Develop an innovative Method for Visual Perception evaluation in a Physical-Based Virtual Environment

Fedaa Abd-Alhamid[a,b], Michael Kent[c], Christopher Bennett[a], John Calautit[a], Yupeng Wu[a,*]

[a]Department of Architecture and Built Environment, Faculty of Engineering, University of Nottingham, University Park, Nottingham, NG7 2RD, United Kingdom.

[b]Department of Architecture, Faculty of Architecture and Design, Al-Zaytoonah University of Jordan, Amman, Jordan.

[c]Berkeley Education Alliance for Research in Singapore, Singapore.

*Corresponding author: laxfa10@nottingham.ac.uk, fedaa.abdalhameed@yahoo.com (Fedaa Abd-Alhamid), Yupeng.Wu@nottingham.ac.uk, Jackwuyp@googlemail.com (Yupeng Wu)

Abstract

With the developments in virtual reality technologies, significant researches have been conducted for human response on indoor luminous environment using head-mounted display to replace those in real environment. However, the limited resolution and luminance values offered by the devices might affect the perceived appearance and high-order impressions in the simulated virtual environment. In this study, a simulated 3-dimensional virtual office was compared against a real one. Both settings presented similar physical and luminous conditions to twenty participants (N=20). The study investigated subjective and objective visual responses and participants' interaction with the virtual environment based on measurements of perceived presence. Subjective assessments included questions on luminous environment appearance (brightness, colour-temperature, distribution) and high-order perceptions (pleasantness, interest, spaciousness, excitement and complexity). Objective assessments measured contrast-sensitivity and colour-discrimination tasks to assess visual performance across the two representation environments. Results showed no significant differences between the two environments based on the studied parameters, indicating a high level of perceptual accuracy of appearance and high-order perceptions. Minor physical symptoms related to the headset

use and high level of perceived presence were found, indicating the proposed methodology's capability to provide realistic immersive environments. Although attributes regarding scene quality: colours, details, and contrast were perceived significantly different to the real environment, objective tasks showed that similar contrast and colour appearance can be produced in the virtual environment with minor impact on fine-details due to limited resolution. Virtual reality maybe a promising alternative representation medium to investigate visual perceptions as the overall appearance of the scene can still be correctly acquired.

1. Introduction

Continual research has led to the development of simulated virtual environments that could be comparable to the experiences felt in real physical environments. This can be achieved by the use of photographs [1-3], 2- or 3-dimensional rendered images displayed on screen [3-5], or reduced scale mock-ups [6-10]. Recently, studies have been using immersive Virtual Reality (VR) as an alternatively media to present the visual setting in indoor lighting studies [11-14]

VR technology can display the visual stimuli in a more comparable field of view (i.e., subjects are immersed in the scene and the visual stimuli can be presented to the same scale of the original environment); an essential parameter when evaluating virtual and real environments [1]. Also, it provides subjects with stereoscopy vision (3-dimensional), which provides the depth perception that cannot be obtained when mesoscopic two-dimensional scenes are assessed [15]. Moreover, it allows the interaction between subjects and the presented scene which can greatly improve the realism for user-experience studies, in which the interaction and immersion are important factors [4, 13, 16, 17].

The mobility of VR technology provides the flexibility of apparatus allocation allowing

the reproducibility of consistent conductions of the experiments [13]. Also, the immersivness

provided by the VR minimises the artificial nature of the experimental setting as subjects

cannot see the setup or the experimenter, which leads to the control of personal factors that can

affect the results as people tend to change their behaviour in the presence of others [18].

VR also allows for more control over different environmental factors [11]. For example,

the variation in daylight conditions caused by changing sky conditions, which is one of the

main challenges in experimental studies using windows [13, 19]. Hence, VR can maintain the

levels of illumination observed within the windows and the surrounding environment, which

can affect visual perception and the assessment of investigated stimuli [19]. Additionally, the

rapid change of visual stimuli in VR environment reduces the time needed to perform the

experiment. This overcomes the limited settings that are usually available for researchers when

real environments are used. Also, the experimenter can replicate the same experimental setup

with wide range of visual stimuli and the same surrounding environmental factors.

The literature suggests that the VR immersive environment could be used as a

representative method to study luminous environments in terms of scale, immersiveness, and

controlled luminous conditions. These factors should first be assessed to validate the use of this

technology when they are compared against real visual environment, however, few studies

have been conducted in this area. In a study by Chamilothori *et al.* [13], stereoscopic

physically-based renderings were evaluated in terms of four lighting impressions (i.e.,

pleasantness, interest, excitement, and complexity); along with satisfaction with amount of

window view and found no significant difference on these parameters when compared to the

real scenes. Presence, appearance attributes and perceptual impressions of lighting using an

immersive 360°-video displayed smartphone VR were compared to a real environment in

recent study [14]. Three reference scenes with average illuminance 800 lux on the work plane

(75 cm above the floor) and three correlated colour temperature (CCT) were used: warm white (3000 K), neutral white (4000 K) and cool white (5500 K). Two lighting attributes (open/close and diffuse/glaring), presence, and overall satisfaction were perceived with no significant difference. However, in this study, it is important to highlight that the VR scene was not calibrated with photometric measurements of the real one. In another study [11], The difference in performance between bright and dark real office was assessed against the difference in bright and dark simulated rendered 3-dimonsional environments. The study results showed no statistically significant differences between the two comparisons. However, the comparisons that were performed did not consider a direct evaluation of the same conditions (e.g., bright real versus bright virtual) between real and virtual conditions, nor did the authors state the exact luminous conditions in both environments (i.e. luminances or illuminances), which highly affects task performance [20-22]. The same limitations were found in a study on lighting preferences for task performance [12]. Validation studies are limited and mainly focus on one specific aspect of the luminous environment. A replication of the results is required to further confirm the applicability of VR when used to evaluate the luminous conditions of any visual environment. Also, the studies were either limited to subjective evaluation or lack the physical calibration of the VR content. Visual quality attributes (i.e. colour, contrast, or detail) are yet to be validated.

The main aim of this study is to evaluate the differences in visual perception (subjectively and objectively) under real and virtual reality conditions by comparing a simulated 3-dimensional virtual office developed using physically based image technique against the real office. Several criteria were used to assess the luminous environment in a more comprehensive approach (i.e., appearance, high-order perceptions and visual quality). This evaluation was accomplished by: (1) Creating controlled luminous conditions in a typical office room under artificial lighting conditions; (2) Developing a replica of the office room (in (1)) based on its

physical and luminous conditions in a 3-dimensional virtual simulated setting using physically

based images; (3) Evaluating the same subjective and objective visual responses under real (1)

and virtual (2) conditions; and (4) Evaluating other parameters related to the use of immersive

virtual reality environments.

2. Experimental Method

2.1. Experimental Setup

To assess the research objectives, a test-room with controlled luminous environment was

prepared. The physical and photometric conditions of this test-room were replicated and

presented within a virtual reality environment. An experiment under controlled artificial

lighting conditions was considered appropriate for this study; as opposed to relying on daylight

from real windows, whereby several uncontrollable parameters (i.e., spectral properties of the

source, light intensity, etc.) would continuously change over time [21, 23, 24]. Also, other

extraneous variables (e.g., noise, temperature, and humidity) could be carefully controlled.

2.2. Test room

An office-like test-room located in Energy Technology Building (University of

Nottingham, UK) was used. The room had internal dimensions of 4.35 m x 2.85 m and a floor

to ceiling height of 3.2 m (Figure 1). The internal walls of the room had reflectance (ρ)

properties: $\rho_{wall} \approx 0.7$, $\rho_{floor} \approx 0.1$, and $\rho_{ceiling} \approx 0.8$. To mask daylight entering the room, the

window was covered with opaque matte-white paper with similar reflectance properties to the

walls $\rho_{paper} \approx 0.7$. The reflectance properties were estimated using the Munsell values [25]. The

room contained furniture to resemble an office working environment. Visual tasks were

mounted onto one of the room's walls at a height of 1.2 m from the floor. A standard office

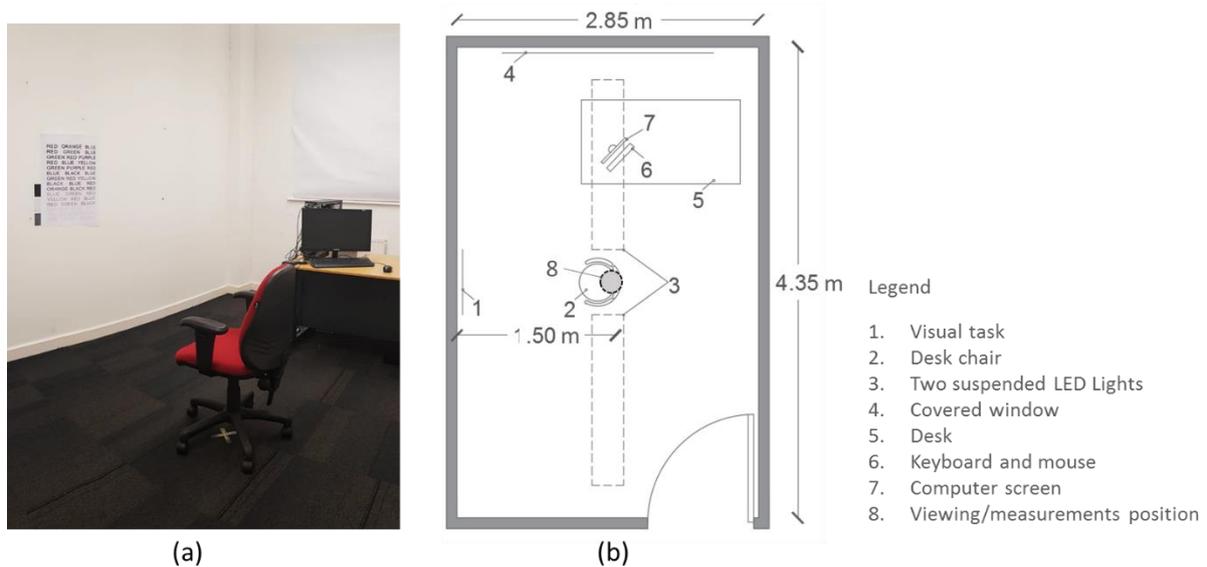desk chair was placed perpendicular to the centre of the task and acted as the viewing position.

