## Title

The effect of thermochromic windows on visual performance and sustained attention

**Permalink**
https://escholarship.org/uc/item/9kt889fn

## Authors

Liang, Runqi
Kent, Michael
Wilson, Robin
et al.

**Publication Date**
2021-04-01

**DOI**
10.1016/j.enbuild.2021.110778

# The Effect of Thermochromic Windows on Visual Performance and Sustained Attention

Runqi Liang [a,b]*, Michael Kent [b,c], Robin Wilson [b], Yupeng Wu [b]*,

a. College of Architecture and Urban Planning, Tongji University, Shanghai, China
b. Department of Architecture and Built Environment, Faculty of Engineering, University of Nottingham, University Park, Nottingham, NG7 2RD, UK
c. Berkeley Education Alliance for Research in Singapore, Singapore
*Corresponding author: Yupeng.Wu@nottingham.ac.uk Y. Wu Runqi_Liang@tongji.edu.cn (R. Liang)

## Abstract

Thermochromic windows have been widely studied as a technology that can potentially offer increases in energy conservation and provide a desirable luminous environment inside buildings. However, there has been little attention placed on how the tinted states of thermochromic glazing influence occupant behaviour and visual perception. An experiment under controlled conditions was designed to test the influence of different thermochromic tint states on human response. By using a controllable artificial window, five typical luminous conditions were set up, including clear (no tint) and two different levels of blue and bronze tint states, respectively, which produced different room colour temperatures. Thirty-one subjects were recruited who completed three visual tasks, including a visual acuity and a colour naming tasks using the coloured Landolt ring chart and a sustained attention test using the d2 test. Subjective assessments were also collected using questionnaires. Statistical analyses showed the across the thermochromic window conditions, no significant differences in performance were found for the visual acuity and d2 tests. However, there was a significant effect for the colour naming task from the Landolt ring test. Under blue tint conditions, subjects reported higher alertness and produced fewer errors. More natural and acceptable lighting conditions were found under the bronze-tinted conditions. Therefore, when developing innovative thermochromic windows applied in buildings, it is also important to cater for the visual requirements of the occupants in the space, not only energy efficiency goals.

https://doi.org/10.1016/j.enbuild.2021.110778
https://escholarship.org/uc/item/9kt889fn

**Keywords**: Thermochromic windows; Colour temperatures; Subjective assessment; Sustained attention; Visual performance

## 1. Introduction

Although windows allow daylight to enter a building, they also represent an area of "thermal weakness" in its façade. In some cases, approximately 60% of total energy loss can be attributed to heat flows through windows. The energy savings offered by advanced windows can significantly reduce the scale of these losses [1, 2]. Thermochromic (TC) windows are receiving increased attention to better understand how they may be used to provide dynamic spectral solar control. Unlike conventional windows, TC windows are able to change their optical properties in response to ambient temperature, adjusting the spectral solar radiation transmitted into an indoor space [3]. The dynamic change of solar gain and daylight offered by TC windows has the potential to meet indoor comfort requirements in different climates [4, 5]. For example, $VO_2$ based TC window (tinted state) blocks the near-infrared (NIR) radiation entering the room in a hot season but allowing the visible light in for daylight, therefore reducing cooling and artificial lighting energy demand. On the other hand, in a cooler season, both visible light and NIR heat (clear state) can penetrate into the room for daylight and passive heating [6].

A number of developed TC materials are appropriate for use in buildings [2, 5, 7, 8]. According to their spectral properties, TC windows are able to be classified into three types: 1) the ones can adjust solar transmittance within NIR wavelength (780-2500nm), but almost no change of visible transmittance (380-780nm), e.g. $VO_2$ based TC windows [2-4]; 2) The ones offer control of visible transmittance, particularly, e.g. Ionic liquid (IL) based composite films [9, 10]; 3) TC windows enable the adjustment of both visible and NIR radiation, named Suntuitive glass system, extruded polyvinyl butyral (PVB) interlayer realized TC window, transferring from clear to dark blue during the transition [11]. Significant researches have been carried out to developed $VO_2$ and IL-based prototype smart windows under lab scale, many studies prove their potential on energy conservation and daylighting regulation[12, 13]. It is also reported that various TC materials exhibit a degree of selective absorption resulting in transmitted light that is tinted, typically yellow/brown or blue/green. For example, pure $VO_2$ TC materials tend to be yellow/brownish at the tinted state, but by adding different metal dopant such as tungsten and gold for temperature or visible modular control, their colour change to green/blueish [14, 15], IL-based films appears green/ blue or bronze at tinted state, depending on the film temperature [9, 10]. As both the amount of daylight and its spectral (colour) properties can be affected, it is necessary for designers to carefully consider which type of TC glazing is suitable to illuminate the interiors of buildings and how its properties complement both climate and building type [16]. These factors also affect how building users experience views out of buildings onto their surroundings [17, 18].

Although illuminance criteria are considered to be one of the main design factors used to provide illumination in office environments, correlated colour temperature (CCT) also influences performance of visual tasks and subjective ratings of naturalness, arousal, and pleasantness [19-21]. Recommended illuminance levels in the Chartered Institute of Building Services Engineers (CIBSE) Guide A for different spaces used for work and study range between 300-500 lux[22]. The commonly used classifications of CCT for artificial light sources are: warm (<3300 K), intermediate (3300 K<CCT< 5300 K), and cold (CCT5300K) [22]. There is a significant body of research exploring human visual performance in environments illuminated with different CCT light sources. This has been used to provide reliable guidance for the design of artificial lighting that in addition to supporting task performance, also seeks to improve occupant physiological and psychological wellbeing [23-25]. One important finding from these studies shows that higher CCT (i.e., light sources with a colder appearance) tend to be perceived as unpleasant when used to provide high levels of illuminance [26-28]. Conversely, warmer colour temperature lighting has positive effects on mood [29] and sources with a CCT around 6500 K provide the most visually comfortable environment [30]. Additionally, CCT is able to affect the thermal comfort  it is found that higher CCT improved thermal comfort at 24 and 26$^{o}$C [31], HVAC energy consumption can be reduced by adjusting CCT of artificial lighting [32]. Despite these links to CCT, standards rarely provide guidance on what CCT and illuminance levels are suitable in indoor daylit spaces.

Studies have been undertaken to evaluate the effects on sustained attention in an office environment illuminated using light sources with different CCT [29, 33-38]. Early studies, which tested visual performance, cognitive ability, and mood, found that there was no significant difference in performance at illumination levels of 500 lux using sources with CCT ranging from 3000 K to 5000 K [33, 34]. A study carried out by Rautkyla [35] with undergraduate students indicated that different CCTs (4000 K and 17000 K) and prior exposure to daylight affected alertness. Additionally, seasonal differences were discovered during the spring. CCT was found to have less effect on alertness for tests performed in the autumn, but the 17000 K source was found to induce more alertness than the source with a CCT of 4000 K.

