx, I. (2018). A Comparison of Methodologies to Investigate the Influence of Light on the Atmosphere of a Space. *LEUKOS - Journal of Illuminating Engineering Society of North America*, *14*(3), 167–191. https://doi.org/10.1080/15502724.2017.1385399

Sullivan, J. T., & Donn, M. (2016). Light Distribution and Spatial Brightness: Relative Importance Of The Walls, Ceiling, And Floor. *Cie*, *September*, 59–69.

Sullivan, J., & Donn, M. (2018). *Measuring the Effect of Light Distribution on Spatial Brightness*. *December 2019*, 356–366. https://doi.org/10.25039/x44.2017.op49

Tiller, D. K., & Veitch, J. A. (1995). Perceived room brightness: Pilot study on the effect of luminance distribution. *Lighting Research and Technology*, *27*(2), 93–101.

Todorović, D. (1987). The Craik-O’Brien-Cornsweet effect: New varieties and their theoretical implications. *Perception & Psychophysics*, *42*(6), 545–560. https://doi.org/10.3758/BF03207986

Van de Perre, L., Smet, K. A. G., Hanselaer, P., Dujardin, M., & Ryckaert, W. R. (2023). The effect of correlated colour temperature and wall luminance on spatial brightness and scene preference in a windowless office setup. *Lighting Research & Technology*, 1–21. https://doi.org/10.1177/14771535231154479

Van Ooyen, M. H. F., Van De Weijgert, J. A. C., & Begemann, S. H. A. (1987). Preferred Luminances in Offices. *Journal of the Illuminating Engineering Society*, *16*(2), 152–156. https://doi.org/10.1080/00994480.1987.10748695

Vogels, I. (2008). Atmosphere Metrics: a tool to quantify perceived atmosphere. *International Symposium Creating an Atmosphere*, *July*. https://doi.org/10.1007/978-1-4020-6593-4

Ward, G. J. (1994). The RADIANCE lighting simulation and rendering system. *In Proceedings of the 21st annual conference on Computer graphics and interactive techniques* (pp. 459-472). [computer software]

Webster, M. A. (2015). Visual Adaptation. *Annual Review of Vision Science*, *1*(1), 547–567. doi:10.1146/annurev-vision-082114-035509

Weerakkody Y, Murphy A, Iqbal S, et al. (2009). *Mach bands*. Reference article, Radiopaedia.org https://doi.org/10.53347/rID-7804

Wong, M. O., Du, J., Zhang, Z. Q., Liu, Y. Q., Chen, S. M., & Lee, S. H. (2019). An experience-based interactive lighting design approach using BIM and VR: a case study. *Earth and Environmental Science*, 238. https://doi.org/10.1088/1755-1315/238/1/012006

Wright, B. D. L. J. (1994). Reasonable mean-square fit values. *Rasch Meas Trans*, 8, 370.

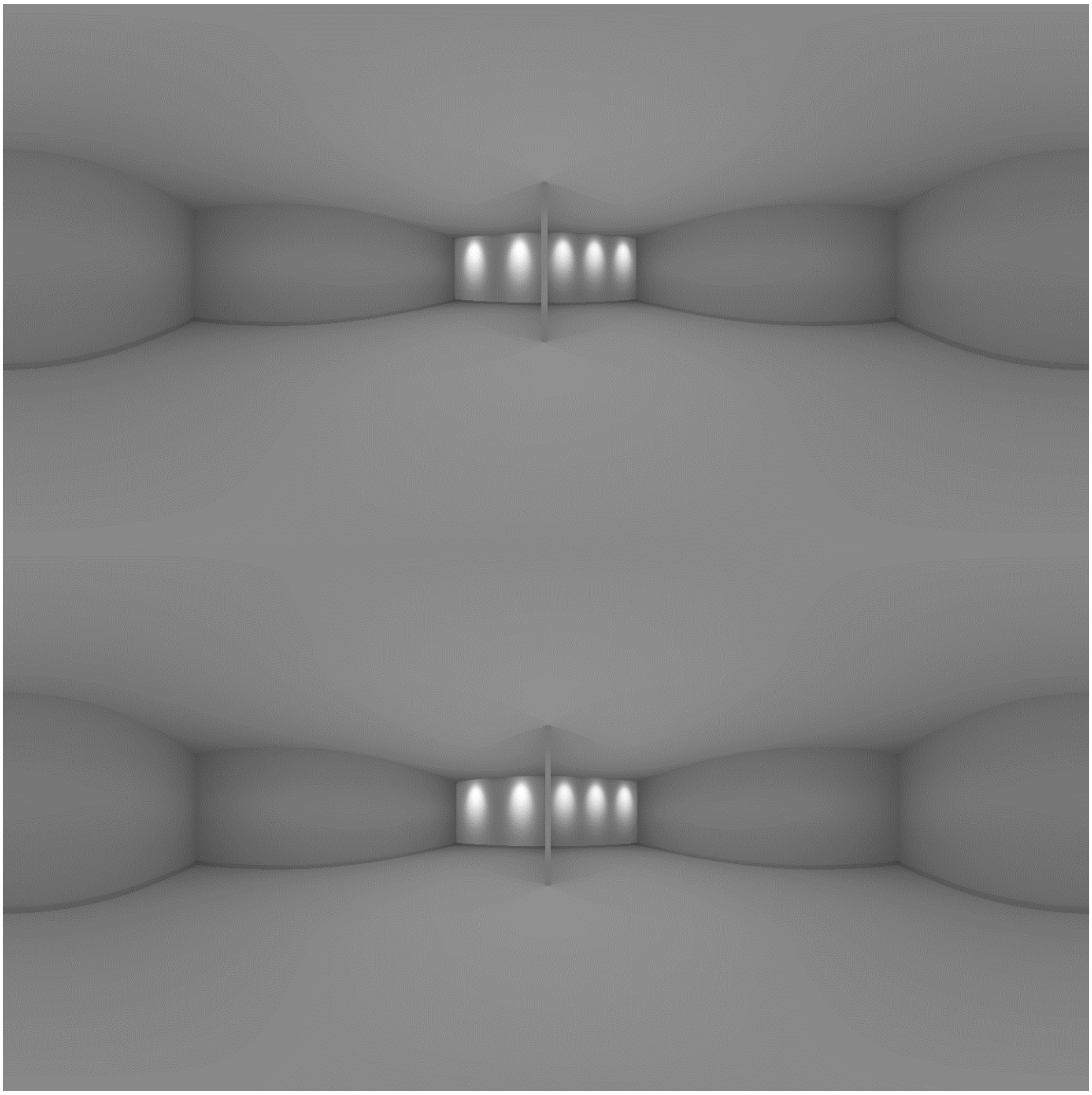
Wright, B. D., & Linacre, J. M. (1989). Observations are always ordinal; measurements, however, must be interval. *Archives of Physical Medicine and Rehabilitation*, *70*(12), 857–860.

Wright, B. D., & Linacre, J. M. (1996). *The Rasch Model as a foundation for the Lexile Framework.*

Yao, Q., Zhong, B., Shi, Y., & Ju, J. (2017). Evaluation of Several Different Types of Uniformity Metrics and Their Correlation with Subjective Perceptions. *LEUKOS - Journal of Illuminating Engineering Society of North America*, *13*(1), 33–45. https://doi.org/10.1080/15502724.2016.1172488

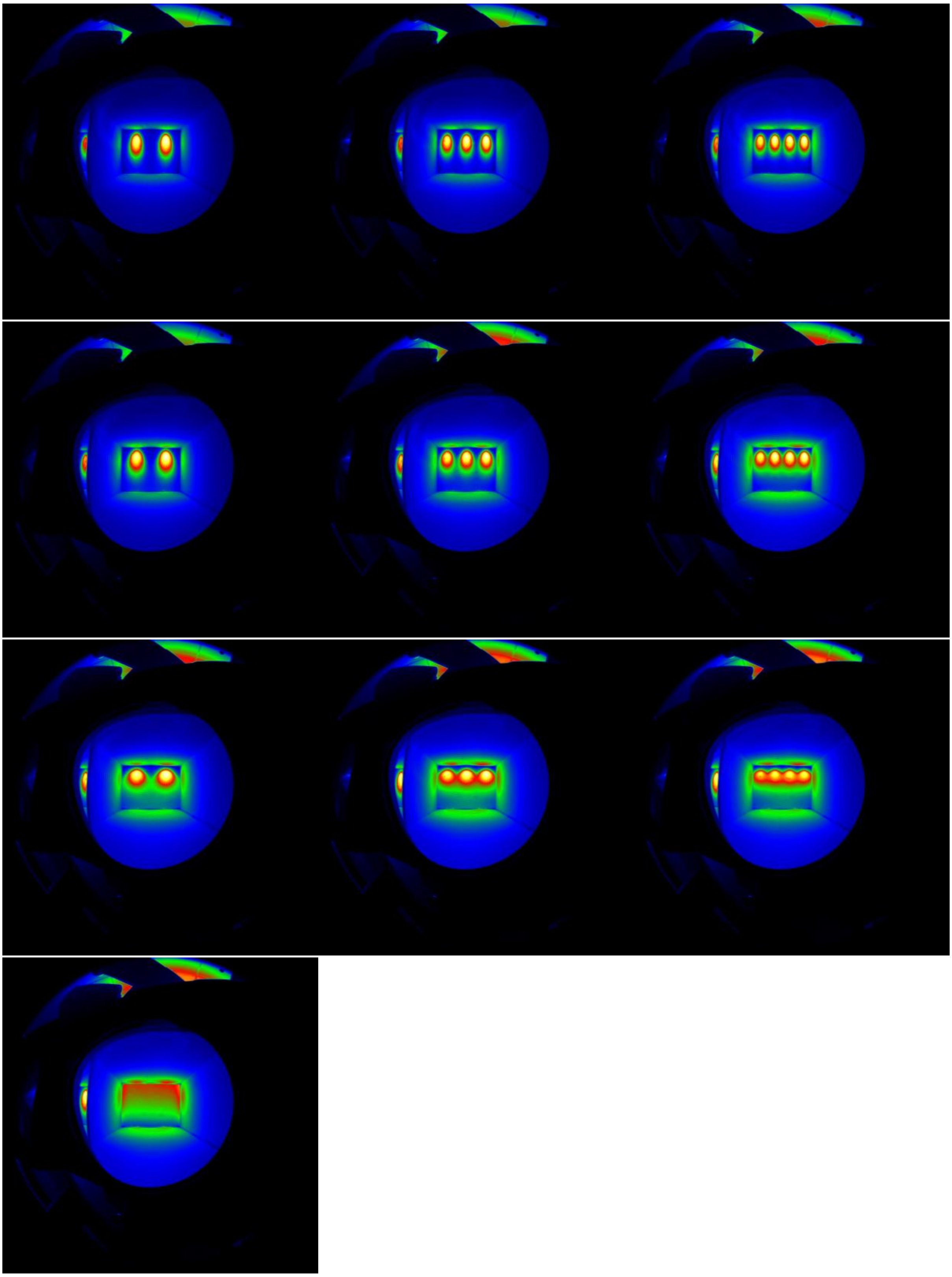
Yilmaz, C. (n.d.) *Office-lighting checklist: How to deliver workplace lighting that meets – and exceeds – the standards*. Philips. URL https://www.lighting.philips.com/main/support/connect/lighting-technology/lighting-design-and-quality/office-lighting-checklist