Figure 1. (a) Internal view of the experimental room, (b) layout of the experimental room

2.3. Photometric Measurements

The following photometric equipment were used to measure the luminous environment of the test-room: 1) Canon EOS 5D camera equipped with a fish-eye lens (Sigma 4.5 mm f/3.5 EX DG) mounted on a tripod; 2) Hagner S3 photometer; 3) Minolta chroma-meter CL-200; 4) Skye DataHog 2 illuminance data-logger. The camera was mounted on tripod at 1.20 m from the floor corresponding to subject's seated eye level [26] and 1.5 m from the task wall.

Using conventional photographic methods, the camera image pixels can be used to obtain luminance measurements of any visual environment [27]. Such cameras have the ability to capture a large range of luminance values that will be stored within the image pixels correlated with measurements of different points in the captured scene [28]. To measure the luminous environment of the test-room, a high dynamic range image (HDRI) was obtained from seven low dynamic range images (LDRI) with different exposure values by varying the camera shutter speed. The used camera settings are indicated in Table 1 along with the exposure values (EV) which were calculated using aperture size (f) and exposure time or shutter speed (v) based on Equation (1) [29].

$$EV = 3.32 \, log10 \left(\frac{f2}{v} \right)$$ (1)

Table 1. Charge coupled device (CCD) camera settings for each of the seven LDRI

Image	White balance (K)	Sensitivity (ISO)	Exposure time (1/s)	Aperture (f/n)	Exposure Value (EV)
1			1/400		12.98
2			1/125		11.30
3			1/40		9.66
4	4500	100	1/13	4.5	8.04
5			1/4		6.34
6			0.8		4.66
7			2.5		3.02

The lowest sensitivity (ISO) 100 was used to reduce the noise in the HDRI with fixed and correct white balance (i.e., correct colour temperature) to maintain consistent colour space transitions [27]. For the camera white balance, 4500 K was used as the light colour temperature which was measured using the chroma-meter CL-200 (accuracy ± 0.02 %). The seven LDRI were combined into a HDRI using Photosphere software [30]. Photosphere generates a camera response curve based on the LDRI series that shows the relationship between the pixel and its related luminance value, which can be calibrated using a single point luminance measurement within the visual scene. This luminance value was taken by calibrated Hagner S3 photometer (accuracy ± 0.03 %) and was used to calibrate the HDRI.

Since participants will evaluate the luminous conditions of the entire room, the HDRI were captured six times at different viewing directions to cover the visual scenery (Figure 6). The resultant six HDRI images were combined using PTguiPro software – stitching software that supports HDR format [31] – and the resulting HDRI for the entire scene was calibrated with Photosphere.

An average of 50 independent luminance measurements were taken using a Hagner S3 photometer using (0.40 x 0.40 m) grid for divergent targets and (0.05 x 0.05 m) grid on the task area for the convergent targets from the camera position, and 13 points were selected to
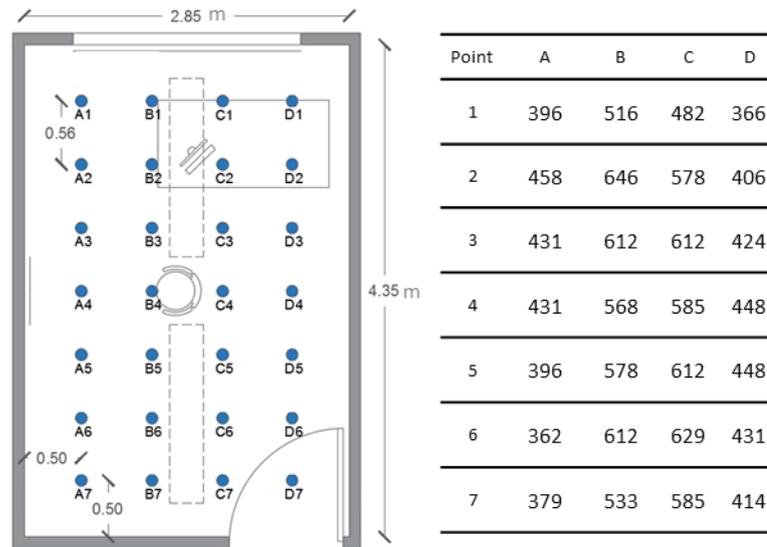
calculate the difference that represents coloured, greyscale, and low and high luminance targets to compare them to corresponding points in the resulting HDRI image of the entire scene, extracted using Photosphere software (Figure 2). This method was informed by Inanici [27]. Percentage error [32] in luminance for each point between spot-point luminance measurements and the resulting HDRI scene was calculated and the resulting average error was 9.5 %, which is within the considered acceptable margin of average errors between 5 and 12 % [1, 27].

Figure 2. Panoramic image illustrating measurements points locations in the test room

The illuminance received at the camera lens and the illuminance received at the lux meter sensor were also compared to further validate the luminance captured by the images [26]. Using the software Evalglare [33], the illuminance received at the camera lens could be obtained. A chroma-meter CL-200 was mounted on a tripod at a height of 1.2 m to measure the vertical illuminance at the same position the camera was mounted facing the direction of the visual task. Measured vertical illuminance and calculated vertical illuminance were equal to: 220 and 219 lux, respectively. This indicates the integrity of the used images. The minimum, maximum and mean average luminances of the entire scene – as calculated from the HDRIs – were equal to 0.015, 23.9 and 7050 cd/m^2, respectively.

1   Twenty eight individual measurements of horizontal illuminance on a regular grid at 0.8

2   m height from the floor level [26] were conducted, the average values were 498 lux (Figure 3).

3   The light correlated colour temperature was 4500 K measured with the chroma-meter. The

4   average illuminance is close to normal office lighting, which is considered to be 'neutral' in

5   terms of both brightness [34] and perceived colour on the Kruithof chart [35].

Point	A	B	C	D
1	396	516	482	366
2	458	646	578	406
3	431	612	612	424
4	431	568	585	448
5	396	578	612	448
6	362	612	629	431
7	379	533	585	414

6

7   **Figure 3.** Horizontal illuminance grid points and their corresponding values measured in lux

8

9   **2.4. Visual Tasks**

10   Two tasks were used in this study. The characters contrast test presented on an achromatic

11   chart (with black and white chart characters) and Stroop test with a chromatic chart (with

12   coloured chart characters) (Figure 4). Both tasks have been used in lighting studies [24, 36-

13   38]. The tasks were mounted at 1.5 m distance from the observer position. The text size was

14   20 mm creating a 0.76° angular size produced by character height, which is within the range

15   needed for fluent reading (between 0.2 to 2°) [39].

Figure 4. Sample of contrast characters (left) and colour recognition tasks (right). For the character contrast test, the rows are denoted by five different contrast groups ranging from Black (first and second rows) to Grey 4 (ninth and tenth rows)

The tasks were used to measure subjects' performance using number of correct responses and completion time [40]. For character contrast, task subjects were asked to read words to measure their cognitive performance, which have no significant meaning in the experiment (i.e., words representing colours) [22]. The words were randomly allocated to counterbalance any learning effect. Subjects were instructed to read the words, attempting to name even those they were uncertain of without any time constraints. The answers were recorded with a Dictaphone to measure accuracy and the rate of time was measured using a stopwatch.

For the chromatic chart, Stroop test with three colours of words: Red, green, blue (RGB) and black was used. The colour represent the three main components of the RGB colour model [20, 21]. The same size and position of the previous task were used and the words were again randomly allocated. The three colours were measured using an Ocean Optics spectrometer USB2000+VIS-NIR-ES (Resolution: 0.1-10 nm varies by configuration) and Halogen Light source HL-2000 (0.25 % Stability of optical output), and had the following positions in the Chromaticity diagram: black (x= 0.306, y= 0.319), red (x= 0.490, y= 0.300), blue (x= 0.210, y= 0.190), green (x= 0.301, y= 0.483), as shown in the chromaticity chart in Figure 5.
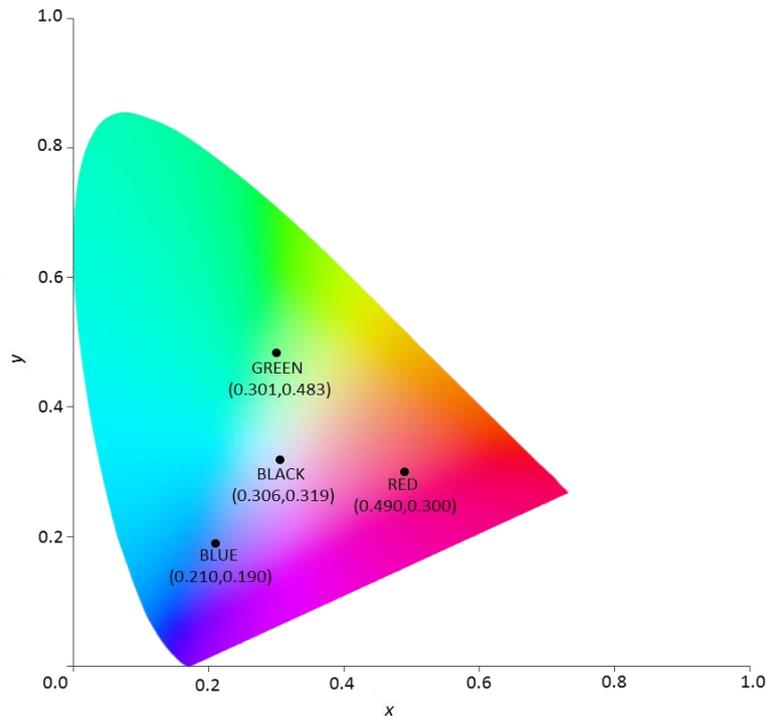
Figure 5. The position of the selected three colours on the chromaticity chart under a standard D65 light source

For this task, subjects were required to identify the colours of each word and their colour discrimination ability was measured by the words' colour correctly named divided by the total number of words' colours that could have been correctly named [40]. Subjects were instructed to name the colours of the words, attempting to name even the ones they were uncertain of, without any time constrains. For a second time, their answers were recorded with a Dictaphone to be analysed for accuracy, and rate of speed was measured using a stopwatch.

The achromatic chart uses a Stroop effect, which increases the difficulty of the colour naming task [41]. This may also influence stress levels. Hence, self-reported levels were recorded using the stress and positive and negative affect schedule (PANAS) [42].