Sleegers *et al.* [36] studied concentration levels using student subjects who completed tasks under three different CCT conditions: 2900, 6500 and 12000 K. The source that approximated daylight with a CCT of around 6500 K resulted in higher levels of concentration compared to tests conducted under warm (2900 K) and cold (12000 K) light sources. Shamsul *et al.* [37] explored how visual task performance and subjective comfort were affected by sources with CCTs of 3000, 4000 and 6500 K. They concluded that colder light sources can result in increased alertness and productivity when performing tasks such as typing. Huang *et al.* [38] presented subjects with the Chu attention test under CCT conditions of, respectively, 2700, 4300, and 6500 K, and found that sources at 4300 K resulted in better sustained attention. Smolders and Kort investigated alertness and arousal levels under two CCT conditions (2700 and 6000 K) using tests performed in the morning and afternoon, and found that warmer lighting induce more positive affect [29]. Yu [39] found that the use of light sources with higher CCT can improve reading concentration. Evaluations made by subjects and physiological measures showed that in an environment illuminated with a higher CCT source (6000 K) subjects showed higher levels of vitality while under a warmer lighting condition (2700 K) higher ratings of mood were reported.

**Table 1.** Studies about the effect of CCTs on human sustained attention

| CCTs (K) | Luminance conditions | Variables | Conclusions | Reference |
|---|---|---|---|---|
| **3000~5000** | 500lux | visual performance; cognitive ability; mood | No significant difference | [33, 34] |
| **4000;17000** | 1000lux | Alertness | Higher CCT increase alertness; CCT has less effect on alertness during autumn | [35] |
| **2900; 6500; 12000** | 300lux;<br>650lux;<br>1000lux | concentration levels | 6500K induces higher concentration | [36] |
| **3000; 4000; 6500** | N/A | visual task performance; subjective comfort | Higher CCT increases alertness and productivity | [37] |
| **2700; 4300; 6500** | 500lux | sustained attention | 4300K increases attention | [38] |
| **2700; 6000** | 150lux;<br>300lux | alertness and arousal levels | Lower CCT induces more positive effect | [29] |
| **2700; 6000** | 500lux | reading concentration | Higher CCT increases concentration | [39] |

We believe there is a need to understand how changes to the spectral content of daylight brought about by transmission through TC glazing materials in their tinted state influences occupant behaviour and visual perception. The literature suggests that there may be a link between sustained attention and the CCT of the light sources to which subjects are exposed. This paper reports on a controlled laboratory experiment designed to replicate the effects of two types of TC windows. The window types selected were based on $VO_2$ TC materials, which can tint to give a bronze hue, and ionic liquid-based TC materials, which can tint to produce a bluish hue. A third condition representing traditional clear glazing was used as a control. Subjects were presented with visual tasks and asked to provide subjective assessment using questionnaires. Performance under these conditions was compared to determine whether visual performance and sustained attention were affected by the luminous environment. This work aims to provide an understanding of the TC windows' colour on occupants' perception, therefore, providing recommendation for future TC window development.

## 2. Experimental Method

### *2.1 Experimental set up*

This study made use of a test room with dimensions of 1.5 m x 1.2 m x 2.1 m constructed in the laboratory of Energy Technologies Building at the University of Nottingham (UK) within which subjects were asked to complete visual tasks. The size of the experimental chamber was approximately 1/3 that of a typical 11m$^2$ minimum working space, meeting the requirement for models designed for use in subject studies where subjects are expected to perceive and provide an assessment of a daylit environment [40, 41]. The temperature inside the chamber was monitored and maintained to be approximately 25 °C, and humidity within 45%-55%, meeting the requirement of moderate thermal comfort shown in CIBSE guide A[22]. The experimental chamber was illuminated using an artificial window comprising an array of six LED lamps behind a diffuser, generating light with a CCT close to 6500 K [42]. Although LED lamps cannot provide that full-spectrum radiation than real sunlight (380-780nm), according to the sensitivity of human eyes (under photopic condition, 400-700nm, peak at 550nm)[26], the measured LED spectrum (400-750nm) is sufficient for human eyes [42], which is suitable for the purpose of the proposed studies.

Two types of TC materials were selected to be investigated in this experiment, 1) $VO_2$ nanoparticle TC film [43], which has a relatively large modulation within NIR wavelength, is one of the most promising TC film, tinting to be yellow/brown before and after the transition;

2) Composite film of ionic-liquid based TC containing $[bmin]_2NiCl_4$ used for VIR control, appears clear at the temperature below 25°C, and gradually tinting to blue with temperature increasing to 75 °C [11]. This work aims to mainly evaluate the occupants' perception for the selected thermochromic windows under tinted states to understand their potential for future building application. To better control the indoor luminous environment and reduce the possible effects of temperature fluctuation on TC film states, corresponding coloured films replicating two tinted levels of each selected TC films with similar spectral properties were applied for the proposed test. Spectrum transmittance of the selected bronze and blue films were measured by a calibrated Ocean Optics Spectrometer USB2000+UV-VIS, proving that they closely match with actual TC film [42].

Using the artificial window, four TC conditions were generated by applying coloured films over the window aperture. These are illustrated in Figure 1. where: a) $VO_2$_S represents the $VO_2$ TC in its switched state, produced by placing two layers of bronze-tinted film over the window to deliver light with a CCT close to 3300 K; b) $VO_2$_U represents the same TC in its unswitched state, produced using a single layer of tinted film to generate light with a CCT close to 4000 K; c) Clear represents the reference case produced using the artificial window without the presence of a tinted film; d) IL_U represents an IL TC in its unswitched state, produced using a single layer of blue-tinted film delivering light with a CCT close to 7000 K; e) IL_S represents the IL TC in its switched state and was produced using two layers of film to generate light with a CCT close to 10000 K. The CCTs produced by the different window conditions were measured using a Konica Minolta CL-200A Chroma meter and are presented in Table 2.

Illuminance levels on the wall opposite the window where one of the test sheets used in the visual tasks was mounted were kept at approximately 350 lux. These levels of illumination are within the recommended illuminance levels for office tasks given in CIBSE

Guide A (300-500 lux) [22]. The literature suggests that under these fixed levels of illumination, the colour temperatures associated with the different types of window would influence human perception of the luminous environment and impact visual task performance [44].