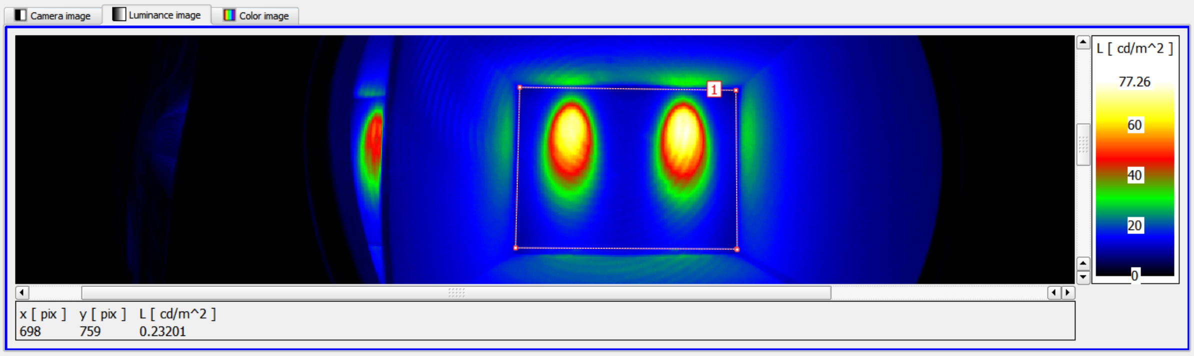
Appendix I



Appendix II

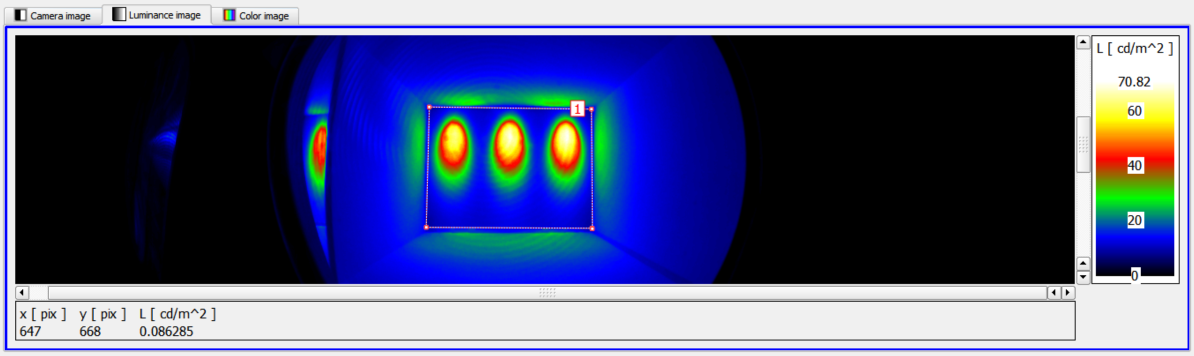
Equal mean set luminance photos and corresponding measurements:





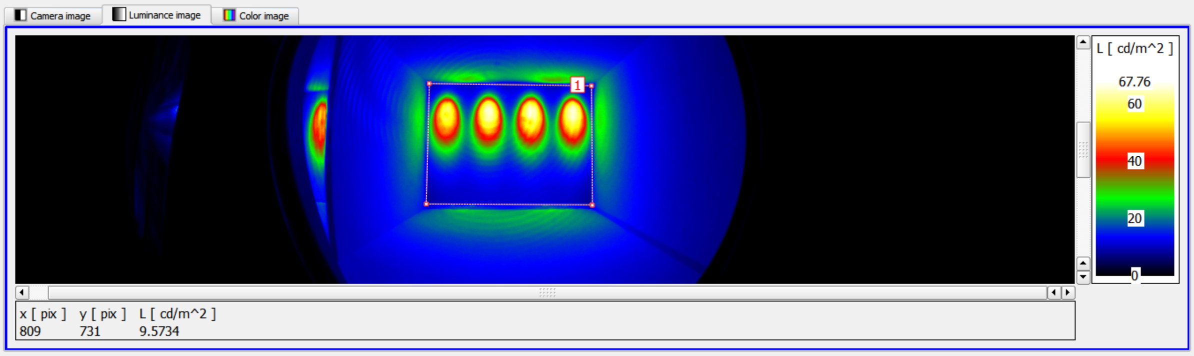
Overview | Last capture | Quantile object

Stat.No.	Parameter	Image	Region	1. Quantile	2. Quantile	3. Quantile	Area	Min	Max	Mean	Disp
							pix^2	cd/m^2	cd/m^2	cd/m^2	cd/m^2
1	Qua_Gr[1]	Luminance image	1	11.5	16.72	62.58	101400	7.028	77.26	24.1	15.88



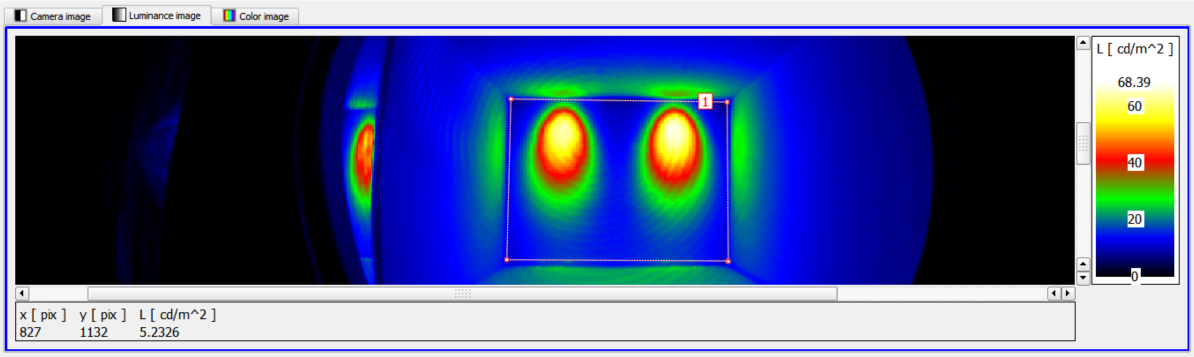
Overview | Last capture | Quantile object

Stat.No.	Parameter	Image	Region	1. Quantile	2. Quantile	3. Quantile	Area	Min	Max	Mean	Disp
							pix^2	cd/m^2	cd/m^2	cd/m^2	cd/m^2
1	Qua_Gr[1]	Luminance image	1	11.39	17.19	57.65	101400	7.332	70.82	23.73	14.51



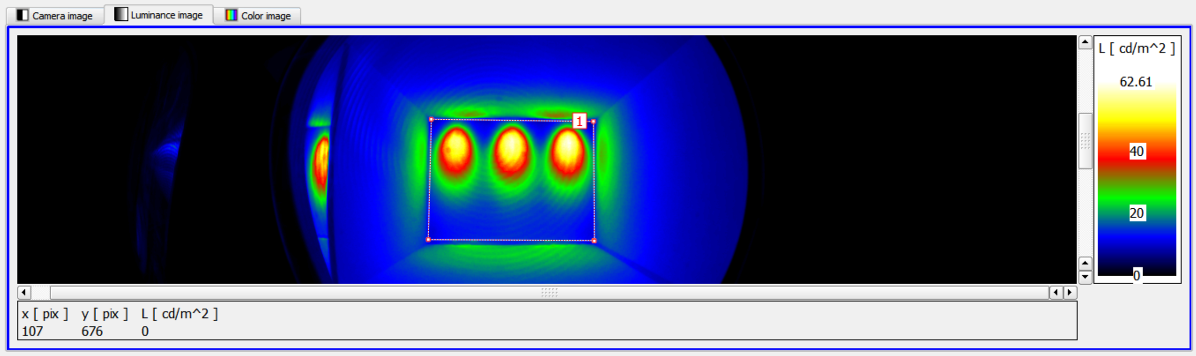
Overview | Last capture | Quantile object