2.5. Virtual Environment

In order to create a virtual environment that replicates the real conditions, physically-based images used on screens [1-3] and physically-based renderings [3, 11, 13] have been used in literature. Literature studies indicate that photographs are more accurate in representing the

1   different luminous conditions than renderings particularly in interactive panoramic view on

2   screen method [3, 4]. Hence, physically-based imaging method was adopted in this study to

3   replicate the real luminosity in 360° 3-dimensional virtual environment.

4   *2.5.1. Physically-Based Images*

5   To create the physically-based virtual luminous environment that replicates the conditions

6   described in reality, four instruments were used: (1) Hagner S3 photometer to measure

7   luminance; (2) DSLR camera equipped with a fish-eye lens and mounted on tripod; (3) HTC-

8   Vive (VR) headset; and (4) Minolta chroma-meter CL-200.

9   A total of six HDRI vertical and horizontal HDRI were created (Figure 4) with each

10   combined from seven LDRI with the same camera settings described in Table 1. The images

11   were taken with fish-eye lens covering 180° in each direction. All images were taken from the

12   same viewing position, aligning the entrance pupil axis to the rotation axis to minimise the

13   differences between the various pictures composing the 360° view [3] (Figure 6).

14   The resultant HDRI images cannot be directly viewed in the virtual reality head-mounted

15   display due to the limited luminance ranges that can be displayed (~216 cd/m^2) [43], which is

16   a common issue with available VR head-mounted displays [13]. To account for this, a tone-

17   mapping process was used, which applies algorithms that compress large ranges of luminance

18   values of the actual scene contained in the HDRI into a lower dynamic range. This allows

19   images to be displayed within conventional devices, while reproducing visual impressions

20   similar to those experienced in real environments [44-46]. Reinhard tone-mapping operator

21   [47] was used for its ability to preserve details and naturalness of the processed images when

22   compared to real scenes [3, 48]. The selected Reinhard tone mapping was applied using

23   Luminance-hdr software [49]. Certain parameters are left for the user to determine such as

24   gamma and key value, which influence the resultant tone-mapped image.
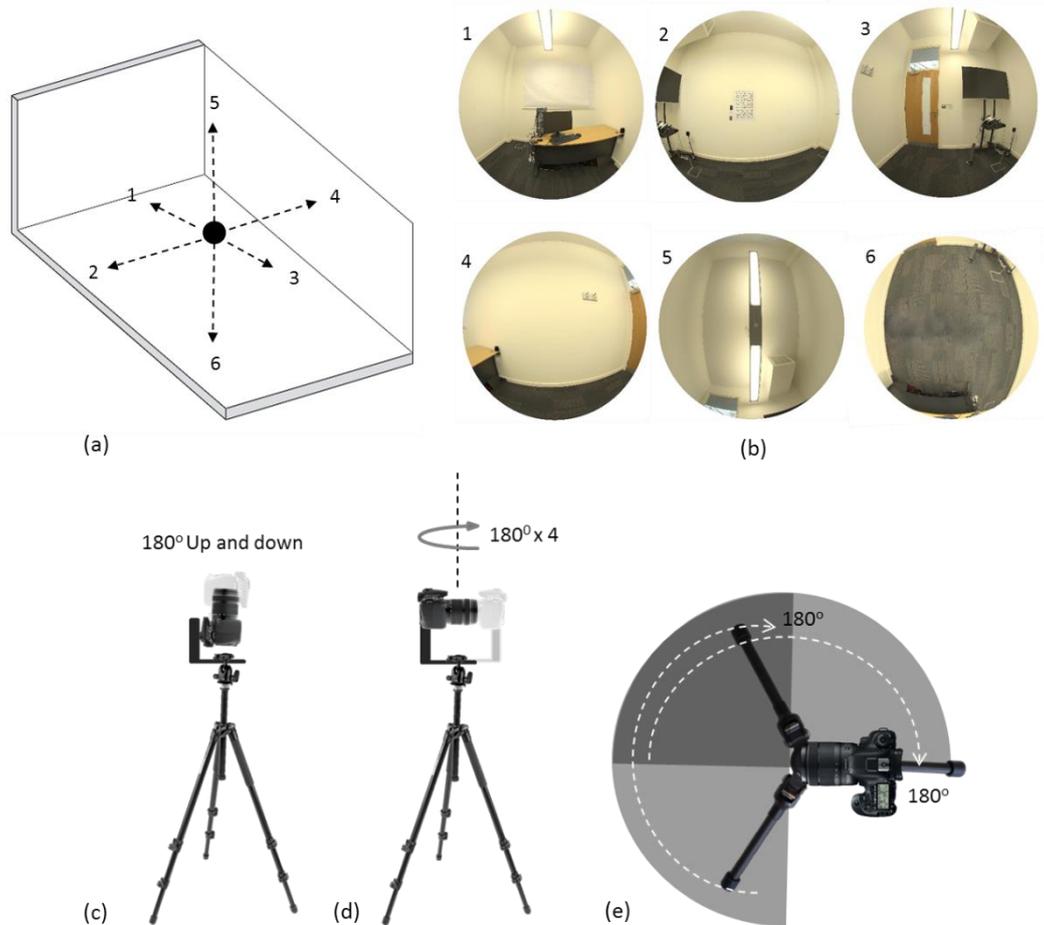
Figure 6. (a) Positions of camera with the six view directions, (b) resulting six fish-eye images, (c) Up and down camera position, (d) entrance pupil axis alignment with rotation axis, and (e) indication of covered view angle for each camera position

A gamma correction of 2.2 was determined for the VR screen using the screen response curve. According to literature, [3, 43], this curve can be obtained by displaying different RGB grey values ranging from (0, 0, 0) to (255, 255, 255) on the screen and measuring the resultant luminance values. For the VR headset used in this study, eight different shades of grey were used, and their corresponding luminance values were measured at the centre of the full field of the lens using Hagner S3 photometer in completely dark room (Figure 7). It should be noted that, 2.2 gamma value is usually used to simulate the human contrast sensitivity curve [44]. However, it is not always the same for all screens and it is more accurate to be measured. Also, the same value was found in human visual perception study for VR display that uses same

1   screen type (OLED) [13]. This curve can be used to calculate the difference between the real

2   luminance values and those resultant from images displayed in the VR headset.
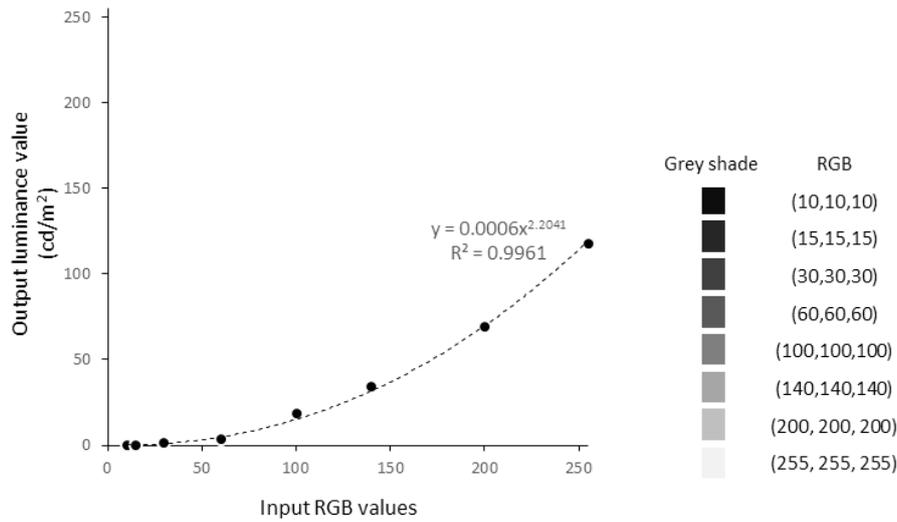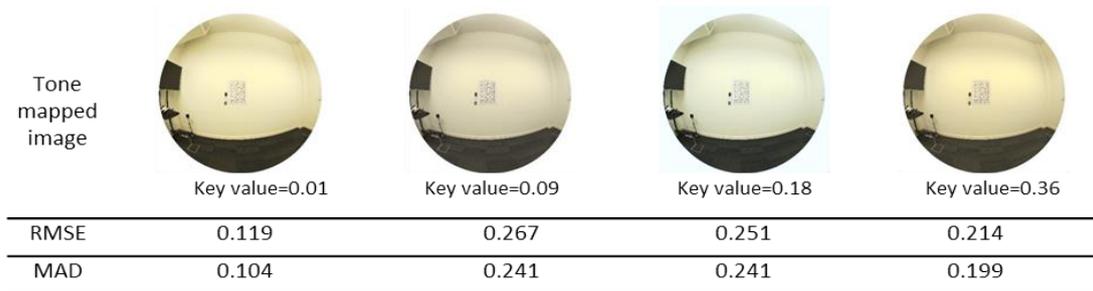
3
4   **Figure 7.** Gamma curve (luminance response curve for the used VR headset as measured at
5   the middle of the lens) using Hagner S3 photometer

6

7   For the key value, a value of 0.01 was applied after a few adjustments were performed as

8   it found to create similar contrast ratios for both the grey and coloured tasks reported in section

9   2.4 - based on Root-Mean-Square-Error (RMSE) [50] and Mean Absolute Deviation (MAD)

10 [51]- between the real contrast and those resultant from tone-mapped image. This aimed to

11 create similar luminous conditions in virtual and real environments. Using a key value that

12 presents similar contrast was selected as it is considered the main factor in image perception

13 preference according to literature [45].

14   To explore the impact of different key values on colourfulness of the resulting scenes, the

15 colourfulness of these images was calculated using a MATLAB code detailed in [52], which

16 quantifies the effect that image processing (i.e., tone-mapping) has on colour perception. The

17 initial room design incorporated some colours besides the colourful task to allow test of the

18 colour representation within the virtual environment, an important aspect of visual-quality

19 representation [53]. The resulting values were compared with the colourfulness of the correct

20 exposed panorama at the correct white balance which produces true colours. A key value of

0.01 was found to create the most accurate contrast and colourfulness (Figure 8) and was used in tone-mapping process. This method was applied for the six HDRI and without any colour adjustment to the resulting images to limit any bias in the image processing procedure.