The illuminance produced by the light sources and their CCTs are plotted on a Kruithof diagram at the foot of Figure 1 [44, 45]. It can be seen that, the conditions produced by the $VO_2$ TC glazing in both its switched and unswitched states can be found to lie within the "pleasing" area at the centre of the chart. The IL TC glazing produces conditions that fall into the region where sources are likely to be perceived as "bluish" where there is the potential for discomfort to be experienced.

a. $VO_2$_S b. $VO_2$_U c. Clear d. IL_U e. IL_S

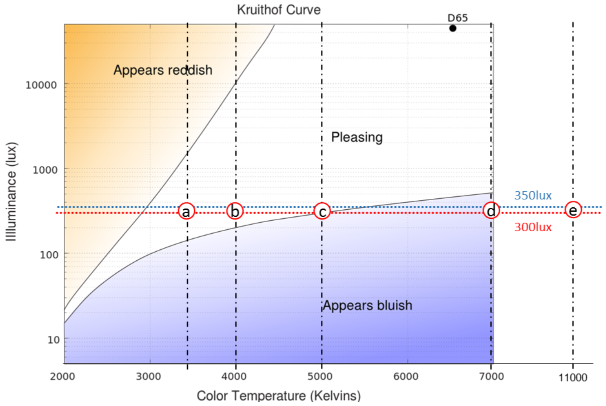


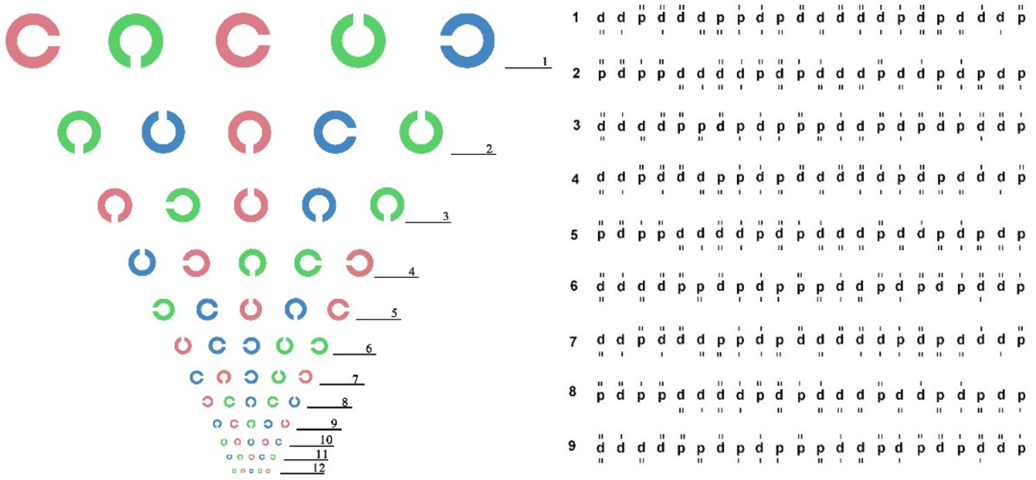
**Figure 1.** Above: Photographs showing the five experimental conditions representing independent variables used in this experiment. The door remained closed during the experiment. Below: The expected human perception of CCT for the five lighting conditions (modern version of the Kruithof curve based on [46] and modified by the authors).

**Table 2.** Vertical and horizontal illuminance measured under the five experimental conditions.

| Window condition | | Vertical surface | | Horizontal surface | |
|---|---|---|---|---|---|
| | | Illuminance (lux) | CCT (K) | Illuminance (lux) | CCT (K) |
| a. | $VO_2$_S | 350 (5 %) | 3410 | 300 (5 %) | 3295 |
| b. | $VO_2$_U | | 4053 | | 3981 |
| c. | Clear | | 4903 | | 4827 |
| d. | IL_U | | 6934 | | 6704 |
| e. | IL_S | | 10997 | | 10557 |

### *2.2. Visual task stimuli: Landolt ring chart and d2 test of attention*

Subjects were asked to perform the two visual tasks shown in Figure 2 (i.e., a Landolt ring test and a d2 test of attention), under each of the five window conditions. A chromatic Landolt chart (Figure 2a) was used to test visual acuity (VA), where subjects were required to identify the position of the gaps in the rings, and colour naming (CN), where subjects were required to identify the colour of each ring. The design of this test was informed by the work of Fotios and Cheal [19]. The rings were coloured green (chromaticity coordinates as given from the chromaticity coordinates, x = 0.401, y = 0.323), blue (x = 0.219, y = 0.231), and red (x = 0.401, y = 0.323). The chart presents subjects with twelve rows of Landolt rings with size decreasing from 8.0 M to 0.63 M (where the unit M specifies the height of typeset materials and is equivalent to 1.5 mm), decreasing in steps of 0.1 log unit per row. Each row contains five rings with at least one of each colour and the rings each have at least one gap positioned in one of four permitted directions, i.e. up, down, left, or right.



(a) Chart of chromatic Landolt rings

(b) Partial sample sheet of the d2 test

**Figure 2.** (a) Chromatic Landolt ring chart used to test visual acuity (gap detection) and colour naming and (b) d2 test for attention.

For both the VA and CN tasks, subjects were asked to provide responses in terms of gap direction and colour for all 60 rings on the chart, working along the rows starting at the top left of the chart. If they were unable to identify the features of a ring, they were asked to make a guess and then move on to the next.

Figure 2(b) shows a sample from the d2 attention task employed in the experiment. Its first application was as a cancellation task [47], and it was subsequently proposed as a useful neuropsychological measurement of sustained attention, concentration processes, and visual scanning speed [48]. The d2 test is conducted using paper and pencil, and observers are asked to cross out any target letter, e.g., the character ‘d’ with two dashes placed above and/or below it, while ignoring all non-target characters (i.e., a ‘d’ with more or fewer than two dashes, and ‘p’ with any dashes) that are interspersed around it. The full version of the d2 test used in this experiment has 14 rows in total, with 47 characters in each row. Observers were asked to complete each row within 20 seconds and move to the next without pausing after the 20 seconds elapse [48].

Table 3 provides the luminance of the background and coloured rings produced under the five window conditions. It may be seen that the luminance of the coloured rings exceeds 25 cd/m$^2$ and so sits within the range of human photopic vision (i.e., luminance > 5 cd/m$^2$) [26]. The Table also provides the contrasts between the rings and the background, which were calculated using equation (1). Under the IL TC conditions, green and blue rings had relatively lower levels of visual contrast than the red ring, while under the $VO_2$ TC condition, the red rings had a lower contrast relative to the other two colours.