Stat.No.	Parameter	Image	Region	1. Quantile	2. Quantile	3. Quantile	Area	Min	Max	Mean	Disp
							pix^2	cd/m^2	cd/m^2	cd/m^2	cd/m^2
1	Qua_Gr[1]	Luminance image	1	10.68	18.69	55.63	101400	7.467	67.76	23.97	14.3



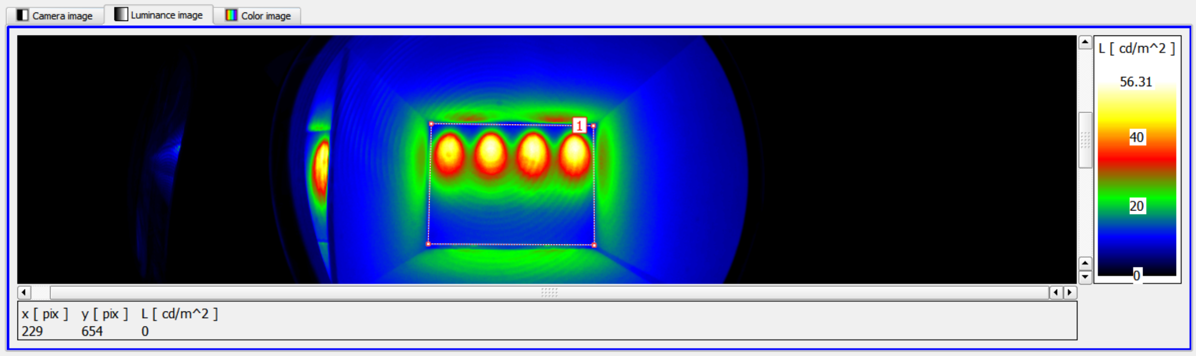
Overview | Last capture | Quantile object

Stat.No.	Parameter	Image	Region	1. Quantile	2. Quantile	3. Quantile	Area	Min	Max	Mean	Disp
							pix^2	cd/m^2	cd/m^2	cd/m^2	cd/m^2
1	Qua_Gr[1]	Luminance image	1	11.59	18.31	55.54	101400	5.862	68.39	24.17	13.54



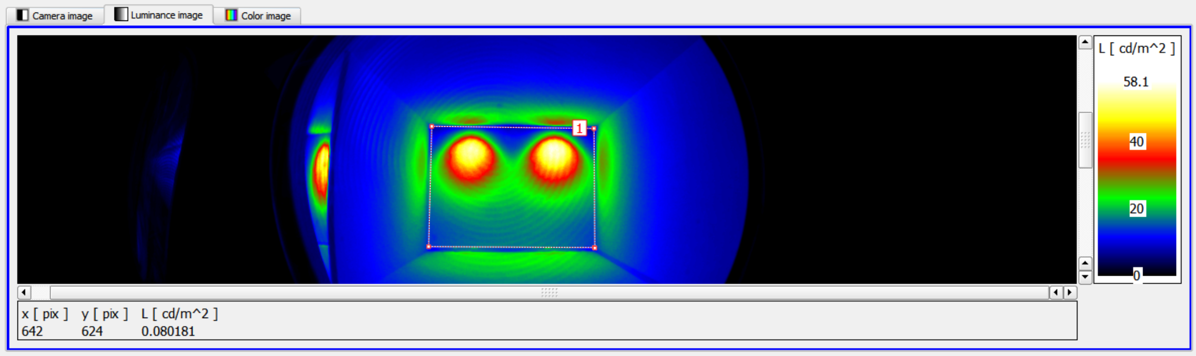
Overview Last capture Quantile object

Stat.No.	Parameter	Image	Region	1. Quantile	2. Quantile	3. Quantile	Area	Min	Max	Mean	Disp
							pix^2	cd/m^2	cd/m^2	cd/m^2	cd/m^2
1	Qua_Gr[1]	Luminance image	1	12.39	19.86	51.16	101400	6.885	62.61	24.1	11.98



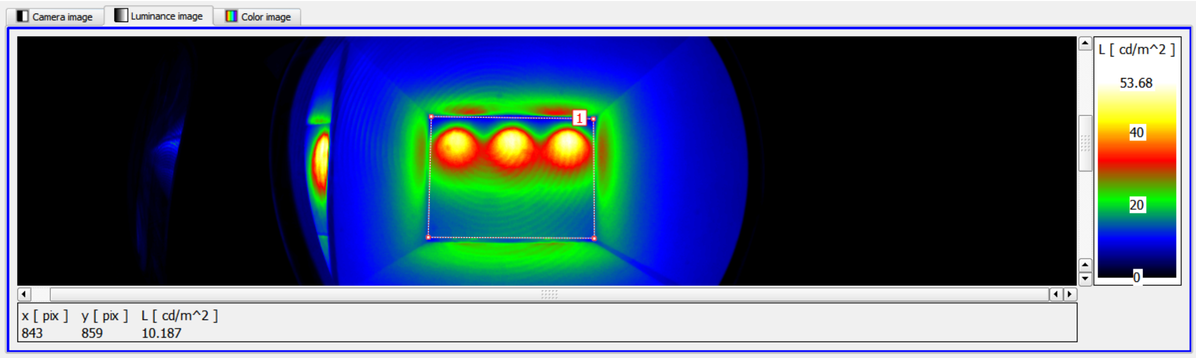
Overview Last capture Quantile object

Stat.No.	Parameter	Image	Region	1. Quantile	2. Quantile	3. Quantile	Area	Min	Max	Mean	Disp
							pix^2	cd/m^2	cd/m^2	cd/m^2	cd/m^2
1	Qua_Gr[1]	Luminance image	1	12.13	20.42	47.21	101400	7.856	56.31	23.77	11.04



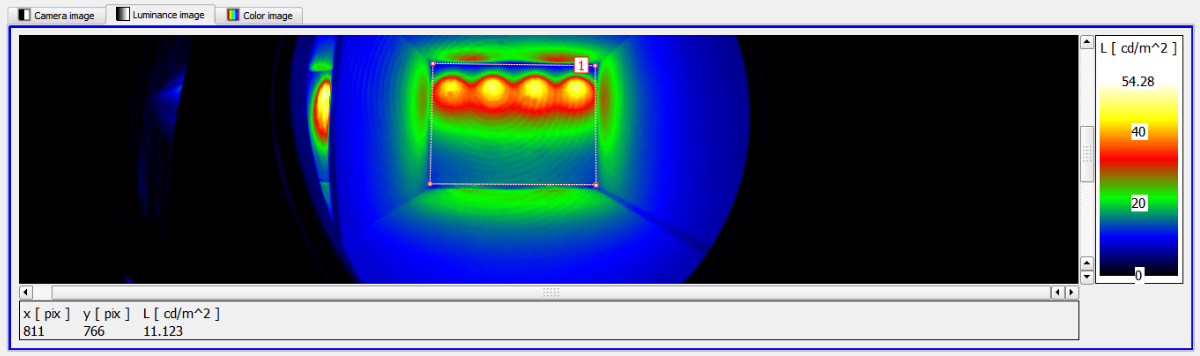
Overview Last capture Quantile object

Stat.No.	Parameter	Image	Region	1. Quantile	2. Quantile	3. Quantile	Area	Min	Max	Mean	Disp
							pix^2	cd/m^2	cd/m^2	cd/m^2	cd/m^2
1	Qua_Gr[1]	Luminance image	1	13.2	20.67	47.89	101400	6.044	58.1	24.07	10.21



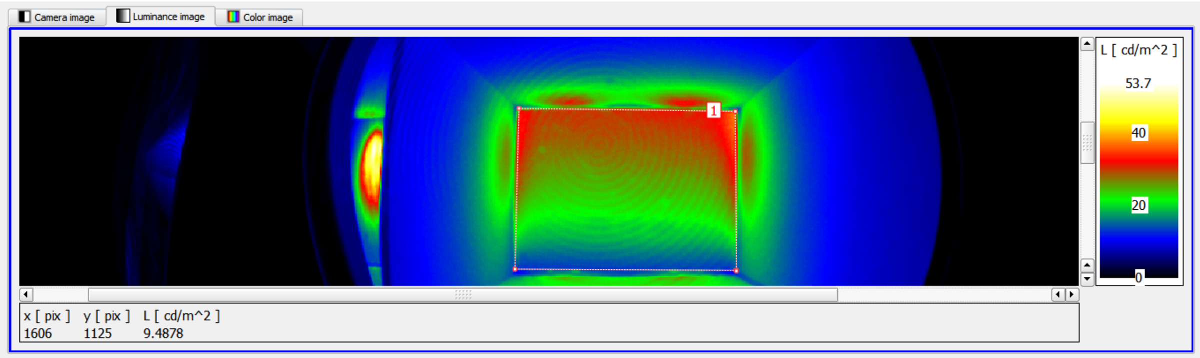
Overview Last capture Quantile object

Stat.No.	Parameter	Image	Region	1. Quantile	2. Quantile	3. Quantile	Area	Min	Max	Mean	Disp
							pix^2	cd/m^2	cd/m^2	cd/m^2	cd/m^2
1	Qua_Gr[1]	Luminance image	1	13.89	20.89	44.08	101400	6.771	53.68	24.1	9.596