Tone mapped image	Key value=0.01	Key value=0.09	Key value=0.18	Key value=0.36
RMSE	0.119	0.267	0.251	0.214
MAD	0.104	0.241	0.241	0.199

Figure 8. RMSE and MAD between the real contrast and those resulting by tone-mapped images with different key values

The resultant six tone-mapped images were combined into 360° panorama using PTguiPro software with an additional image for the floor to mask the tripod area in the final image. In order to create the depth perception from two-dimensional images, the previous process was conducted twice (i.e., taking the six HDRI, tone-mapping process, and stitching into 360° panorama) from two viewpoints 65 mm horizontally apart to reproduce the distance between the centres of observer's eyes [54] (Figure 7(a)).
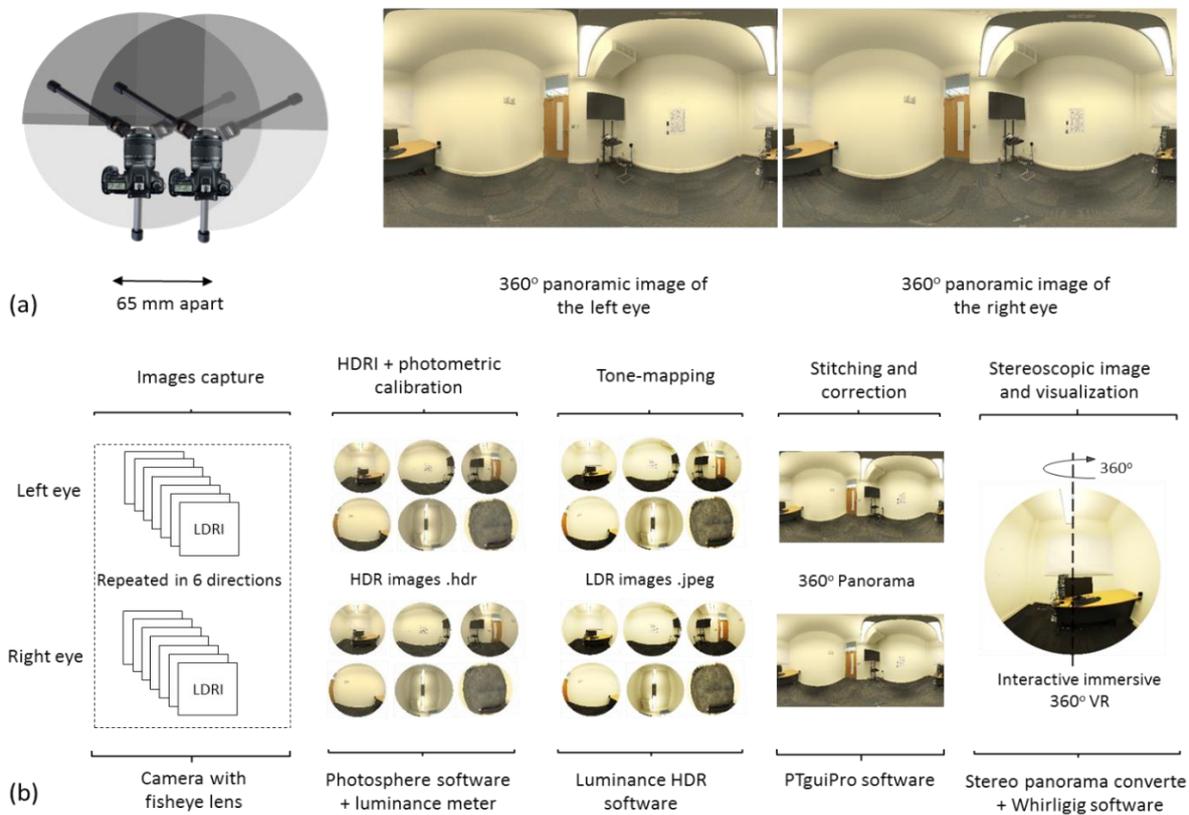
The resulting stereoscopic image (difference between the images for left and right eyes) will create the illusion of depth and the resultant image will be perceived as 3-dimensional [3, 13]. However, this method will create the depth in two directions, and minimising the objects in the other two non-stereoscopic view directions will mask this effect. The two 360° images for each eye were combined into stereoscopic image using Stereo Panorama Converter software [55] and were projected in the VR head-mounted display using Whirligig software [56], which supports stereoscopic image viewing that will be perceived as 3-dimensional (Figure 9).

A HTC Vive head-mounted display [57] with computer of two 2.40 GHz possessors and NVIDIA GeForce GTX 1060 card were used along with Whirligig software, which supports the display of stereoscopic images, to display the immersive 360º images. This will create an interactive viewing mode, whereby the viewed part of the scene will correspond to the subject's head position. The VR HTC Vive has a dual AMOLED 3.6'' diagonal screen with a resolution of 1080 x 1200 pixels per eye (2160 x 1200 pixels when combined) and provides 110º nominal field of view.

Figure 9. (a) Illustration of the stereoscopic principle to the left and the resulting stereoscopic image to the right, (b) Process for generation of the virtual environment

The illuminance received at eye (10 mm from the lens) was measured using chroma-meter (CL-200) in a completely dark environment (i.e., absent of any other source of illumination). This was to verify that the illuminance from the VR display was similar to vertical illuminance measurement taken when in the real luminous environment from the same viewing position: 194 compared to 220 lux when the scene is displayed. No changes were made to the projected

1 images so that any evaluation of the luminous environment will be due to the used method.

2

3 *2.5.2. Visual Tasks Properties in the Virtual Environments*

4 Luminance values of the actual task were measured using Hagner S3 Photometer and the

5 contrast ratios were obtained using Weber's formula (Equation (2)) [58] for both charts

6 displayed in Figure 2, which were calculated using the background luminance of the task (L_b)

7 and target luminance of the visual characters (L_t). Resultant values are presented in Table 3.

8 $$C = (L_t - L_b)/L_b \tag{2}$$

9 The visual properties were affected due to limitations of the current virtual head-mounted

10 display as they cannot display HDR images. Hence, tone-mapping process was applied as

11 described earlier in section 2.5.1 using the key value that produced the contrast and

12 colourfulness that resemble the appearance of the tasks in the real environments. Table 2

13 displays the real and virtual luminances and contrast values of the two visual tasks (Figure 2),

14 and the percentage change in contrast across the two conditions.

15 **Table 2.** Luminances and contrast ratios of the different colours used in the tasks

Colour	Real environment luminance cd/m^2	Tone-mapped images relative luminance	Real environment Contrast	Virtual environment Contrast	Percentage change (%)
Black	9	0.06	-0.88	-0.89	1
Grey 1	11	0.11	-0.85	-0.79	7
Grey 2	26	0.27	-0.65	-0.47	28
Grey 3	48	0.42	-0.37	-0.18	50
Grey 4	65	0.49	-0.13	-0.04	69
Red	23	0.20	-0.69	-0.61	11
Green	19	0.21	-0.74	-0.59	21
Blue	12	0.12	-0.84	-0.76	10
White (background)	75	0.51			

16

17 **2.6. Questionnaires**

18 For the two environments and after completing each visual task, subjects were asked to

19 report their stress using a visual analogue scale (VAS) and to report their Affect levels using

20 the positive and negative affect schedule (PANAS) [59-61]. PANAS short form [42] was used

1   which can be found in the Appendix. Subjects were asked to evaluate five negative and five

2   positive affects using 5-point Likert scales, whereby one indicates or not at all and five indicates

3   extremely. Five-point Likert scales using semantic bipolar words were used for a total of 12

4   questionnaire items related to visual perceptions and were selected based on the literature

5   (Table 3).

6   **Table 3.** Questions used in luminous environment subjective assessment

Parameter		Question	Bipolar descriptions	References
Visual-quality perception	Details	The words on the coloured chart were	Blurry---Sharp	[3, 21, 45, 62]
	Contrast	The contrast of the coloured task was	Low---High	
	Colours	How would you describe the colours in the room	Faded---Strong	
		The overall variety of colours in the room was	Low---High	
Perception of the lighting appearance	Colour temperature	The lighting in this room feels	Cool--Warm	
	Brightness	I perceive the space lighting to be to be	Dark---Bright	[1, 2, 63]
	Distribution	How the lighting distribution in the room appeared	Uneven --- Uniform	[1, 63]
Perception impressions of the room	Pleasantness	I perceive the room as a whole to be	Unpleasant ---Pleasant	[1, 2, 13, 64, 65]
	Interest	I perceive the room to be	Dull ---Interesting	[2, 13, 64, 65]
	Complexity	I perceive the room as a whole to be	Simple ---Complex	[13, 65]
	Excitement	I perceive the room as a whole to be	Tense---Calm	[2, 13, 65]
	Spaciousness	I perceive the room to be	Narrow ---Spacious	[2, 3]

7

8   Two additional questionnaires were used in the experiment for the reported presence [66]

9   and Simulator Sickness Questionnaire (SSQ) [67]. The presence questionnaire was used after

10   the subject had completed each test condition (i.e., real and virtual) on three parameters

11   (realness, spatial presence, and involvement) [11, 13, 66] in comparison to the real environment

12   (Table 4). The SSQ was completed twice before and after immersion in the virtual environment

13   to assess any physical discomfort associated with the immersion in the virtual environment.

14   **Table 4.** Reported presence questionnaire

Parameter	Statement	Bipolar descriptions
Realness [11, 13, 66, 68]	Your experiences in the virtual environment were consistent with your real-world experiences	fully disagree—fully agree
Spatial presence [29, 34, 78, 80]	I felt "being there" in the virtual environment	fully disagree—fully agree
Involvement	The virtual space has become reality for me	fully disagree—fully agree

[29, 34, 78, 80]