According to Weber’s formula, contrast ($C$) is calculated using the background luminance ($L_b$) and target luminance ($L_t$) of each chromatic ring:

https://doi.org/10.1016/j.enbuild.2021.110778
https://escholarship.org/uc/item/9kt889fn

**Table 3**. Background and target (green, red and blue rings) luminance, as well as corresponding calculated contrast. (Lv **(**units cd/m$^2$**),** luminance were measured at the position of observer, and C is the contrast ratio from equation 1).

| | Ring colour | VO$_2$_S | | VO$_2$_U | | Clear | | IL_U | | IL_S | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Lv | C | Lv | C | Lv | C | Lv | C | Lv | C |
| $L_t$ | Green | 37.03 | -0.59 | 37.67 | -0.59 | 38.92 | -0.57 | 39.79 | -0.56 | 38.99 | -0.54 |
| | Red | 46.42 | -0.49 | 44.49 | -0.52 | 39.47 | -0.57 | 40.08 | -0.56 | 36.42 | -0.57 |
| | Blue | 25.19 | -0.72 | 25.16 | -0.73 | 27.26 | -0.70 | 28.12 | -0.69 | 27.56 | -0.67 |
| $L_b$ | Background | 91.39 | | 92.42 | | 90.85 | | 90.2 | | 84.52 | |

Since the amount of time given to complete the d2 test was fixed (i.e., 20 seconds per row on the test sheet), accuracy was the parameter selected to evaluate the performance of subjects. Accuracy in this study was defined by the total number of incorrect responses given by subjects i.e. total errors (TE) [49]. TE indicated the sum of all errors comprising omission (number of target symbols not cancelled) and commission (number of non-target symbols cancelled). Although the total number of correct responses was also evaluated, it was found that subjects consistently completed all the characters on the d2 test sheet within the time limit given to them. Hence, only TE was evaluated in the final analysis.

Because every subject provided responses for all 60 rings in the Landolt test, TE was again used as the sole parameter of accuracy. However, speed (i.e., the time required to provide responses for all 60 rings) was also evaluated as a secondary parameter of performance. Therefore, it was assumed that subjects who completed the tests making fewer errors and had a faster completion time generally performed better. For the d2 test, subjects with the smallest number of errors were considered to have a higher performance.

### *2.3. Subjective assessments*

The questionnaires used in this study consisted of two parts. Part I focussed on evaluating subjects' perception of the test-room luminous environment while tackling the Landolt ring tasks and was administered following their completion under the fifth window condition. Fifteen questions were posed as shown in Table 4.

Part II of the questionnaire presented a further five questions and was completed by the subjects following completion of the final d2 test. The first four questions echoed those in Part I but were framed in the context of the completing the d2 tests.

The questions investigated visual comfort and concentration levels and sought responses exploring brightness, eyestrain, headache, glare, naturalness, preciseness, uniformity, colour temperature, alertness, pleasantness, and overall comfort [50, 51]. Responses were collected via 5-point Likert scales, with bipolar descriptors anchored to the ends of each scale. On each scale, 1 corresponded to the lowest criterion, 5 to the highest and 3 represented a neutral response. All scales ran from left to right in terms of order of magnitude and when performing the statistical analysis, it was assumed that a single point on the scale corresponded to a numerical value of 1.

**Table 4.** Questionnaire Part I: 15 questions exploring subjects' perception of the luminous environment after completing Landolt ring tasks on a vertical surface. Questionnaire Part II: 5 questions exploring subjects' perception of the luminous environment after completing the d2 test of attention on a horizontal surface.

| **Survey questionnaire: Part I** | | |
|---|---|---|
| | Questions | Bipolar descriptions (range [1 - 5]) |
| **1** | Whilst reading the chart, I found the brightness on the chart to be: | Insufficient-------------Sufficient |
| **2** | I felt glare whilst reading the chart: | Strongly disagree---Strongly agree |
| **3** | I had eyestrain whilst reading the chart: | Strongly disagree---Strongly agree |
| **4** | I had a headache after reading the chart: | Strongly disagree---Strongly agree |
| **5** | How would you describe the colours in the chart? | Artificial------------------Natural |
| **6** | The coloured rings on the chart seem to be: | Blurry----------------------Precise |
| **7** | I perceive the room as a whole to be: | Dark------------------------Bright |
| **8** | How would you describe the light distribution in this room? | Uneven-------------------Uniform |
| **9** | How would you describe the feel of the lighting in the room? | Cool ------------------------ Warm |
| **10** | The lighting in the room makes me feel: | Sleepy------------------------Alert |

| | | |
|---|---|---|
| **11** | The lighting conditions in this room make me feel: | Excited------------------------Calm |
| **12** | The lighting conditions in this room make me feel: | Unpleasant----------------Pleasant |
| **13** | On a working day, I predict that I could work under these lighting conditions for: | <1 hour; 1-3 hours; 4-5 hours;<br>6-7 hours; >7 hours |
| **14** | Overall, I find the lighting conditions of this room to be: | Uncomfortable------Comfortable |
| **15** | Do you think this lighting environment is appropriate to conduct office work in? | Unacceptable---------Acceptable |
| **Survey questionnaire: Part II** | | |
| | Questions | Bipolar descriptions (range [1 - 5]) |
| **16** | Whilst doing the ‘d2 test’, I found the brightness on the test sheet to be: | Insufficient-------------Sufficient |
| **17** | I felt glare whilst doing the ‘d2 test’: | Strongly disagree---Strongly agree |
| **18** | I had eyestrain whilst doing the ‘d2 test’: | Strongly disagree---Strongly agree |
| **19** | I had a headache after doing the ‘d2 test’: | Strongly disagree---Strongly agree |
| **20** | How easy was it for you to concentrate when doing the ‘d2 test’? | Difficult --------------------Easy |

### *2.4. Procedure*

A total of 31 subjects were invited to participate in this experiment, which was conducted during October 2017. Subjects were all postgraduate research students or staff working in the Energy Technologies Building at the University of Nottingham (UK) and were aged between 21 and 42 (mean = 30 years, standard deviation = 5.69). Within the sample, 18 were male,13 were female, 20 wore corrective lenses, and none reported colour vision deficiency. The tests were carried out during working hours (9:00-11:00 am and 1:00-3:00 pm) of the day, ensuring that subjects are in good health conditions without fatigue at the beginning of the test. (i.e., effect caused by the time of the day has been tested and analysed. Statistical analysis showed that there is no significant difference ($p>0.05$) caused by time of the day. It proved that this effect can be neglected in this experiment).