Overview | Last capture | Quantile object

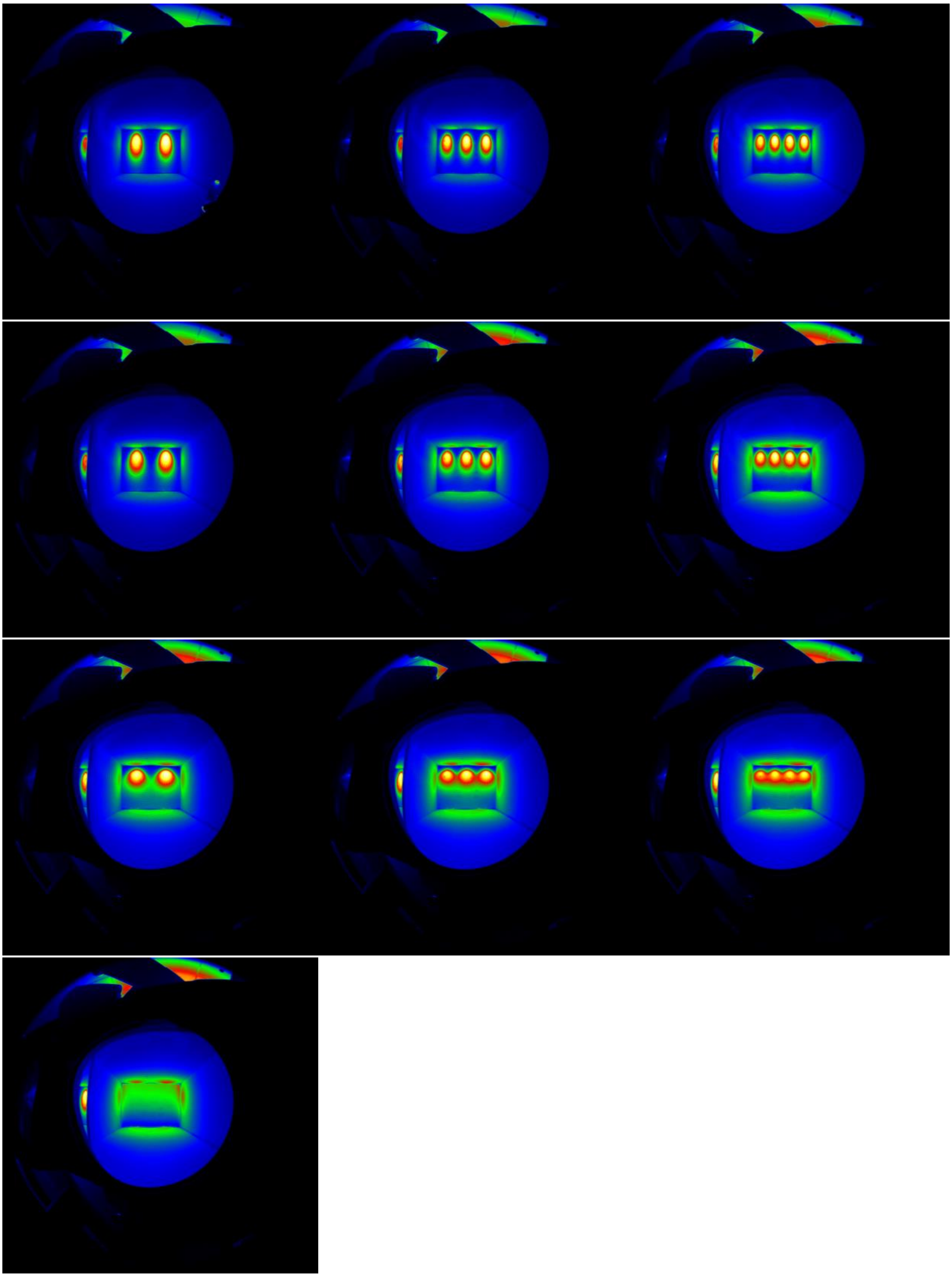
Stat.No.	Parameter	Image	Region	1. Quantile	2. Quantile	3. Quantile	Area	Min	Max	Mean	Disp
							pix^2	cd/m^2	cd/m^2	cd/m^2	cd/m^2
1	Qua_Gr[1]	Luminance image	1	13.95	20.85	41.85	101400	6.965	49.26	24.08	9.356

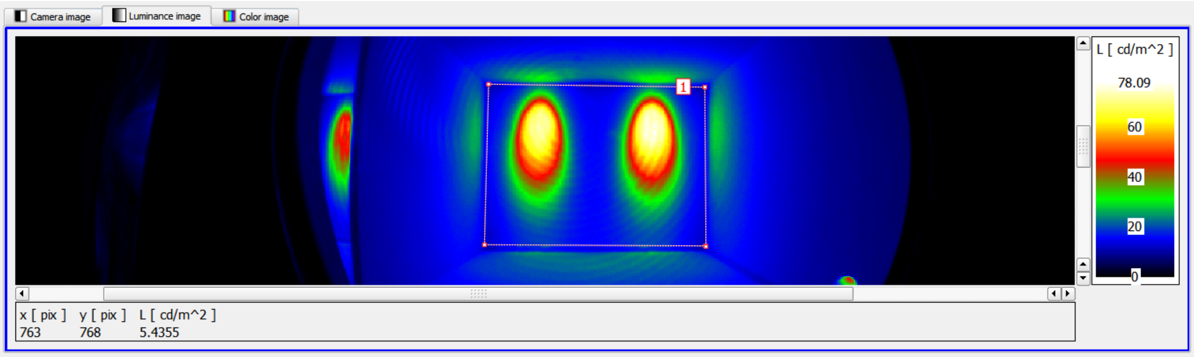


Overview | Last capture | Quantile object

Stat.No.	Parameter	Image	Region	1. Quantile	2. Quantile	3. Quantile	Area	Min	Max	Mean	Disp
							pix^2	cd/m^2	cd/m^2	cd/m^2	cd/m^2
1	Qua_Gr[1]	Luminance image	1	14.52	24.78	31.14	101400	12.09	37.66	23.99	5.29

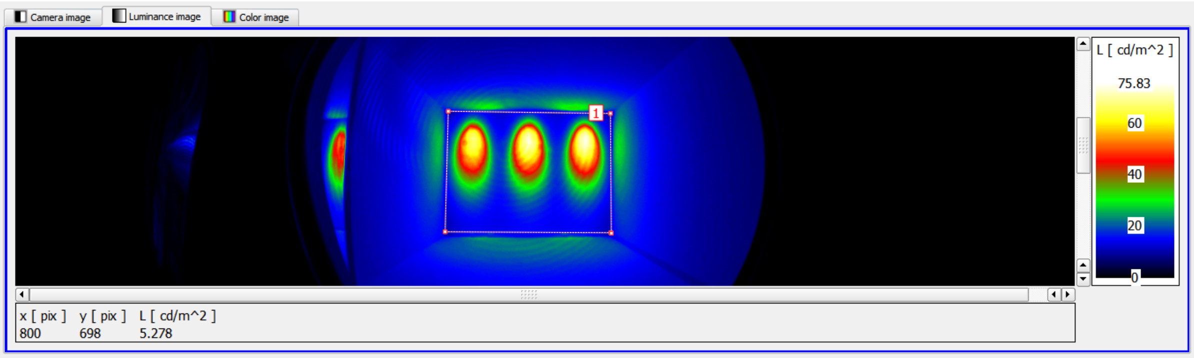
Equal median set luminance photos and corresponding measurements:





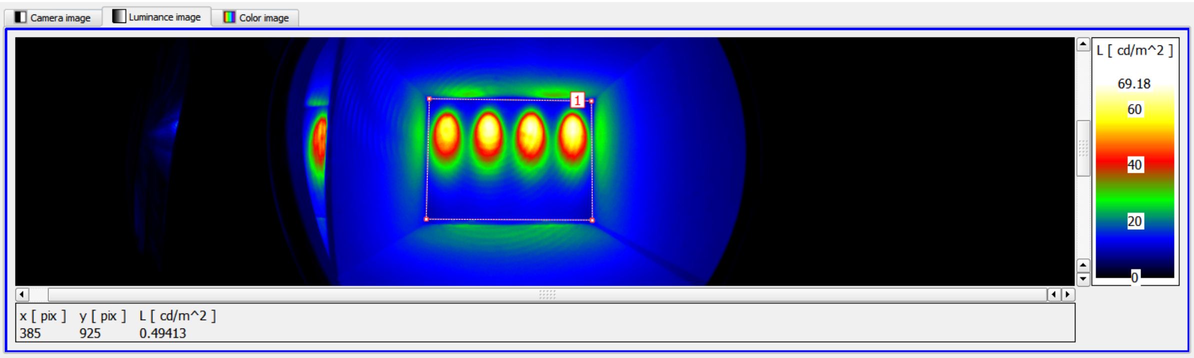
Overview | Last capture | Quantile object