The methodology workflow and the several steps used to in this study are summarized in Figure 10 below.
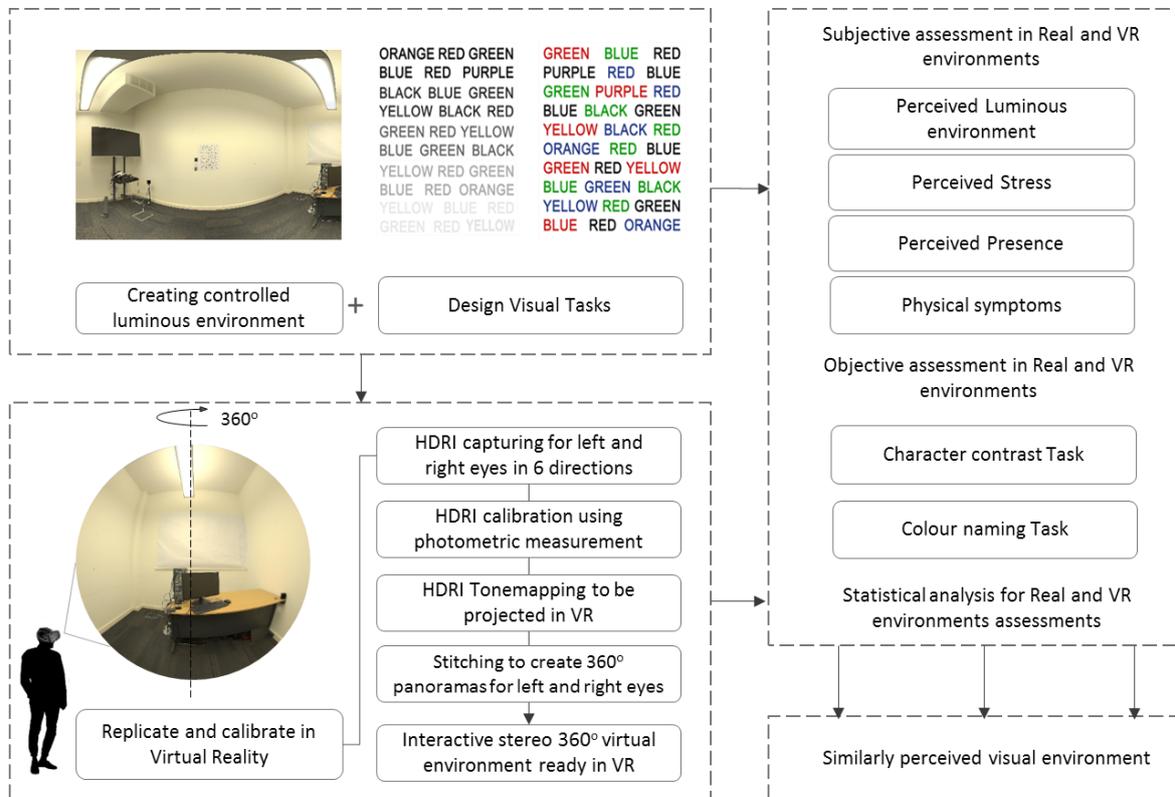

Figure 10. Illustration of the methodology workflow used in the experiment setup design

3. Experiment Design

The study used repeated measure design with the same participant taking part in two conditions to reduce random variability in the collected data [18]. The visual representation environment was the independent variable with two conditional variables: real environment and immersive 3-dimensional environment. The subjects were randomly assigned to test order to counterbalance the effect of presenting order of the stimuli between participants [18].

3.1. Experimental Procedure

The experimental procedure and questionnaires used in the study were assessed and approved by the University of Nottingham Ethics Committee. A total of 20 subjects voluntarily

1 participated in the test. The sample was recruited from Energy Technologies Building and

2 Engineering Faculty from the University of Nottingham using online advertisements.

3   The experimental procedure and duration are shown in Table 5. At the beginning of each

4 session, subjects read the experimental instructions and signed a consent form. Afterwards,

5 subjects completed the post experiment questionnaire on demographic information (age,

6 gender, and academic background), vision problems (e.g., colour blindness) and vision

7 correction, followed by SSQ on physical symptoms.

8   Subjects were undergraduate and postgraduate students, 10 males and 10 females with

9 mean age of 26 years (SD= 6.24) and were from different ethnic backgrounds. None of the

10 participants reported any colour vision problems and eight participants wore corrective glasses

11 during the experiment. Only subjects who did not have epilepsy or suffer from migraines,

12 motion sickness, and dizziness or sleep disorders were permitted to participate. An explanation

13 of the tasks was given to the subjects, using samples with different versions of the tasks used

14 in the experiment, before starting the experiment counterbalancing errors that could occur by

15 unfamiliarity with the test [21]. Each session lasted approximately 40 minutes.

16 **Table 5.** Experiment detailed procedure and duration

Time progress in minutes	Activity	Duration in minutes
0-10	Welcoming and introduction, signing the consent form and completing the pre-test participant questionnaires and SSQ	10
10-12	Demonstration of the experiment to make sure subjects understand the procedures	2
12-17	Participant complete the task in the first environment and the experimenter record responses	5
17-22	Participant complete perception questionnaires and SSQ in case of VR	5
22-27	Participant rest outside the experiment room and experimenter prepare for the second condition	5
27-32	Participant complete the task in the second environment and the experimenter record responses	5
32-37	Participant complete perception questionnaires and SSQ in case of VR	5
37-39	Participant complete presence questionnaire	2
40	The end of the experiment. The participant will be thanked for their time, led to the door and told they are free to leave	1

17

When the subjects started with the virtual reality condition, they were not given any prior information of the real environment to make sure they saw only the VR first. In both conditions, subjects were invited to set on a rotating chair (Figure 1) and were instructed to look around to explore the surrounding environment. Two minutes were allowed before starting the task performance to allow adjusting to the luminous environment and at least 90 % of chromatic adaptation to be reached, as suggested by visual studies [20, 69].

In the characters contrast task, subjects read a total of 45 characters, beginning from the top left corner. In colour discrimination task, the procedure was repeated but instead they were asked to name the words colours of the characters. In both tasks, the number of words was the same but their position and the associated colour or contrast were randomly assigned to counterbalance any learning effects or errors related to the task design [18].

After completing each condition, participants were asked to complete the stress and PANAS questionnaire and a series of questions in random order regarding the perception of visual-quality, lighting, and impressions of the room resultant from the assessed luminous environment (Table 3). After the virtual condition, participants completed another SSQ. Subjects removed the VR headset to complete the questionnaires and were allowed to refer to the VR at any point of the questionnaire to help provide their responses. A similar procedure was followed when evaluating the real environment. Subjects were given a five minute break between conditions and were asked to report their sense of presence in the virtual environment at the end of both conditions (Table 4). Before leaving, participants were asked to sign a consent form indicating the absence of any discomfort that might have been caused by the VR.

4. Results

For visual task performance, no errors were found in both environments for characters contrast task, while colour naming errors in Stroop task were only detected in virtual environment. Hence, visual performance was analysed only by the time spent to complete the tasks, and colour naming errors were analysed separately. The subjective responses to different questionnaires were also analysed for the perception of visual-quality, lighting, and impressions of the room resultant from both luminous environments, SSQ before and after immersion in virtual reality, and sense of presence.

4.1. Task Performance

In Figure 11, box-plots of the results shows the outliers (circles) and the tendencies for the statistical values (e.g., 25th percentile, median, and 75th percentile) [70], indicating lower levels for rate of time for Characters Contrast (CC) and Colour Naming (CN) tasks in virtual reality environment. This suggests a high rate of visual performance under this condition.
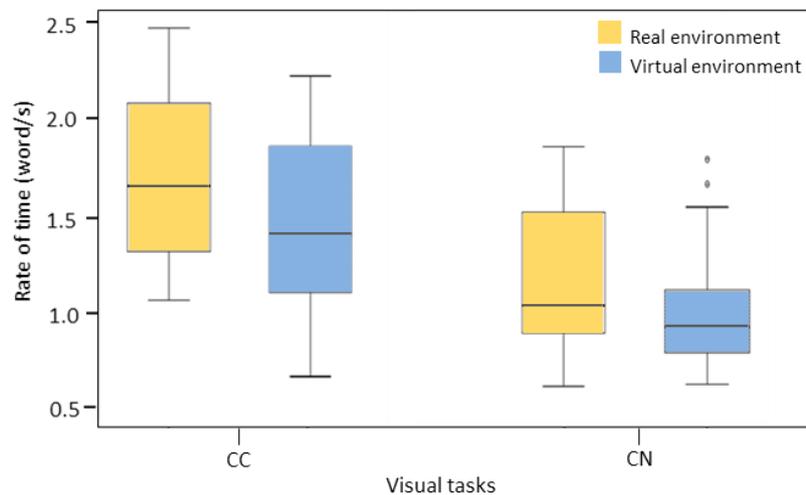
Figure 11. Boxplots presenting the rate of time for the two tasks in real and virtual environments.

To determine whether the mean average values were a reliable indicator of the data distribution, statistical tests (Kolmogorov-Smirnov and Shapiro-Wilk) and graphical plots (Q-

Q) were used. The Levine's test was also applied to determine whether the variances in the data across the independent variables were homogeneous (i.e., approximately equal). When the assumptions of normality and homogeneity of variance were met, the mean average parameter was considered to be a reliable indicator of the data distribution [18]. When these assumptions were not met, non-parametric tests were used as the mean average parameter was not a reliable estimator of the data distribution [18, 71]. These tests were used to determine whether the differences in errors detected were statistically significant across the two conditions. The effect size, r was reported along with statistical significant values to provide a standardised measure of the differences across the two conditions [70]. The interpretation of the effect sizes were derived using thresholds given in the literature: 'small' $(0.20 \leq r < 0.50)$, 'moderate' $(0.5 \leq r < 0.80)$, and 'large' $(r \geq 0.80)$ effect sizes, respectively [72].

The results indicate a highly significant difference between CC task performance in the two environments. Task speed in real environment was different than in virtual environment: ΔMean= 0.24, SD= 0.28), $t(19)$ =3.88, **p<0.01**, **r=0.65** (moderate effect size).

For CN task, the non-parametric Wilcoxon Signed-Rank Test was used to analyse the data as the assumption of normal distribution of the data was violated. Similar to CC, the results for CN indicate a highly significant difference in task performance across the two environments. Task speed in real environment was different than in virtual environment: ΔM_{dn}= 0.19, positive ranks= 4, negative ranks= 15, ties= 0, z_{score}= 2.76, **p<0.01**, **r= 0.45** (small effect size).

These results along with box-and-whisker plots (Figure 8) provides evidence that the difference in task performance was moderately significant in CC and weakly significant in CN as participants needed more time to complete the same task in the virtual condition.

For CN task, errors were found in virtual environment condition (M= 2.63, SD= 2.11). No errors were detected in the real environment, which implies there may have been a difference

1 in colour perception within virtual environment. The initial analysis during pilot testing of CN

2 in the virtual environment revealed that errors were made between the black and blue colours.