All subjects were required to finish five sessions in repeated steps. The duration of each session was approximately 10 minutes and between every two sessions, a five-minute break was provided. For each subject, the experiment commenced with a brief introduction to provide an overview of the study. General information about their recent history and experience was collected using a short questionnaire. An explanation of how to complete the

Landolt ring and d2 tests was given, and following this, each subject was asked to perform three practise trails.

The main part of the experiment then commenced. The procedure for administrating the tests is illustrated in Figure 3. This shows that condition (c), which is the clear window condition shown in Table 2, was always presented to subjects first. In this block, they performed the three visual tasks and completed the two parts of the questionnaire. Because this experiment was primarily designed to evaluate the influence of TC windows, the clear condition was used as a training exercise. This allowed subjects to familiarise themselves with the test procedures before then being exposed to the other window conditions. The clear condition (c) when performing the statistical analyses thereby served as a trial session, reducing the influence of practise effects – particularly for the d2 test [52] – when the visual tasks were performed under the other four conditions.

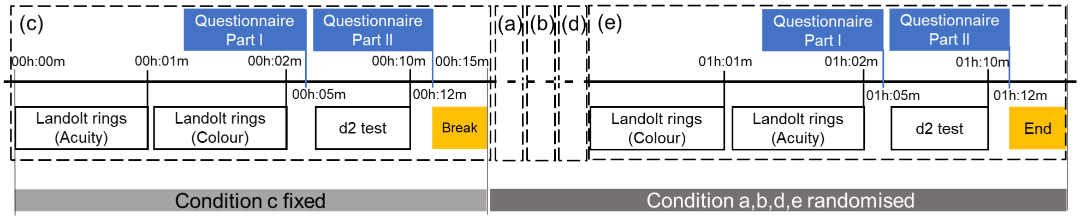


**Figure 3.** Sequence of experimental procedures performed in the order from left to right. Blocks (a), (b), (c), (d) and (e) correspond to the five window conditions shown in Figure 1, whereby condition (c) is the clear condition that is always presented first. The remaining TC conditions (a), (b), (d) and (e) were randomised.

The first block was then repeated for one of the four TC window conditions (i.e. (a), (b), (d) and (e), respectively), which were presented in a randomised sequence. The experimental session for each subject took approximately 1 hour and 12 minutes to complete.

### *2.5. Statistical analyses*

In analysing the data collected under the four TC window conditions (excluding the clear condition), inferential tests were used to determine whether the TC window conditions had a statistically significant influence on visual performance and the subjective assessments given by subjects. Since statistical tests (Kolmogorov-Smirnov and Shapiro-Wilk) showed that the sampling distribution of the data was not normally distributed about the mean, the non-parametric Friedman's Analysis of Variance ANOVA was used to analyse the data [53, 54]. Use of a 5-point Likert scale to collect the subjective assessments meant that the same non-parametric analyses could also be used to analyse these data [55].

The conventional alpha thresholds were used to denote differences that were statistically significant, whereby *p*-values below 0.05, 0.01 and 0.001 were weakly significant, significant and highly significant, respectively. Values above 0.05 were not statistically significant (n.s.).

The Wilcoxon signed-rank test was then used to determine if the differences between each TC window condition were significantly different from each other. Since similar hypotheses were applied across multiple analyses using the Wilcoxon signed-rank test, Bonferroni-Holms corrections were applied to adjust the threshold at which significant results were declared [56].

To help interpret the magnitude of the differences use was made of the effect size values as well as the number of positive, negative and tied ranks [57]. The effect size is a standardised value that can be used to determine the magnitude of the differences being evaluated. In this study, the Pearson's coefficient, $r$, was utilised as the effect size indicator. The calculated ranks show the direction of the differences according to the number that are positive and negative, and a tied rank indicates that there was no difference across the repeated TC condition. To interpret the magnitude of the differences use was made of tables given by Ferguson [49], which gives thresholds referring to 'small', 'moderate' and 'large' effect sizes ($r \geq 0.20$, 0.50 and 0.80 respectively). Values below 0.20 were considered to represent 'negligible' effect sizes.

Subjects were trained to complete the d2 test under the clear window condition prior to completing the main tests, and the data show that higher numbers of errors were made. The clear window condition was therefore excluded from the analyses as it is believed it was influenced by practise effects. Because the survey responses may not have been influenced in the same way by unwanted procedural effects, the subjective assessments given under the clear window condition were included in the analyses.

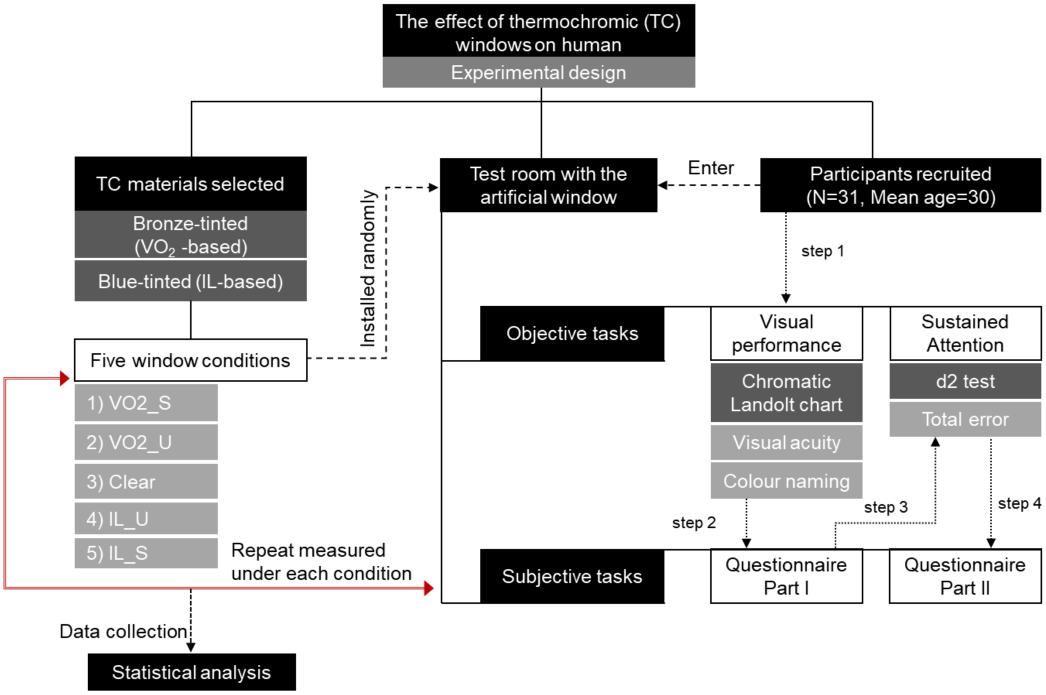


**Figure 4**. The overall research flowchart

## 3. Results

### *3.1. Visual performance: Landolt ring and d2 tests*

The top half of Figure 5 presents boxplots of TE for the test of visual acuity (plot (a)), colour naming (plot (b)), and for the d2 test (plot (c)). In each plot, the results from the corresponding Friedman's ANOVA tests are shown. These were performed twice: once to compare all five window conditions and once to compare the four TC conditions. In plot (a), no significant difference is detected in the TE recorded for the visual acuity test. This result suggests that there was no difference between the five window conditions, or the four TC conditions in relation to their effect on visual acuity as measured using the Landolt ring test.