Stat.No.	Parameter	Image	Region	1. Quantile	2. Quantile	3. Quantile	Area	Min	Max	Mean	Disp
							pix^2	cd/m^2	cd/m^2	cd/m^2	cd/m^2
1	Qua_Gr[1]	Luminance image	1	14.3	20.7	66.71	101400	9.076	78.09	28.07	16.26



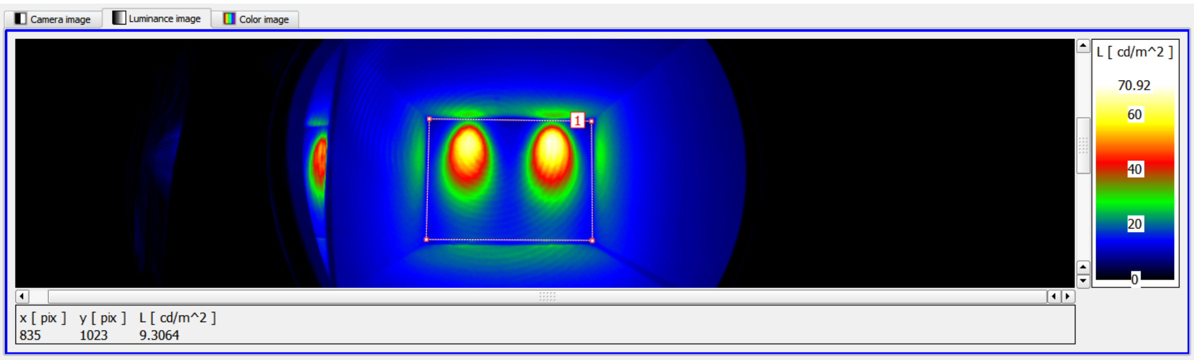
Overview | Last capture | Quantile object

Stat.No.	Parameter	Image	Region	1. Quantile	2. Quantile	3. Quantile	Area	Min	Max	Mean	Disp
							pix^2	cd/m^2	cd/m^2	cd/m^2	cd/m^2
1	Qua_Gr[1]	Luminance image	1	13.67	20.33	62.56	101400	8.981	75.83	27.22	15.39



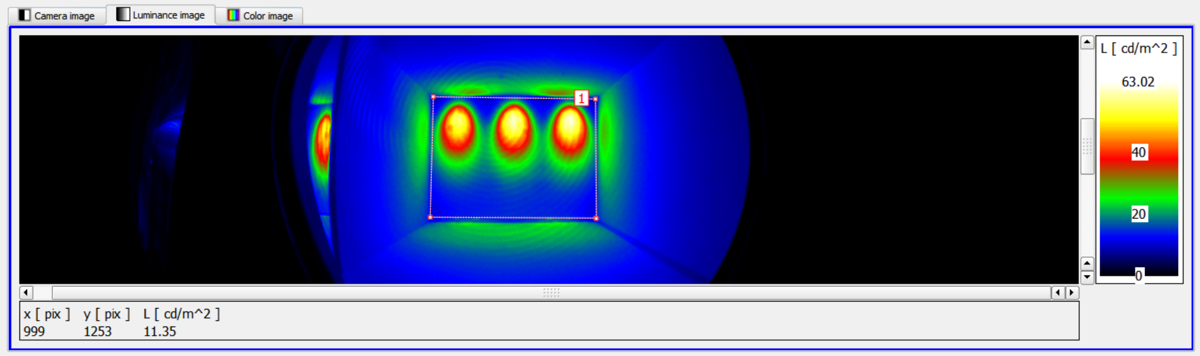
Overview | Last capture | Quantile object

Stat.No.	Parameter	Image	Region	1. Quantile	2. Quantile	3. Quantile	Area	Min	Max	Mean	Disp
							pix^2	cd/m^2	cd/m^2	cd/m^2	cd/m^2
1	Qua_Gr[1]	Luminance image	1	11.75	20.24	57.44	101400	8.173	69.18	25.59	14.57



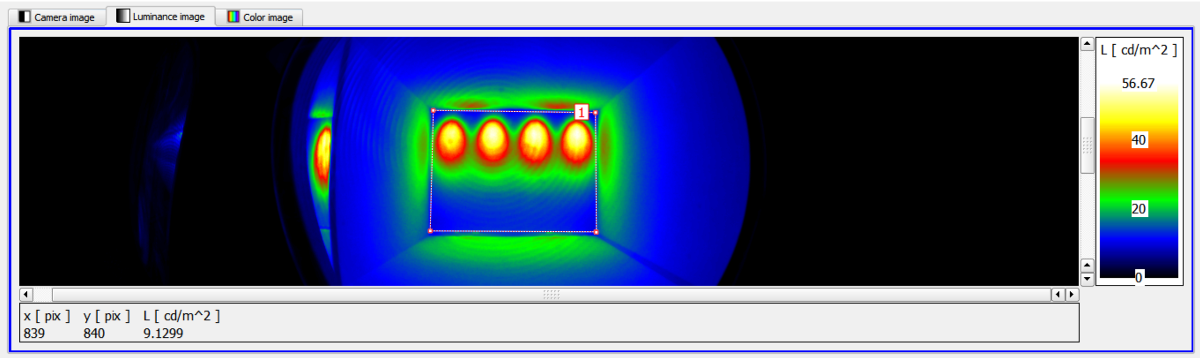
Overview | Last capture | Quantile object

Stat.No.	Parameter	Image	Region	1. Quantile	2. Quantile	3. Quantile	Area	Min	Max	Mean	Disp
							pix^2	cd/m^2	cd/m^2	cd/m^2	cd/m^2
1	Qua_Gr[1]	Luminance image	1	12.71	19.87	57.85	101400	6.764	70.92	26.03	14.03



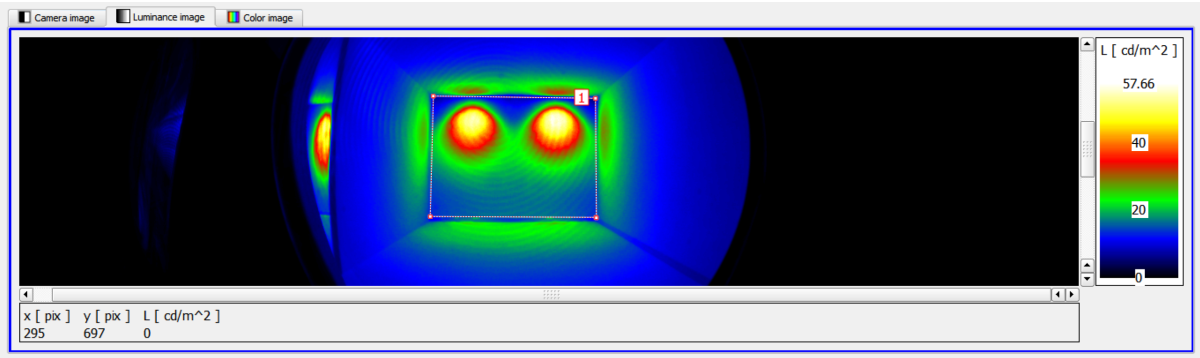
Overview | Last capture | Quantile object

Stat.No.	Parameter	Image	Region	1. Quantile	2. Quantile	3. Quantile	Area	Min	Max	Mean	Disp
							pix^2	cd/m^2	cd/m^2	cd/m^2	cd/m^2
1	Qua_Gr[1]	Luminance image	1	12.4	19.85	51.11	101400	7.098	63.02	24.11	11.97



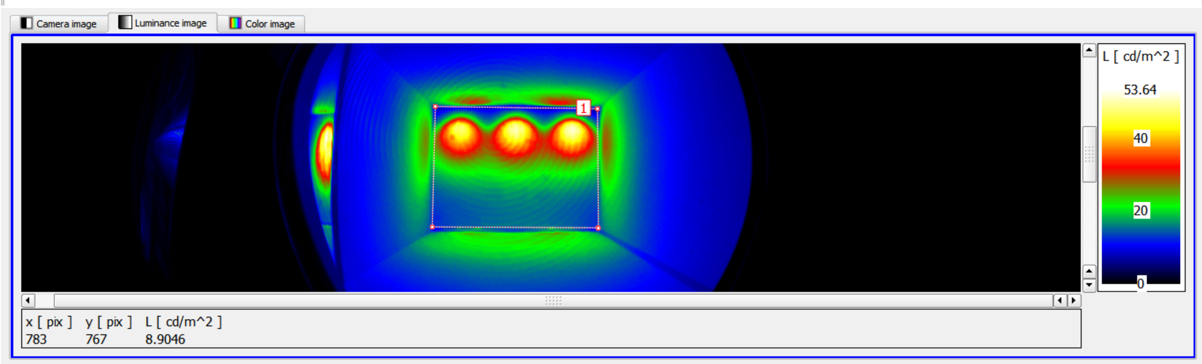
Overview | Last capture | Quantile object

Stat.No.	Parameter	Image	Region	1. Quantile	2. Quantile	3. Quantile	Area	Min	Max	Mean	Disp
							pix^2	cd/m^2	cd/m^2	cd/m^2	cd/m^2
1	Qua_Gr[1]	Luminance image	1	11.62	19.71	47.3	101400	7.698	56.67	23.36	11.26