3 This might be affected by the low resolution of the VR (i.e., with the characters size at similar

4 RGB for the real task, participates were not able to clearly distinguish between those colours

5 in the VR. To investigate this, additional experiment was conducted to explore whether the

6 tone-mapping or the limited resolution of the VR have affected the colours discernment of the

7 virtual environment represented here by the task characters (i.e., to understand whether a

8 change in colour discernments applies for larger targets or only a result of combined low

9 resolution and small details). Participants were invited to perform simple CN in real and virtual

10 conditions using different widths of strokes with the same character height and identical colours

11 and tone-mapping process of previous test (Figure 12). The same experimental setting and

12 procedure was followed. It was found that participants made no errors in the real conditions to

13 the contrary to the virtual one. In the virtual condition, participants were able to name colours

14 correctly up until the 4th row with the same width size used in the previous experiment (Stroop

15 task described in section 2.4). For the 5th row, errors were detected for all the colours that were

16 presented.
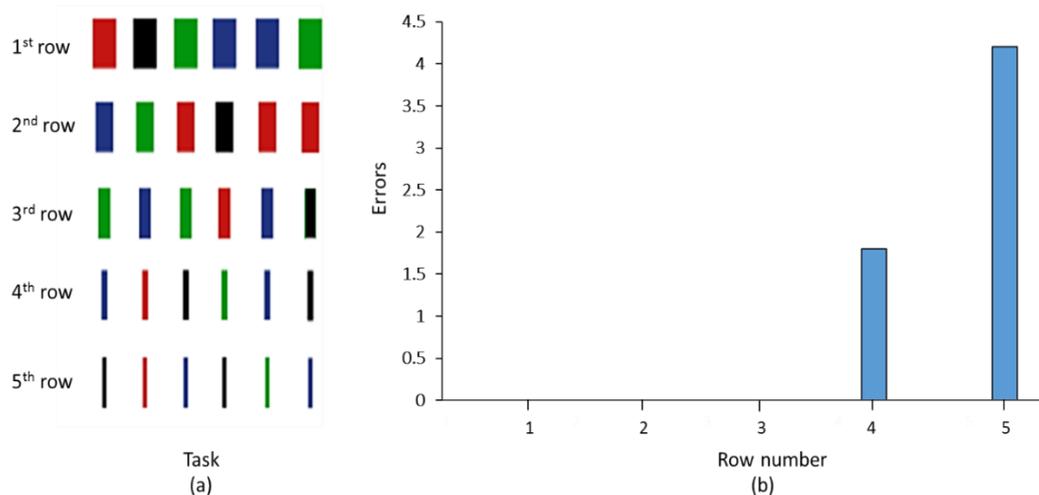
17
18 **Figure 12.** (a) Second colour discrimination task, (b) Mean numbers of errors detected

19

This provided evidence that the difference in CN performance may have been affected by the limited resolution impact on colour discernment of fine details. Both tasks indicate that the visual performance under a given luminous environment in terms of accuracy (freedom of errors) can be replicated in virtual environment. However, the task size should be carefully designed as the resolution was found to be affecting colour's identification for fine details.

4.2. Subjective Perception of Luminous Environments

Since questionnaires related to the luminous environment (i.e., presence and SSQ) were measured using ordinal scales, non-parametric Wilcoxon Signed-Rank test was used to analyse the data. Table 6 presents the results of the Wilcoxon Signed-Rank test and effect sizes.

Three questions had statistically significant differences between the two environments for parameters: 'Details' (moderate effect size); 'Contrast' (moderate effect size); and 'Colourfulness (small effect size). Colour variety was perceived similarly with no significant difference (tied ranks= 11). Although the contrast responses were lower in VR, the objective assessment of the contrast task showed no difference in accuracy in CC task (i.e., no errors were made in both environment). Visual information was correctly extracted, however, it was slower in the virtual environment (i.e., more time needed to complete the task in virtual condition). These questions were related to visual-quality perception of the scene. Other perceptual aspects of the luminous environment including perception of the lighting appearance and perception impressions of the room were perceived similarly in both environments.

Table 6. Results of the Wilcoxon Signed-Rank test for responses to questions.

Parameter	VR(M_{dn})	R(M_{dn})	p-value	Negative	Positive	Ties	z_{score}	Effect size r
Brightness	4	4	0.57 n.s.	3	6	11	-0.59	-0.09
Distribution	4	4	0.58 n.s.	5	4	11	-0.58	-0.09
Colour Temp.	3	3	0.10 n.s.	10	4	6	-1.66	**-0.26***
Pleasantness	3	3	0.27 n.s.	8	4	8	-1.11	-0.18
Interest	2.5	2.5	0.61 n.s.	3	5	12	-0.51	-0.08
Excitement	3	3	0.62 n.s.	5	4	11	-0.49	-0.08
Complexity	2	3	0.10 n.s.	3	7	10	-1.65	**-0.26***
Details	2	4	0.00***	19	1	0	-3.84	**-0.61****
Contrast	2	4	0.00***	15	1	4	-3.54	**-0.56****

Colourfulness	2	3	0.01**	13	3	4	-2.53	**-0.39***
Colour Variety	2.5	3	0.80 n.s.	5	4	11	-0.25	-0.04
Spaciousness	2.5	3	0.11 n.s.	6	2	12	-1.61	**-0.26***

VR= Virtual environment and R= real environment
p-values: ***highly significant; **statistically significant; * weakly significant; n.s. not significant
Effect size: *** Large; ** Moderate; *Small

Regarding lighting appearance, no significant differences were found for brightness and distributions. For many variables, the comparisons showed no large differences across the two conditions as indicated by the higher number of tied ranks in Table 6. However, colour temperature results indicate a tendency towards the negative scale (cool). Questions on lighting high-order perceptions were perceived similarly with 'small' or 'negligible' effect sizes. Pleasantness, interest, and excitement demonstrated a considerably high number of tied ranks (i.e. no differences across the two groups): 8, 12, and 11, respectively. Complexity and spaciousness were perceived slightly different with small effect size) with 10 and 12 tied ranks.

In summary, the differences in lighting appearance and high-order perceptions were not statistically significant and generally the effect sizes were of a 'small' or 'negligible' magnitude.

4.3.Reported Sense of Presence

The attributes of presence: realness, spatial presence, and involvement were measured using ordinal data with 5-point Likert scales however, the mean and standard deviation along with percentages will be reported to allow comparisons to be made with previous studies. Seventy five percent of the 20 participants reported in the positive scale that they felt being there in the virtual environment (Mean= 3.74, SD= 0.99) and that their experience in VR was consistent with real-world (Mean= 3.40, SD= 0.89). Seventy percent reported in the positive scale that virtual environment moderately becomes reality for them (Mean= 3.21, SD= 0.91). These results are similar to those reported in literature [13, 73], which suggests that the used

1 methodology provided acceptable immersive environment in VR compared to the real

2 environment and that participants had sense of presence during the virtual environment.

3 **4.4. Reported Simulator Sickness Symptoms**

4 Wilcoxon Signed-Rank tests were used to analyse the reported simulator sickness

5 symptoms (Table 7). The following symptoms were significantly different before and after

6 using the VR: 'General Discomfort', 'Eye Strain', 'Difficulty Focussing', 'Fullness of the

7 Head', 'Blurred Vision', 'Dizziness Eyes Open', 'Dizziness Eyes Closed', and 'Vertigo' with

8 small effect sizes.

9 **Table 7.** Results of the Wilcoxon Signed-Rank tests for responses to questions on simulator
10 sickness questionnaire

Parameter	VR(M_{dn})	R(M_{dn})	p-value	Negative	Positive	Ties	z_{score}	Effect size, r
General Discomfort	1	1	0.01**	0	7	13	-2.65	**-0.41***
Fatigue	1	1	0.66 n.s.	2	3	15	-0.45	-0.07
Headache	1	1	0.56 n.s.	1	2	16	-0.58	-0.09
Eye Strain	1	1	0.01**	0	7	13	-2.53	**-0.40***
Difficulty Focussing	1	1	0.01**	0	7	13	-2.53	**-0.40***
Salvation Increasing	1	1	0.41 n.s.	1	2	17	-0.82	-0.12
Sweating	1	1	1.00 n.s.	0	0	20	0.00	0.00
Nausea	1	1	1.00 n.s.	1	1	18	0.00	0.00
Difficulty Concentrating	1	1	1.00 n.s.	1	1	18	0.00	0.00
Fullness of the Head	1	1	0.05*	1	6	13	-1.93	**-0.31***
Blurred Vision	1	1	0.10 n.s.	0	3	17	-1.63	**-0.26***
Dizziness Eyes Open	1	1	0.08 n.s.	0	3	17	-1.73	**-0.27***
Dizziness Eyes Closed	1	1	0.10 n.s.	0	3	17	-1.63	**-0.26***
Vertigo	1	1	0.16 n.s.	0	2	18	-1.41	**-0.22***
Stomach Awareness	1	1	1.00 n.s.	0	0	20	0.00	0.00
Burping	1	1	1.00 n.s.	0	0	20	0.00	0.00

VR= Virtual environment and R= real environment
p-values: **statistically significant; * weakly significant; n.s. not significant
Effect size: *** Large; ** Moderate; *Small

11

12 General discomfort may have been reported differently because most participants were

13 using VR for the first time. In a previous studies [13, 74], similar findings were reported for

14 eye strain and dizziness. Nevertheless, it should be noted that all reported symptoms were

15 denoted by small effect sizes and a high number of ties (tied ranks >13) for all symptoms, in

16 other words, when the evaluations across both conditions were the same. Although these

17 symptoms have been associated with virtual reality application, they are generally minor and

short-lived [74]. In fact, before leaving the experiment setting, all participants reported that any

discomfort that was experienced during the VR trial has subdued.

4.5.Perceived Stress after the Task Performance

The Wilcoxon Signed-Rank test was used to analyse the data. Similar levels of stress were

reported in both environments after completing the tasks. The results indicate no significant

difference in PA and minor increase in NA and VAS stress in VR with small magnitude (Table

8). This indicates that the use of VR alone does not impose any change in the post task stress

towards the positive direction (i.e., less stress) [75], and implies that any change in stress levels

in the virtual environment may also be experienced in the real one.

Table 8. Results of the Wilcoxon Signed-Rank tests for responses to stress and PANAS scale

Parameter	VR(M_{dn})	R(M_{dn})	p-value	Negative	Positive	Ties	z_{score}	Effect size
Stress	29.50	17.50	0.04*	7	12	1	-2.095	**-0.33***
PA	16	16.50	0.06 n.s.	7	8	5	-0.057	-0.01
NA	6	5	0.17 n.s.	4	10	6	-1.380	**-0.21***

VR= Virtual environment and R= real environment
p-values: * statistically significant; n.s. not significant
Effect size: *** Large; ** Moderate; *Small

5. Discussion

The results of this study show relatively similar subjective and objective visual responses

between the real and virtual environments and provide evidence that the virtual environment

could be considered as an alternative method when investigating visual responses.