The results for the colour naming test in plot (b) indicate that the responses to the clear condition (i.e., the window that was presented first to subjects) show an increased incidence of error. This could be explained by a practise effect, whereby more errors were made under the clear condition as subjects were familiarising themselves with the test procedure. The Friedman's ANOVA test demonstrated that differences in errors made across the five window conditions (shown in Table 2) were statistically significant. In the separate analysis that focussed on analysing the four window conditions without the clear window case, the differences across the four TC windows were also statistically significant. Since these four conditions were randomised in the experiment, any practise effects are minimised across these comparisons, which indicates that the differences in TE are a consequence of the window properties. This suggests that colour discrimination was influenced by the different TC window conditions. For the analysis of speed of task completion for the visual acuity and colour naming tasks, plot (d) and plot (e) show that no significant differences were found across the four TC window conditions. When considering the clear condition, the colour naming task, plot (e) shows no significant difference but the visual acuity test showed that subjects took significantly longer to complete the task compared to the TC window cases. This may again be due to the practise effect.

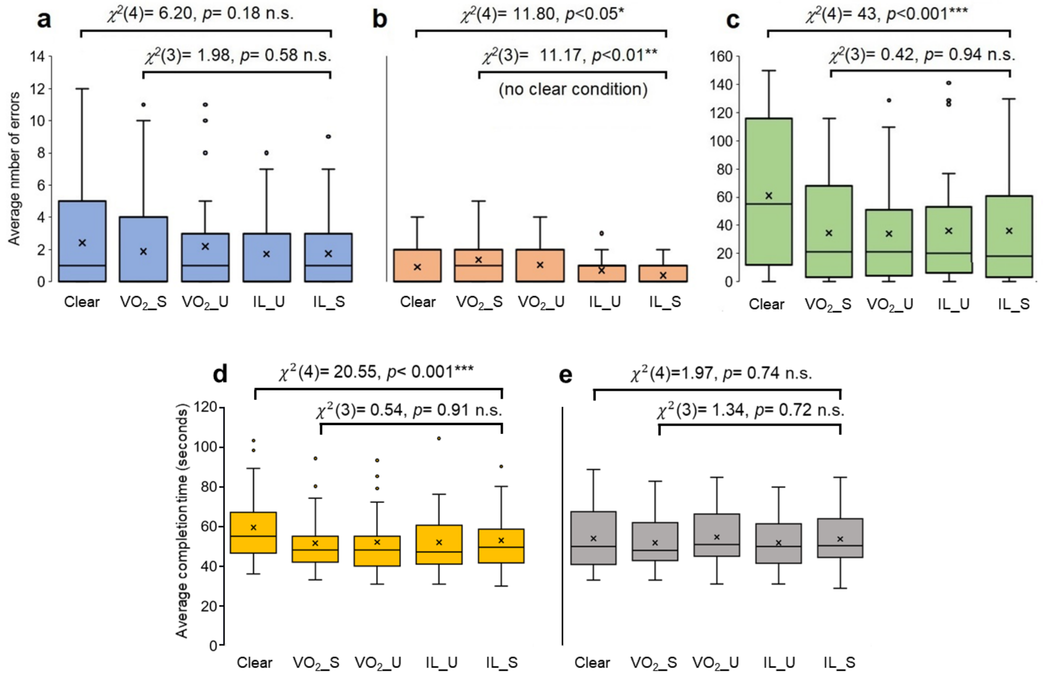


**Figure 5.** Top: Boxplots showing the minimum, 25th quartile, median (line) and mean (cross) averages, 75th quartile, and maximum values of error occurred during the visual acuity (a) and colour naming (b) tasks using the Landolt ring test and TE (c) for the d2 test under the five different window conditions. Bottom: Boxplots showing the average values of time spent performing the visual acuity (d) and colour naming (e) tasks for the Landolt ring test. The Friedman's ANOVA test $\chi^2$ value is shown in each plot twice to compare the five conditions (with the clear case) and four TC conditions (without the clear case).

Table 5 shows the Wilcoxon signed-rank test results for the colour naming task. This provides the descriptive statistics (i.e. the median and mean differences), statistical significance (*p*-value), ranks (positive, negative and ties), and the effect size (*r*). These analyses compared only the four TC window conditions against each other to determine their influence on TE from the Landolt ring test. It may be seen that two out of the six cases were found to be significantly different: the $VO_2$_S vs. IL_U and $VO_2$_S vs. IL_S conditions. Looking at the related boxplots in Figure 5, plot (b), it may be seen that subjects made fewer errors under the IL_U and IL_S conditions than under the $VO_2$_S condition (i.e., they performed with higher accuracy under the blue-tinted TC conditions). The remaining four pairwise comparisons show non-significant differences across the TC windows conditions.

Although plot (b) shows that subjects made a higher number of errors under the $VO_2$_U condition as compared with the IL_S condition, the statistical comparison of the two showed only a small effect size and was not statistically significant after the correction applied to the significance level.

**Table 5**. Wilcoxon signed-rank test used to test the difference in TE made in the colour naming task using the Landolt ring test.

| Comparisons | ΔMedian | ΔMean | *p*-value | Positive | Negative | Ties | Effect Size (*r*) |
|---|---|---|---|---|---|---|---|
| IL_S vs. IL_U | -1 | -0.65 | 0.06 n.s. | 4 | 10 | 17 | **-0.24** |
| $VO_2$_U vs.IL_U | -1 | 0.32 | 0.16 n.s. | 9 | 8 | 14 | -0.18 |
| $VO_2$_S vs.IL_U | 0 | 0.65 | 1.1 * | 13 | 5 | 13 | **-0.33** |
| $VO_2$_U vs.IL_S | 0 | 0.65 | 0.02 n.s. | 12 | 5 | 14 | **-0.30** |
| $VO_2$_S vs.IL_S | 1 | 0.97 | 0.001*** | 16 | 2 | 13 | **-0.44** |
| $VO_2$_S vs.$VO_2$_U | 1 | 0.32 | 0.28 n.s. | 11 | 7 | 13 | -0.14 |

Bonferroni-Holms corrected: ***highly significant; **significant; *weakly significant; n.s.= not significant
$0.20<r<0.50$= “small” effect size; $r<0.20$= “negligible” effect size

A higher average TE was found for the acuity task, as shown in plot (a), than the colour naming task, as shown in plot (b). Although there was no significant difference in TE within each task across the four TC windows, subjects consistently gave more incorrect responses for the acuity task than for the colour naming task.