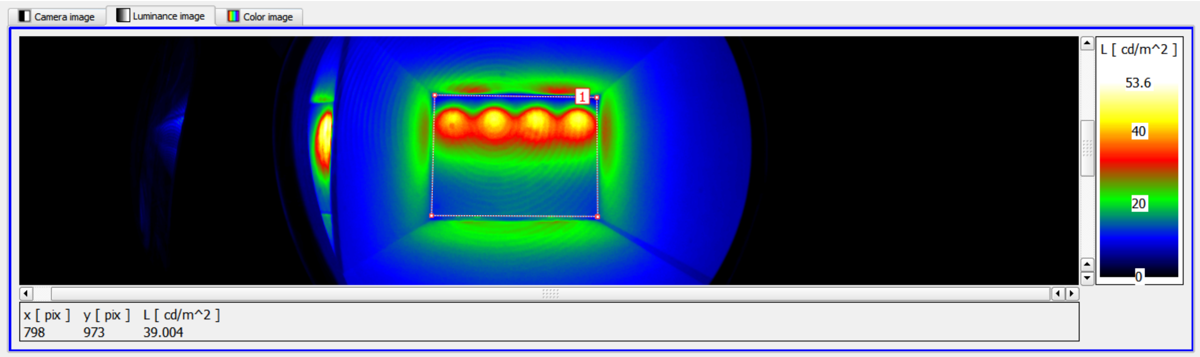
Overview | Last capture | Quantile object

Stat.No.	Parameter	Image	Region	1. Quantile	2. Quantile	3. Quantile	Area	Min	Max	Mean	Disp
							pix^2	cd/m^2	cd/m^2	cd/m^2	cd/m^2
1	Qua_Gr[1]	Luminance image	1	12.66	19.85	47.35	101400	5.757	57.66	23.37	10.23



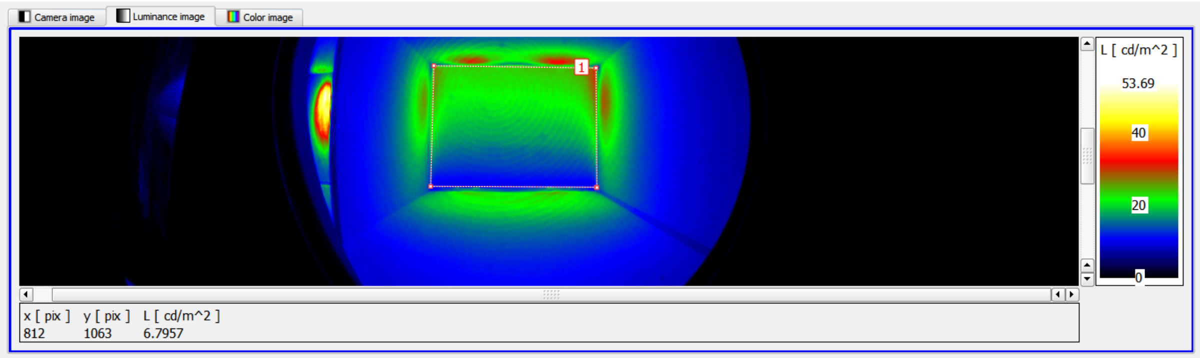
Overview | Last capture | Quantile object

Stat.No.	Parameter	Image	Region	1. Quantile	2. Quantile	3. Quantile	Area	Min	Max	Mean	Disp
							pix^2	cd/m^2	cd/m^2	cd/m^2	cd/m^2
1	Qua_Gr[1]	Luminance image	1	13.08	19.55	42.42	101400	6.211	52.57	22.84	9.365



Overview | Last capture | Quantile object

Stat.No.	Parameter	Image	Region	1. Quantile	2. Quantile	3. Quantile	Area	Min	Max	Mean	Disp
							pix^2	cd/m^2	cd/m^2	cd/m^2	cd/m^2
1	Qua_Gr[1]	Luminance image	1	13.15	19.62	40.28	101400	6.59	47.85	22.88	9.057



Overview | Last capture | Quantile object

Stat.No.	Parameter	Image	Region	1. Quantile	2. Quantile	3. Quantile	Area	Min	Max	Mean	Disp
							pix^2	cd/m^2	cd/m^2	cd/m^2	cd/m^2
1	Qua_Gr[1]	Luminance image	1	10.83	19.13	24.33	101400	9.511	31.08	18.53	4.337

Appendix III

Person statistics for equal mean set of conditions on condition level.

Total Score	Total Count	Obsvd Average	Fair(M) Average	+ Measure	Model S.E.	Infit MnSq	Infit ZStd	Outfit MnSq	Outfit ZStd	Estim. Discrm	Displ.	Correlation PtMea	Correlation PtExp	Nu	idnr
51	90	.57	.51 A	.00	.22	1.82	9.0	1.95	8.8	-2.65	.26	-.70	.35	6	6
34	90	.38	.51 A	.00	.22	1.61	7.2	1.73	7.0	-1.75	-.60	-.44	.35	4	4
34	90	.38	.51 A	.00	.22	1.53	6.3	1.60	6.0	-1.33	-.60	-.31	.35	28	28
58	90	.64	.51 A	.00	.22	1.47	5.7	1.56	5.6	-1.11	.61	-.28	.35	12	12
44	90	.49	.51 A	.00	.22	1.44	5.3	1.52	5.2	-.96	-.09	-.21	.35	2	2
50	90	.56	.51 A	.00	.22	1.31	3.9	1.34	3.6	-.37	.21	-.05	.35	25	25
30	90	.33	.51 A	.00	.22	1.05	.6	1.08	.9	.76	-.80	.31	.35	14	14
41	90	.46	.51 A	.00	.22	1.05	.7	1.06	.7	.77	-.24	.29	.35	19	19
44	90	.49	.51 A	.00	.22	1.02	.2	1.04	.4	.90	-.09	.32	.35	11	11
57	90	.63	.51 A	.00	.22	1.03	.4	1.04	.4	.87	.56	.30	.35	9	9
63	90	.70	.51 A	.00	.22	.97	-.3	.96	-.4	1.12	.87	.39	.35	15	15
32	90	.36	.51 A	.00	.22	.96	-.5	.96	-.4	1.18	-.70	.43	.35	10	10
49	90	.54	.51 A	.00	.22	.96	-.5	.98	-.2	1.16	.16	.39	.35	3	3
36	90	.40	.51 A	.00	.22	.94	-.8	.93	-.8	1.28	-.50	.44	.35	21	21
52	90	.58	.51 A	.00	.22	.89	-1.5	.89	-1.3	1.47	.31	.48	.35	17	17
47	90	.52	.51 A	.00	.22	.87	-1.8	.88	-1.4	1.55	.06	.50	.35	8	8
38	90	.42	.51 A	.00	.22	.88	-1.7	.86	-1.6	1.53	-.39	.51	.35	7	7
47	90	.52	.51 A	.00	.22	.86	-1.9	.84	-1.9	1.61	.06	.52	.35	5	5
46	90	.51	.51 A	.00	.22	.82	-2.6	.79	-2.6	1.81	.01	.57	.35	22	22
62	90	.69	.51 A	.00	.22	.80	-2.9	.77	-2.9	1.89	.82	.62	.35	13	13
48	90	.53	.51 A	.00	.22	.78	-3.2	.76	-3.0	1.96	.11	.61	.35	24	24
44	90	.49	.51 A	.00	.22	.75	-3.7	.72	-3.6	2.11	-.09	.66	.35	23	23
53	90	.59	.51 A	.00	.22	.73	-4.1	.69	-4.0	2.19	.36	.69	.35	1	1
41	90	.46	.51 A	.00	.22	.71	-4.4	.69	-4.1	2.26	-.24	.71	.35	26	26
51	90	.57	.51 A	.00	.22	.70	-4.5	.68	-4.2	2.29	.26	.71	.35	20	20
43	90	.48	.51 A	.00	.22	.69	-4.8	.66	-4.5	2.36	-.14	.73	.35	16	16
42	90	.47	.51 A	.00	.22	.67	-5.2	.64	-4.9	2.46	-.19	.76	.35	18	18
46	90	.51	.51 A	.00	.22	.64	-5.8	.61	-5.4	2.59	.01	.79	.35	27	27
45.8	90.0	.51	.51	.00	.22	1.00	-.4	1.01	-.3			.35		Mean (Count: 28)	
8.4	.0	.09	.00	.00	.00	.31	4.0	.35	3.8			.39		S.D. (Population)	
8.5	.0	.09	.00	.00	.00	.31	4.1	.36	3.9			.40		S.D. (Sample)	

Model, Populn: RMSE .22 Adj (True) S.D. .00 Separation .00 Strata .33 Reliability .00
Model, Sample: RMSE .22 Adj (True) S.D. .00 Separation .00 Strata .33 Reliability .00
Model, Fixed (all same) chi-squared: .0 d.f.: 27 significance (probability): 1.00

Person statistics for equal mean set of conditions on factor level.