Comparisons of subjective measures across the real and virtual environments (i.e.,

pleasantness, interest, excitement, and complexity) showed no statistically significant

differences with small or negligible effect sizes. These findings support those found in a past

study [13]. The difference in perceived spaciousness of the room was also not statistically

significant and had a small effect size. This parameter was not included in other representation

media (i.e., 2-dimensional screen and 2-dimensional interactive panorama [3]), which suggests that a satisfactory representation of size perception could be produced in the virtual setting.

Participants were able to give correct responses when performing visual tasks presented in the virtual environment, albeit with a slower rate of time. The visual information is acquired from the scene based on its shape, contrast and colour [53]. Hence, even with the lower resolutions as those provided by VR, the overall appearance of the scene can still be correctly perceived. For fine details, the low resolution has more impact and subjects needed more time to perceive the visual information. This was confirmed during the debriefing session.

While commenting on their experience in the virtual environment, participants stated that it felt like looking through "fuzzy glass into the actual room" and that they were "aware of the small squares forming the lens". This may indicate a limitation of the current available VR lenses and its resolution. This was also reflected in their response to questions regarding the quality of the scene (i.e. contrast, colourfulness, and details). The contrast ratios between the two environments were acceptable and had no effect on the accuracy in character contrast task as participants had no errors in reading different contrast ratios. However, the low resolution did impact their performance on colour naming for some colours and the follow-up test showed that the applied method can replicate the colours in the actual environment, as perceived colours of the entire scene. However, this is not true for some colours in the case of fine details.

Another limitation of the discussed method in this study is the difference in luminance values between the real and virtual environment due to the limited luminance that can be produced with similar types of displays ($\sim$216 cd/m^2) [43] and to approximately 118 cd/m^2 with the used software [56] in this study as calculated in the response curve indicated in section 2.5.1. This restricted the use of this technology when considering the evaluation of glare caused by high luminances and enforces the use of tone-mapping process. As HDRI images cannot be

displayed directly in the VR scene, its dynamic range must be compressed to the dynamic range of the display. Hence, the selection of the tone-mapping operator is essential and unbiased objective selection of different parameters such as contrast and colourfulness as proposed in this study should be used in order to replicate the results as both affect the preference of presented scenes [45, 46].

Regarding the three attributes of presence: realness, spatial presence and involvement, respectively, showed fairly acceptable results compared to previous studies [13, 73], which indicate that the used methodology was adequate to create an immersive environment and that participants had sense of presence within the virtual environment.

General discomfort, eye strain and difficulty focussing were slightly higher after using the VR. General discomfort was reported differently as a result of non-familiarity with the technology, as indicated by the subjects at the end of the experiment. In previous studies [13, 74] similar findings were also reported. Although the effect was minor for all symptoms ($0.20 \leq r < 0.50$), future studies should consider these effects (for example, allow participants to familiarise themselves with the device prior to the main test).

Similar levels of post-task stress using VAS and PANAS scale were reported in both environments. However, the self-reported stress and PANAS scale results were not consistent, minor increase in stress was reported in VAS with less magnitude in PANAS. Hence, objective measures are encouraged to be used along with previous scales to assess stress and effect levels such as biofeedback [76-79] and eye-tracking [80-82] technologies.

6. Conclusion

In this study, a novel method for evaluating the use of immersive virtual environment as a replacement of real luminous environment was introduced using a physically-based 360º imaging technique. Objective tasks performance was conducted and subjective responses to

30

perception of scene visual-quality, lighting, and personal impressions of the same test-room were collected along with presence and physical symptoms questionnaires to compare the virtual to the real environment.

The main findings of this study are:

- Participants took relatively longer time to complete the same visual tasks when using VR than when it was presented in the real environment.

- The subjective assessments showed no significant difference for the perception of the lighting and the perception impressions of the room between the two environments.

- A significant difference was found in visual-quality attributes assessment (i.e., details, contrast, and colour vividness). Nevertheless, the analysis of contrast task and colour naming tasks indicated that both colours and contrast replications were acceptable. Hence, the responses may be affected by the limited resolution of the current VR head set.

- The use of VR had minor effects on reported physical symptoms and produced similar stress and positive and negative affect levels, which indicates the adequacy of the proposed methods; however, a more objective assessment could be used in future studies to accurately measure the stress, positive and negative affects and high order light perceptions.

These findings were based on objective and subjective evaluation of twenty participants. Similar number of participants have been used in lighting research 83,84], and the results were statistically relevant and were interrupted based on a more conservative approach (i.e., reporting the effect sizes as an additional measure for significant difference along with *p*-values), yet, some caution should be acknowledged when trying to generalise these research findings.

These results encourage the enhancement of the VR lens technology as the result of this study along with the mobility of the used device has a promised outcome that can replace the continuously changed real day lit environment.

In conclusion, the proposed method looks promising as an alternative to investigate real luminous environments and guarantee reproducibility of the experiment. Further studies on more levels of the investigated parameters (i.e., brightness, colour temperature, etc.) are encouraged to add more validity to the used method. The development of VR headset with higher resolution and HDRI screens could benefit more attributes of lighting to be studies (e.g. glare), however it is yet to be developed.

Acknowledgement:

Declaration of Interest: None.

References

1. Newsham, G.R., et al., *Comparing lighting quality evaluations of real scenes with those from high dynamic range and conventional images.* ACM Transactions on Applied Perception (TAP), 2010. **7**(2): p. 13.
2. Cetegen, D., J. Veitch, and G. Newsham, *View size and office illuminance effects on employee satisfaction.* Proceedings of Balkan light, 2008: p. 242-252.
3. Cauwerts, C., *Influence of presentation modes on visual perceptions of daylit spaces.* Université catholique de Louvain (UCL), Louvain-la-Neuve, Belgium, 2013.
4. Murdoch, M.J., M.G. Stokkermans, and M. Lambooij, *Towards perceptual accuracy in 3D visualizations of illuminated indoor environments.* Journal of Solid State Lighting, 2015. **2**(1): p. 12.
5. Dogrusoy, I.T. and M. Tureyen, *A field study on determination of preferences for windows in office environments.* Building and Environment, 2007. **42**(10): p. 3660-3668.
6. Arsenault, H., M. Hébert, and M.-C. Dubois, *Effects of glazing colour type on perception of daylight quality, arousal, and switch-on patterns of electric light in office rooms.* Building and environment, 2012. **56**: p. 223-231.

7. Ludlow, A., *The functions of windows in buildings.* Lighting Research & Technology, 1976. **8**(2): p. 57-68.

8. Keighley, E.C., *Visual requirements and reduced fenestration in office buildings — A study of window shape.* Building Science, 1973. **8**(4): p. 311-320.

9. Keighley, E.C., *Visual requirements and reduced fenestration in offices — a study of multiple apertures and window area.* Building Science, 1973. **8**(4): p. 321-331.

10. Ne'Eman, E. and R.G. Hopkinson, *Critical minimum acceptable window size: a study of window design and provision of a view.* Lighting Research & Technology, 1970. **2**(1): p. 17-27.

11. Heydarian, A., et al., *Immersive virtual environments versus physical built environments: A benchmarking study for building design and user-built environment explorations.* Automation in Construction, 2015. **54**: p. 116-126.

12. Heydarian, A., et al., *Towards understanding end-user lighting preferences in office spaces by using immersive virtual environments.* Computing in Civil Engineering, 2015. **20152015**: p. 475-482.

13. Chamilothori, K., J. Wienold, and M. Andersen, *Adequacy of immersive virtual reality for the perception of daylit spaces: Comparison of real and virtual environments.* LEUKOS, 2018: p. 1-24.

14. Chen, Y., Z. Cui, and L. Hao, *Virtual reality in lighting research: Comparing physical and virtual lighting environments.* Lighting Research & Technology, 2019: p. 1477153518825387.

15. Lambooij, M., et al., *Evaluation of stereoscopic images: beyond 2D quality.* IEEE Transactions on broadcasting, 2011. **57**(2): p. 432-444.

16. Bishop, I.D. and B. Rohrmann, *Subjective responses to simulated and real environments: a comparison.* Landscape and urban planning, 2003. **65**(4): p. 261-277.

17. Kort, Y.A.d., et al., *Virtual laboratories: Comparability of real and virtual environments for environmental psychology.* Presence: Teleoperators & Virtual Environments, 2003. **12**(4): p. 360-373.

18. Field, A. and G. Hole, *How to design and report experiments.* 2002: Sage Ltd ISBN: 9780761973836.

19. Guidolin, E., *Impact of window amount and size on user perception, daylighting and energy demand in an office space.* University of Padua, Thesis, 2014.

20. Fotios, S. and C. Cheal, *Lighting for subsidiary streets: investigation of lamps of different SPD. Part 1—Visual Performance.* Lighting Research & Technology, 2007. **39**(3): p. 215-232.

21. Liang, R., et al., *Development of experimental methods for quantifying the human response to chromatic glazing.* Building and Environment, 2019. **147**: p. 199-210.

22. Bailey, I., R. Clear, and S. Berman, *Size as a determinant of reading speed.* Journal of the Illuminating Engineering Society, 1993. **22**(2): p. 102-117.

23. Tuaycharoen, N. and P. Tregenza, *Discomfort glare from interesting images.* Lighting Research & Technology, 2005. **37**(4): p. 329-338.

24. Kent, M., et al., *Temporal variables and personal factors in glare sensation.* Lighting Research & Technology, 2016. **48**(6): p. 689-710.

25. Tregenza, P. and D. Loe, *The design of lighting, published by E&FN SPON.* 1998, London.

26. Wienold, J. and J. Christoffersen, *Evaluation methods and development of a new glare prediction model for daylight environments with the use of CCD cameras.* Energy and Buildings, 2006. **38**(7): p. 743-757.

27. Inanici, M., *Evaluation of high dynamic range photography as a luminance data acquisition system.* Lighting Research & Technology, 2006. **38**(2): p. 123-134.

28. Jacobs, A., *High dynamic range imaging and its application in building research.* Advances in building energy research, 2007. **1**(1): p. 177-202.

29. Coutelier, B. and D. Dumortier, *Luminance calibration of the Nikon Coolpix 990 digital camera. Application to glare evaluation.* Proceedings of Energy Efficiency and Healthy Buildings in Sustainable Cities, Lyon, October 23e26 October, 2002.