The boxplot of TE for the d2 test is shown in plot (c). Friedman’s ANOVA tests were performed to compare the five window conditions (i.e., including the clear condition) and repeated for the four TC conditions (i.e., without clear condition). The analyses detected that there was a statistically significant difference across the five window conditions, but no significant difference across the four TC conditions. This indicates that the clear condition was different from the other four TC conditions, which is borne out by inspection of the boxplots in plot (c). Again, the results for the clear window test were excluded from further analyses on the assumption that this reduced learning effects from the d2 test. The absence of any statistically significant difference in the results for the four TC conditions suggests that the luminous environments they created had no effect on concentration as measured by the d2 test of sustained attention. Subjects were given a fixed period of 20s to complete each row of the d2 tests and as a consequence, a comparison of time taken to complete tasks cannot be made for the different TC conditions.

### *3.2. Subjective assessments*

Subjects were asked to provide subjective ratings to the 15 questions in Part I of the questionnaire after completing each Landolt ring test as indicated in Figure 5. The Friedman's ANOVA test was used to analyse the differences across the five window conditions and this identified the eight questions in Table 6 as showing statistically significant results.

Part II of the questionnaire comprised five questions that were presented following completion of the d2 test. It aimed to investigate subjects' perceptions after having to sustain short periods of concentration under the five different window conditions. Table 6 reports the results of the Friedman's ANOVA test for the responses to the questions and this identified two questions with statistically significant results. Significant results from the pairwise comparisons of responses for the questions identified as having statistical significance from the Part I and Part II questionnaires are provided in Appendix A.

It is interesting to note that the majority of questions exploring visual discomfort (Questions 1, 2, 3, 4 and 17, 18, 19) showed no statistical significance indicating that at an illuminance of 300lux on a workplane, the colour temperatures investigated between 3000K and 10000K did not have an effect on subjects' level of visual discomfort when completing visual acuity, colour identification, and concentration tasks.

**Table 6.** Friedman's ANOVA test on responses to questions with significant results in questionnaire Parts I and II.

| **Questions** | | ***p*-value** |
|---|---|---|
| **Part I** | | |
| (Q5) How would you describe the colours in the chart? | 18.41 | 0.001*** |
| (Q7) I perceive the room as a whole to be: | 21.43 | 0.000*** |
| (Q9) How would you describe the feel of the lighting in the room? | 13.12 | 0.011** |

| | | |
|---|---|---|
| (Q10) The lighting in the room makes me feel: | 13.12 | 0.011** |
| (Q12) The lighting conditions in this room make me feel: | 30.62 | 0.000*** |
| (Q13) On a working day, I predict that I could work under these lighting conditions for: | 20.18 | 0.000*** |
| (Q14) Overall, I find the lighting conditions of this room to be: | 30.19 | 0.000*** |
| (Q15) Do you think this lighting environment is appropriate to conduct office work in? | 25.36 | 0.000*** |
| **Part II** | | |
| (Q16) Whilst doing the 'd2 test', I found the brightness on the test sheet to be: | 11.52 | 0.021* |
| (Q20) How easy was it for you to concentrate when doing the 'd2 test'? | 19.48 | 0.001*** |

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

Note: For pairwise comparisons, see Appendix A for Part I and Appendix B for Part II

To help understand why there was a significant difference in the number of errors found when subjects performed the colour naming task, specific emphasis was placed on the evaluations given to three of the questions found in Table 6. These were Q5, which asked subjects to describe the quality of the colours on the chart, and Q10 and Q15, which related to task performance. Boxplots of the responses to these questions are presented in Figure 6 for the five window conditions.

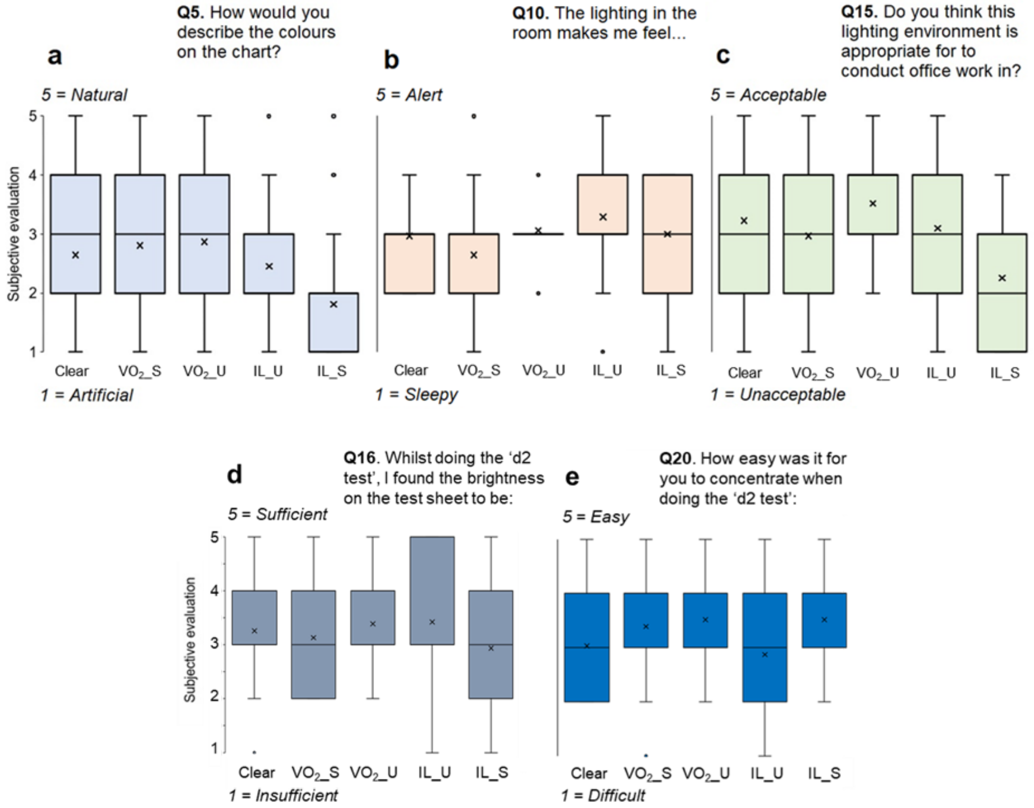


**Figure 6.** Boxplots of the responses to question 5 (a), question 10 (b), and question 15 (c) from Questionnaire Part I (Table 4); question 16(d), question 20 (e) from Questionnaire Part II (Table 4).