Total Score	Total Count	Obsvd Average	Fair(M) Average	+ Measure	Model S.E.	Infit MnSq	Infit ZStd	Outfit MnSq	Outfit ZStd	Estim. Discrm	Displ.	Correlation PtMea	Correlation PtExp	Nu	idnr
42	72	.58	.51 A	.00	.25	1.72	7.6	1.81	7.4	-2.41	.33	-.65	.34	6	6
26	72	.36	.51 A	.00	.25	1.61	6.6	1.67	6.3	-1.86	-.67	-.48	.34	28	28
23	72	.32	.51 A	.00	.25	1.54	5.9	1.62	5.9	-1.56	-.85	-.40	.34	4	4
34	72	.47	.51 A	.00	.25	1.49	5.5	1.55	5.3	-1.33	-.17	-.32	.34	2	2
40	72	.56	.51 A	.00	.25	1.40	4.5	1.45	4.5	-.89	.21	-.21	.34	25	25
50	72	.69	.51 A	.00	.25	1.37	4.2	1.43	4.3	-.77	.83	-.21	.34	12	12
35	72	.49	.51 A	.00	.25	1.13	1.5	1.15	1.5	.39	-.10	.16	.34	19	19
23	72	.32	.51 A	.00	.25	1.10	1.2	1.12	1.3	.52	-.85	.23	.34	10	10
44	72	.61	.51 A	.00	.25	1.07	.8	1.08	.9	.67	.46	.23	.34	9	9
28	72	.39	.51 A	.00	.25	1.02	.2	1.03	.3	.89	-.54	.32	.34	14	14
29	72	.40	.51 A	.00	.25	1.02	.2	1.02	.2	.92	-.48	.33	.34	21	21
50	72	.69	.51 A	.00	.25	.97	-.4	.96	-.3	1.16	.83	.39	.34	15	15
39	72	.54	.51 A	.00	.25	.96	-.4	.95	-.5	1.18	.14	.38	.34	11	11
31	72	.43	.51 A	.00	.25	.94	-.7	.94	-.7	1.27	-.35	.42	.34	7	7
37	72	.51	.51 A	.00	.25	.89	-1.4	.89	-1.2	1.52	.02	.48	.34	8	8
38	72	.53	.51 A	.00	.25	.88	-1.5	.87	-1.5	1.56	.08	.49	.34	22	22
42	72	.58	.51 A	.00	.25	.87	-1.7	.86	-1.6	1.62	.33	.50	.34	17	17
52	72	.72	.51 A	.00	.25	.86	-1.8	.84	-1.9	1.66	.96	.55	.34	13	13
39	72	.54	.51 A	.00	.25	.78	-2.9	.75	-3.0	2.03	.14	.62	.34	24	24
32	72	.44	.51 A	.00	.25	.74	-3.6	.73	-3.3	2.20	-.29	.68	.34	26	26
34	72	.47	.51 A	.00	.25	.74	-3.6	.71	-3.6	2.23	-.17	.68	.34	23	23
34	72	.47	.51 A	.00	.25	.72	-3.8	.70	-3.7	2.28	-.17	.70	.34	16	16
37	72	.51	.51 A	.00	.25	.72	-3.9	.70	-3.8	2.30	.02	.70	.34	5	5
33	72	.46	.51 A	.00	.25	.70	-4.2	.67	-4.2	2.41	-.23	.74	.34	18	18
41	72	.57	.51 A	.00	.25	.70	-4.2	.67	-4.2	2.41	.27	.73	.34	1	1
40	72	.56	.51 A	.00	.25	.70	-4.3	.67	-4.2	2.42	.21	.73	.34	20	20
38	72	.53	.51 A	.00	.25	.67	-4.7	.64	-4.6	2.54	.08	.77	.34	3	3
36	72	.50	.51 A	.00	.25	.65	-5.0	.62	-4.9	2.63	-.04	.79	.34	27	27
36.7	72.0	.51	.51	.00	.25	1.00	-.4	1.00	-.3			.33		Mean (Count: 28)	
7.2	.0	.10	.00	.00	.00	.31	3.7	.34	3.7			.42		S.D. (Population)	
7.4	.0	.10	.00	.00	.00	.31	3.8	.35	3.7			.42		S.D. (Sample)	

Model, Populn: RMSE .25 Adj (True) S.D. .00 Separation .00 Strata .33 Reliability .00
Model, Sample: RMSE .25 Adj (True) S.D. .00 Separation .00 Strata .33 Reliability .00
Model, Fixed (all same) chi-squared: .0 d.f.: 27 significance (probability): 1.00

Person statistics for equal median set of conditions on condition level.

Total Score	Total Count	Obsvd Average	Fair(M) Average	+ Measure	Model S.E.	Infit MnSq	Infit ZStd	Outfit MnSq	Outfit ZStd	Estim. Discrm	Displ.	Correlation PtMea	Correlation PtExp	Nu	idnr
46	90	.51	.51 A	.00	.22	1.31	4.9	1.36	5.1	-1.23	-.01	-.26	.27	18	18
35	90	.39	.51 A	.00	.22	1.24	3.9	1.28	4.0	-.72	-.54	-.12	.27	10	10
48	90	.53	.51 A	.00	.22	1.23	3.7	1.27	3.8	-.65	.09	-.13	.27	11	11
48	90	.53	.51 A	.00	.22	1.14	2.4	1.15	2.2	.00	.09	.03	.27	23	23
40	90	.44	.51 A	.00	.22	1.13	2.2	1.14	2.1	.07	-.30	.06	.27	14	14
47	90	.52	.51 A	.00	.22	1.10	1.7	1.12	1.7	.27	.04	.09	.27	27	27
36	90	.40	.51 A	.00	.22	1.08	1.4	1.10	1.6	.37	-.49	.15	.27	19	19
63	90	.70	.51 A	.00	.22	1.08	1.4	1.10	1.5	.40	.81	.10	.27	13	13
47	90	.52	.51 A	.00	.22	1.08	1.3	1.07	1.0	.49	.04	.15	.27	2	2
50	90	.56	.51 A	.00	.22	1.04	.7	1.03	.4	.74	.18	.20	.27	6	6
39	90	.43	.51 A	.00	.22	1.03	.5	1.04	.5	.77	-.35	.23	.27	26	26
56	90	.62	.51 A	.00	.22	1.03	.6	1.03	.5	.76	.47	.20	.27	15	15
37	90	.41	.51 A	.00	.22	1.03	.5	1.03	.5	.80	-.44	.25	.27	22	22
47	90	.52	.51 A	.00	.22	.98	-.2	.99	-.1	1.11	.04	.29	.27	1	1
41	90	.46	.51 A	.00	.22	.98	-.3	.98	-.3	1.13	-.25	.31	.27	7	7
44	90	.49	.51 A	.00	.22	.99	-.1	.97	-.3	1.08	-.11	.29	.27	4	4
46	90	.51	.51 A	.00	.22	.96	-.7	.95	-.7	1.30	-.01	.34	.27	9	9
49	90	.54	.51 A	.00	.22	.96	-.7	.94	-.9	1.32	.13	.34	.27	12	12
37	90	.41	.51 A	.00	.22	.95	-.8	.94	-.9	1.36	-.44	.38	.27	28	28
46	90	.51	.51 A	.00	.22	.93	-1.2	.92	-1.3	1.52	-.01	.39	.27	25	25
49	90	.54	.51 A	.00	.22	.92	-1.4	.90	-1.6	1.60	.13	.41	.27	3	3
47	90	.52	.51 A	.00	.22	.89	-1.9	.88	-1.8	1.75	.04	.45	.27	16	16
55	90	.61	.51 A	.00	.22	.85	-2.7	.84	-2.7	2.06	.42	.52	.27	20	20
51	90	.57	.51 A	.00	.22	.85	-2.8	.83	-2.8	2.09	.23	.52	.27	5	5
56	90	.62	.51 A	.00	.22	.84	-3.0	.82	-2.9	2.15	.47	.54	.27	24	24
35	90	.39	.51 A	.00	.22	.83	-3.2	.81	-3.2	2.23	-.54	.60	.27	21	21
49	90	.54	.51 A	.00	.22	.81	-3.6	.79	-3.6	2.37	.13	.59	.27	8	8
50	90	.56	.51 A	.00	.22	.75	-4.9	.73	-4.7	2.78	.18	.69	.27	17	17
46.2	90.0	.51	.51	.00	.22	1.00	-.1	1.00	-.1			.27		Mean (Count: 28)	
6.8	.0	.08	.00	.00	.00	.13	2.4	.15	2.4			.23		S.D. (Population)	
6.9	.0	.08	.00	.00	.00	.14	2.4	.16	2.4			.23		S.D. (Sample)	

Model, Populn: RMSE .22 Adj (True) S.D. .00 Separation .00 Strata .33 Reliability .00
Model, Sample: RMSE .22 Adj (True) S.D. .00 Separation .00 Strata .33 Reliability .00
Model, Fixed (all same) chi-squared: .0 d.f.: 27 significance (probability): 1.00

Person statistics for equal median set of conditions on factor level.