30. Ward, G., *Photosphere, Anyhere Software.* 2005.

31. New House Internet Services B.V., *Ptguipro software.* Rotterdam, The Netherlands., 1996.

32. Guang, W., M. Baraldo, and M. Furlanut, *Calculating percentage prediction error: A user's note.* Pharmacological Research, 1995. **32**(4): p. 241-248.

33. Wienold, J., *Daylight Glare in Offices.* . Ph.D. Fraunhofer ISE., 2009.

34. Steidle, A. and L. Werth, *Freedom from constraints: Darkness and dim illumination promote creativity.* Journal of Environmental Psychology, 2013. **35**: p. 67-80.

35. Fotios, S., *A Revised Kruithof Graph Based on Empirical Data.* LEUKOS, 2017. **13**(1): p. 3-17.

36. Jindra, L.F. and V. Zemon, *Contrast sensitivity testing: A more complete assessment of vision.* Journal of Cataract & Refractive Surgery, 1989. **15**(2): p. 141-148.

37. de Vries, A., et al., *Lighting up the office: The effect of wall luminance on room appraisal, office workers' performance, and subjective alertness.* Building and Environment, 2018. **142**: p. 534-543.

38. Schlangen, L., J. Verhaegh, and A. Denissen, *WORKPLACE ILLUMINATION EFFECTS ON ACUITY, COGNITIVE PERFORMANCE AND WELL-BEING IN OLDER AND YOUNG PEOPLE.* Proceedings of 28th CIE Session, 2015. **87**.

39. Legge, G.E. and C.A. Bigelow, *Does print size matter for reading? A review of findings from vision science and typography.* Journal of vision, 2011. **11**(5): p. 8-8.

40. Boyce, P.R., *On measuring task performance.* Coloration Technology, 2011. **127**(2): p. 101-113.

41. Golden, C.J. and S.M. Freshwater, *Stroop color and word test.* 1978.

42. Thompson, E.R., *Development and Validation of an Internationally Reliable Short-Form of the Positive and Negative Affect Schedule (PANAS).* Journal of Cross-Cultural Psychology, 2007. **38**(2): p. 227-242.

43. Ito, H., M. Ogawa, and S. Sunaga, *Evaluation of an organic light-emitting diode display for precise visual stimulation.* Journal of Vision, 2013. **13**(7): p. 6-6.

44. Reinhard, E., et al., *High dynamic range imaging: acquisition, display, and image-based lighting*. Morgan Kaufmann Series in Computer Graphics. 2010.

45. Kuang, J., G.M. Johnson, and M.D. Fairchild. *Image preference scaling for HDR image rendering*. in *Color and Imaging Conference*. 2005. Society for Imaging Science and Technology.

46. Kuang, J., et al., *Evaluating HDR rendering algorithms.* ACM Transactions on Applied Perception (TAP), 2007. **4**(2): p. 9.

47. Reinhard, E., et al., *Photographic tone reproduction for digital images.* ACM transactions on graphics (TOG), 2002. **21**(3): p. 267-276.

48. Yoshida, A., et al. *Perceptual evaluation of tone mapping operators with real-world scenes*. in *Human Vision and Electronic Imaging X*. 2005. International Society for Optics and Photonics.

49. Guiuseppe, R., C. Franco, and A. Davide, *Luminance HDR 2.5.1 software.* 2006-2010.

50. Chai, T. and R.R. Draxler, *Root mean square error (RMSE) or mean absolute error (MAE)?– Arguments against avoiding RMSE in the literature.* Geoscientific model development, 2014. **7**(3): p. 1247-1250.

51. Willmott, C.J. and K. Matsuura, *Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance.* Climate research, 2005. **30**(1): p. 79-82.

52. Hasler, D. and S.E. Suesstrunk. *Measuring colorfulness in natural images*. in *Human vision and electronic imaging VIII*. 2003. International Society for Optics and Photonics.

53. Castelhano, M.S. and J.M. Henderson, *The influence of color on the perception of scene gist.* Journal of Experimental Psychology: Human perception and performance, 2008. **34**(3): p. 660.

54. Dashwood, T., *A beginner's guide to shooting stereoscopic 3D.* Dashwood Cinema Solutions, 2011.

55. Michel, G., *Stereo Panorama Converter software* 2018.

56. Day, P., *Whirligig VR media player.* 2015.

57. Vive.com, *VIVE™ | VIVE Virtual Reality System. [online] Available at: https://www.vive.com/us/product/vive-virtual-reality-system/ [Accessed 29 Oct. 2018].* 2016.

58. Ortiz-Jaramillo, B., et al., *Content-aware contrast ratio measure for images.* Signal Processing: Image Communication, 2018. **62**: p. 51-63.

59.   Boubekri, M., R.B. Hull, and L.L. Boyer, *Impact of window size and sunlight penetration on office workers' mood and satisfaction: A novel way of assessing sunlight.* Environment and Behavior, 1991. **23**(4): p. 474-493.

60.   Wang, N. and M. Boubekri, *Design recommendations based on cognitive, mood and preference assessments in a sunlit workspace.* Lighting Research & Technology, 2011. **43**(1): p. 55-72.

61.   Wang, N. and M. Boubekri, *Investigation of declared seating preference and measured cognitive performance in a sunlit room.* Journal of Environmental Psychology, 2010. **30**(2): p. 226-238.

62.   Odabaşioğlu, S. and N. Olguntürk, *Effects of coloured lighting on the perception of interior spaces.* Perceptual and motor skills, 2015. **120**(1): p. 183-201.

63.   Seetzen, H., et al. *25.3: Observations of luminance, contrast and amplitude resolution of displays.* in *SID Symposium Digest of Technical Papers.* 2006. Wiley Online Library.

64.   Vogel, I., *Atmospheric Metrics: Development of a tool to quantify experienced atmosphere in Westerink et al.* Probing Experience, 2008: p. 25-41.

65.   Flynn, J.E., et al., *Interim study of procedures for investigating the effect of light on impression and behavior.* Journal of the Illuminating Engineering Society, 1973. **3**(1): p. 87-94.

66.   Schubert, T., F. Friedmann, and H. Regenbrecht, *The experience of presence: Factor analytic insights.* Presence: Teleoperators & Virtual Environments, 2001. **10**(3): p. 266-281.

67.   Kennedy, R.S., et al., *Simulator sickness questionnaire: An enhanced method for quantifying simulator sickness.* The international journal of aviation psychology, 1993. **3**(3): p. 203-220.

68.   Witmer, B.G. and M.J. Singer, *Measuring presence in virtual environments: A presence questionnaire.* Presence, 1998. **7**(3): p. 225-240.

69.   Fairchild, M.D. and L. Reniff, *Time course of chromatic adaptation for color-appearance judgments.* Journal of the Optical Society of America A, 1995. **12**(5): p. 824-833.

70.   Field, A., *Discovering statistics using IBM SPSS statistics.* 2013: sage.

71.   Kitchen, C.M.R., *Nonparametric vs parametric tests of location in biomedical research.* American journal of ophthalmology, 2009. **147**(4): p. 571-572.

72.   Ferguson, C.J., *An effect size primer: A guide for clinicians and researchers.* Professional Psychology: Research and Practice, 2009. **40**(5): p. 532.

73.   Kuliga, S.F., et al., *Virtual reality as an empirical research tool — Exploring user experience in a real building and a corresponding virtual model.* Computers, Environment and Urban Systems, 2015. **54**: p. 363-375.

74.   Solimini, A.G., et al., *A survey of visually induced symptoms and associated factors in spectators of three dimensional stereoscopic movies.* BMC public health, 2012. **12**(1): p. 779.

75.   Valtchanov, D., K.R. Barton, and C. Ellard, *Restorative effects of virtual nature settings.* Cyberpsychology, Behavior, and Social Networking, 2010. **13**(5): p. 503-512.

76.   Ulrich, R.S., et al., *Stress recovery during exposure to natural and urban environments.* Journal of Environmental Psychology, 1991. **11**(3): p. 201-230.

77.   Chang, C.-Y. and P.-K. Chen, *Human response to window views and indoor plants in the workplace.* HortScience, 2005. **40**(5): p. 1354-1359.

78.   Kim, J., et al., *The effects of indoor plants and artificial windows in an underground environment.* Building and Environment, 2018. **138**: p. 53-62.

79.   CHANG, C., *Effects of Landscape on Psychological and Physical Responses.* Journal of Therapeutic Horticulture, 1998. **9**: p. 73-76.

80.   Bergstrom, J.R. and A. Schall, *Eye tracking in user experience design.* 2014: Elsevier.

81.   Choi, J.-H., *Investigation of human eye pupil sizes as a measure of visual sensation in the workplace environment with a high lighting colour temperature.* Indoor and Built Environment, 2017. **26**(4): p. 488-501.

82.   Ho, M.-C., et al., *Numerical analysis on color preference and visual comfort from eye tracking technique.* Mathematical Problems in Engineering, 2015. Article ID 861610 2015: p. 4.

83.   Kent, M., Fotios, S. and Altomonte, S. (2018). Order effects when using Hopkinson's multiple criterion scale of discomfort due to glare. Building and Environment, 136, pp.54-61.

84.   Altomonte, S., Kent, M., Tregenza, P. and Wilson, R. (2016). Visual task difficulty and temporal influences in glare response. *Building and Environment*, 95, pp.209-226.

1 **Appendix**

PANAS and self-reported stress

Please draw a vertical line at the appropriate point along the Visual Analogue Scale (VAS).

'Indicate how stressed you feel

0	50	100

Not at all   very stressed
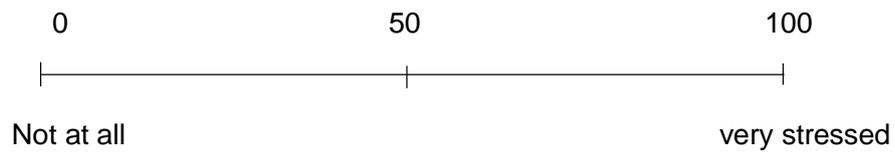
On scale from 1 to 5 please circle how you feel at this moment using the reference scale

1	2	3	4	5
Not at all	a little	moderately	quite a bit	extremely

	1	2	3	4	5
Active	1	2	3	4	5
Attentive	1	2	3	4	5
Inspired	1	2	3	4	5
Alert	1	2	3	4	5
Hostile	1	2	3	4	5
Afraid	1	2	3	4	5
Nervous	1	2	3	4	5
Determined	1	2	3	4	5
Upset	1	2	3	4	5
Ashamed	1	2	3	4	5

2