Total Score	Total Count	Obsvd Average	Fair(M) Average	+ Measure	Model S.E.	Infit MnSq	Infit ZStd	Outfit MnSq	Outfit ZStd	Estim. Discrm	Displ.	Correlation PtMea	Correlation PtExp	Nu	idnr
41	72	.57	.51 A	.00	.24	1.15	3.4	1.16	3.4	-1.40	.19	-.28	.18	6	6
41	72	.57	.51 A	.00	.24	1.16	3.4	1.16	3.4	-1.40	.19	-.28	.18	11	11
28	72	.39	.51 A	.00	.24	1.11	2.4	1.12	2.5	-.71	-.55	-.06	.18	28	28
35	72	.49	.51 A	.00	.24	1.11	2.4	1.11	2.5	-.68	-.15	-.10	.18	4	4
38	72	.53	.51 A	.00	.24	1.08	1.8	1.08	1.7	-.22	.02	-.04	.18	2	2
40	72	.56	.51 A	.00	.24	1.06	1.5	1.07	1.4	.00	.14	-.02	.18	23	23
29	72	.40	.51 A	.00	.24	1.06	1.5	1.07	1.4	.00	-.49	.06	.18	10	10
37	72	.51	.51 A	.00	.24	1.06	1.3	1.06	1.3	.11	-.03	.02	.18	18	18
40	72	.56	.51 A	.00	.24	1.05	1.2	1.06	1.3	.17	.14	.01	.18	12	12
38	72	.53	.51 A	.00	.24	1.03	.6	1.03	.6	.56	.02	.09	.18	25	25
36	72	.50	.51 A	.00	.24	1.02	.5	1.03	.6	.63	-.09	.12	.18	9	9
35	72	.49	.51 A	.00	.24	1.02	.5	1.03	.6	.62	-.15	.13	.18	14	14
31	72	.43	.51 A	.00	.24	1.02	.4	1.02	.5	.69	-.38	.17	.18	19	19
33	72	.46	.51 A	.00	.24	1.01	.2	1.01	.2	.82	-.26	.18	.18	26	26
42	72	.58	.51 A	.00	.24	.98	-.4	.98	-.4	1.28	.25	.20	.18	15	15
51	72	.71	.51 A	.00	.24	.98	-.5	.97	-.5	1.35	.77	.16	.18	13	13
36	72	.50	.51 A	.00	.24	.96	-.9	.96	-.9	1.63	-.09	.30	.18	27	27
27	72	.38	.51 A	.00	.24	.96	-.9	.96	-.9	1.64	-.61	.38	.18	21	21
31	72	.43	.51 A	.00	.24	.95	-1.1	.95	-1.1	1.74	-.38	.36	.18	22	22
40	72	.56	.51 A	.00	.24	.95	-1.1	.95	-1.1	1.75	.14	.29	.18	3	3
33	72	.46	.51 A	.00	.24	.95	-1.2	.94	-1.2	1.84	-.26	.36	.18	7	7
44	72	.61	.51 A	.00	.24	.94	-1.4	.94	-1.4	1.94	.37	.30	.18	20	20
39	72	.54	.51 A	.00	.24	.92	-1.9	.92	-1.9	2.26	.08	.39	.18	16	16
39	72	.54	.51 A	.00	.24	.92	-2.0	.91	-2.0	2.33	.08	.40	.18	8	8
42	72	.58	.51 A	.00	.24	.90	-2.4	.89	-2.4	2.56	.25	.43	.18	1	1
43	72	.60	.51 A	.00	.24	.90	-2.4	.89	-2.4	2.57	.31	.42	.18	5	5
43	72	.60	.51 A	.00	.24	.88	-2.8	.88	-2.8	2.84	.31	.47	.18	24	24
41	72	.57	.51 A	.00	.24	.87	-3.2	.86	-3.2	3.07	.19	.52	.18	17	17
37.6	72.0	.52	.51	.00	.24	1.00	.0	1.00	.0			.18		Mean (Count: 28)	
5.3	.0	.07	.00	.00	.00	.08	1.9	.08	1.9			.21		S.D. (Population)	
5.4	.0	.08	.00	.00	.00	.08	1.9	.08	1.9			.22		S.D. (Sample)	

Model, Populn: RMSE .24 Adj (True) S.D. .00 Separation .00 Strata .33 Reliability .00
Model, Sample: RMSE .24 Adj (True) S.D. .00 Separation .00 Strata .33 Reliability .00
Model, Fixed (all same) chi-squared: .0 d.f.: 27 significance (probability): 1.00

Appendix IV

Full output table of brightness estimates on condition level for the equal mean set of conditions.

```
+-------------------------------------------------------------------------------------------------------------+
| Total   Total   Obsvd  Fair(M)|   +     Model | Infit      Outfit    |Estim.| Correlation |                  |
| Score   Count  Average Average|Measure  S.E.  | MnSq ZStd  MnSq ZStd |Discrm| PtMea PtExp | Nu objects       |
|-------------------------------+---------------+----------------------+------+-------------+------------------|
|  354     504       .70     .69 |   .77    .10 | 1.06   1.4  1.10   1.9 |  .82 |   .38   .44 | 10 uni           |
|  313     504       .62     .62 |   .43    .09 | 1.00    .1  1.01    .2 |  .98 |   .32   .33 |  9 4W            |
|  293     504       .58     .58 |   .27    .09 |  .95 -1.8   .95 -1.7 | 1.28 |   .37   .29 |  7 2W            |
|  292     504       .58     .58 |   .27    .09 | 1.01    .2  1.01    .4 |  .95 |   .28   .29 |  8 3W            |
|  264     504       .52     .52 |   .05    .09 |  .95 -2.0   .95 -1.9 | 1.35 |   .34   .26 |  4 2M            |
|  258     504       .51     .51 A   .00    .09 |  .98 -1.0   .98  -.9 | 1.18 |   .30   .26 |  5 3M            |
|  253     504       .50     .50 |  -.03    .09 |  .98 -1.0   .98 -1.0 | 1.18 |   .30   .26 |  6 4M            |
|  179     504       .36     .36 |  -.61    .10 | 1.06   1.6  1.07   1.8 |  .78 |   .30   .37 |  1 2N            |
|  175     504       .35     .35 |  -.64    .10 | 1.01    .3  1.02    .5 |  .95 |   .36   .38 |  2 3N            |
|  139     504       .28     .29 |  -.95    .10 |  .99   -.1  1.01    .1 | 1.01 |   .48   .48 |  3 4N            |
|-------------------------------+---------------+----------------------+------+-------------+------------------|
|  252.0   504.0     .50     .50 |  -.04    .10 | 1.00   -.2  1.01    .0 |      |   .34       | Mean (Count: 10) |
|   64.5      .0     .13     .12 |   .51    .00 |  .04   1.2   .05   1.3 |      |   .06       | S.D. (Population)|
|   68.0      .0     .13     .13 |   .54    .00 |  .04   1.3   .05   1.4 |      |   .06       | S.D. (Sample)    |
+-------------------------------------------------------------------------------------------------------------+
Model, Populn: RMSE .10  Adj (True) S.D. .50  Separation 5.26  Strata 7.35  Reliability .97
Model, Sample: RMSE .10  Adj (True) S.D. .53  Separation 5.56  Strata 7.74  Reliability .97
Model, Fixed (all same) chi-squared:  269.2  d.f.: 9  significance (probability): .00
Model,  Random (normal) chi-squared:  8.7  d.f.: 8  significance (probability): .37
---------------------------------------------------------------------------------------------------------------
```

Full output table of brightness estimates on condition level for the equal median set of conditions.

```
+-------------------------------------------------------------------------------------------------------------+
| Total   Total   Obsvd  Fair(M)|   +     Model | Infit      Outfit    |Estim.| Correlation |                  |
| Score   Count  Average Average|Measure  S.E.  | MnSq ZStd  MnSq ZStd |Discrm| PtMea PtExp | Nu objects       |
|-------------------------------+---------------+----------------------+------+-------------+------------------|
|  331     504       .66     .65 |   .64    .09 |  .98  -.6   .97  -.7 | 1.09 |   .37   .35 |  4 2M            |
|  283     504       .56     .56 |   .27    .09 | 1.00  -.2  1.00    .0 | 1.04 |   .23   .23 |  7 2W            |
|  283     504       .56     .56 |   .27    .09 | 1.02    .9  1.02    .8 |  .81 |   .19   .23 |  8 3W            |
|  276     504       .55     .55 |   .22    .09 | 1.00    .1  1.00    .1 |  .97 |   .20   .21 |  2 3N            |
|  275     504       .55     .54 |   .21    .09 | 1.01    .3  1.01    .3 |  .92 |   .19   .21 |  9 4W            |
|  272     504       .54     .54 |   .19    .09 | 1.00    .1  1.00    .0 |  .98 |   .21   .21 |  1 2N            |
|  246     504       .49     .49 A   .00    .09 | 1.01    .3  1.01    .4 |  .91 |   .18   .20 |  5 3M            |
|  211     504       .42     .42 |  -.27    .09 | 1.00    .0  1.00    .1 |  .99 |   .25   .25 |  6 4M            |
|  196     504       .39     .40 |  -.38    .09 |  .98  -.8   .97  -.9 | 1.14 |   .32   .29 |  3 4N            |
|  147     504       .29     .31 |  -.78    .10 | 1.01    .2  1.01    .2 |  .97 |   .43   .43 | 10 uni           |
|-------------------------------+---------------+----------------------+------+-------------+------------------|
|  252.0   504.0     .50     .50 |   .04    .09 | 1.00    .1  1.00    .1 |      |   .26       | Mean (Count: 10) |
|   50.5      .0     .10     .09 |   .39    .00 |  .01    .5   .01    .5 |      |   .08       | S.D. (Population)|
|   53.3      .0     .11     .10 |   .41    .00 |  .01    .5   .01    .5 |      |   .09       | S.D. (Sample)    |
+-------------------------------------------------------------------------------------------------------------+
Model, Populn: RMSE .09  Adj (True) S.D. .38  Separation 4.05  Strata 5.74  Reliability .94
Model, Sample: RMSE .09  Adj (True) S.D. .40  Separation 4.28  Strata 6.05  Reliability .95
Model, Fixed (all same) chi-squared:  163.2  d.f.: 9  significance (probability): .00
Model,  Random (normal) chi-squared:  8.5  d.f.: 8  significance (probability): .38
---------------------------------------------------------------------------------------------------------------
```