The results from Q5 describing the fidelity of the colours on the Landort chart are presented in plot (a) and show that the IL_S window created the strongest perception of artificial colour rendering. The subjective evaluation scores for the remaining four window conditions are similar to each other and suggest that while colour rendering was not perceived at natural, it was deemed better than for the IL_S window. Pairwise comparisons of the responses to Q5, given in Appendix A, also show that all statistically significant differences were found only when window conditions were compared against the IL_S case.

The results for Q10 seeking self-assessments of how the five window conditions made participants feel are shown in plot (b). These indicate higher levels of alertness under the IL_U and IL_S conditions, and greater levels of sleepiness under the clear and $VO_2$_S cases. The differences between the $VO_2$_S case and both TC conditions in their unswitched state are evident in the pairwise comparison of responses to Q10, given in Appendix A.

When evaluating the working conditions across the five window conditions through Q15, it is apparent from the plot (c) that the IL_S window was deemed the least acceptable. The remaining three TC window conditions were perceived as more acceptable and received scores similar to those for the clear condition, suggesting that the $VO_2$_S, $VO_2$_U and IL_U windows might be as acceptable as a clear glass window. Again, the pairwise comparisons for Q15, given in Appendix A, show a high proportion of statistically significant difference for comparisons made with the IL_S source.

Even though no significant difference was detected in the performance of subjects in the d2 test, it is interesting to understand their subjective evaluations of the luminous environments they were asked to work within. Figure 6, plot (d) presents the results to Q16 exploring the illumination of the test sheet. The paired comparison identified two significant differences, each involving the IL_S condition which was perceived as delivering insufficient illumination as compared with the two unswitched conditions as shown in Appendix B.

The results for Q20 exploring concentration are presented in Figure 6, plot (e). The paired comparison tests in Appendix B indicate that, among the four TC window conditions, IL_S was given the lowest rating of concentration, which means that subjects felt the IL_S window condition led to a decrease in concentration, again as compared with $VO_2$_S and the two unswitched TC window conditions. The tests also indicate subjects were better able to concentrate under the $VO_2$_U and $VO_2$_S window conditions as compared to the clear window condition, revealing that bronze-tinted TC window is more conducive to concentration.

## 4. Discussion

The clear window was used as a training exercise and it is assumed that the higher levels of error reported are a consequence of learning effects [58, 59]. While this is in keeping with the findings of other studies, this conclusion is not definitive and other explanations are not ruled out. Although data and analysis are presented for the clear window, beyond the observation above, no conclusions are drawn from them.

The results for the TC conditions indicate higher TE occurred in the visual acuity test than in the colour naming test. Although both use the same Landolt rings, the challenges presented are different and higher incidence in TE observed is probably a consequence of gap detection, which was used to quantify visual acuity, being more challenging than colour naming. When subjecting the results to statistical analysis, no significant difference was found for visual acuity performance under the four TC conditions, but significant difference was found for colour naming performance indicating that the latter is affected by the TC conditions. This suggests that the CCTs used in this study, under an illuminance of 300lux, have no effect on acuity but do affect the discrimination of colour.

The results from the d2 test, which show no statistically significant differences in the ability of subjects to concentrate on a task under the four TC conditions, are not consistent with previous findings. These indicate that intermediate (e.g., 4300K) or slightly cold (e.g., 6500K) lighting conditions can lead to higher levels of concentration than lighting conditions that sit above and below these on the Planckian locus [37, 38]. The disparity could be explained by studies that link performance in the d2 test to the level of education of subjects, i.e. subjects with a higher level of education maintain a relatively higher scanning speed and exhibit lower error occurrence when completing the d2 test [60]. In this study, subjects were all university educated to postgraduate level, which may explain why no significant differences were identified across the four TC conditions. The fact that subjects perceived the IL_S condition as being relatively difficult for maintaining concentration may mean that further experiments extending the duration of the d2 test, beyond the relatively short-term task used in this study, to a longer-term task more representative of the work place, may yield different results.

While the results related to task performance did not suggest any significant effects of TC conditions, beyond colour naming, the subjective evaluations made by subjects in relation to the environment and how this affected task performance yielded some interesting observations.

Despite the finding that the bluish/cool conditions created by the IL_S window were perceived to be the most artificial of those presented, this window, in both its switched and unswitched state, i.e. IL_S and IL_U, was judged to promote higher levels of alertness during the experiment. This contrasted against the relatively warmer appearance of the $VO_2$_S condition, which was perceived to create a stronger sensation of 'sleepiness' - a result consistent with studies indicating that increasing CCT can improve the alertness levels of building occupants [29, 61]. The IL_S condition also received the lowest evaluations of acceptability, this being the condition subjects would want to spend the least time in with a majority of respondents indicating they would wish to occupy the space for less than one hour. This contrasts with the $VO_2$_U condition, which received the greatest acceptance, with subjects indicating they could tolerate for between four and five hours. These evaluations are consistent with subjects' opinion on what constituted the most comfortable condition, produced by the $VO_2$_U window with a CCT of 4000K, and uncomfortable condition, produced by the IL_S window condition with CCT of 11000K. Previous studies produced similar findings, suggesting that warm lighting with a CCT of around 3500K is beneficial for visual comfort [23, 62]. The subjective evaluations of concentration show that the IL_S window condition led to a decrease in concentration, similar to previous findings [37, 38].

In summary, when comparing the TC windows against each other, the findings in this study indicated that, blue/cool tinted TC windows (i.e. IL_U and IL_S) create luminous conditions on indoor surfaces that are perceived to be unnatural and create lower the levels of acceptability in the building. However, they may support higher levels of alertness and were found to best support colour naming tasks. For bronze/warm tinted TC windows (e.g. $VO_2$_S and $VO_2$_U), higher levels of acceptability are observed, however, they produced higher levels of sleepiness and more errors were made when the colour naming task was performed.

Although TC glazing represents a technology that can reduce the energy requirements in buildings, the quality of the luminous environment they create, the visual comfort needs of occupants, and the impact they have on task performance needs to be carefully considered. By identifying their wider impacts on building occupants in this study, it would be possible to apply the most suitable type of TC windows for any given building context.

### *4.1. Limitations*

This study was conducted in an environment illuminated under simulated daylight using an artificial window. While this approach does offer controlled conditions under which to explore the effects of the different simulated TC windows, it is not possible to fully replicate the luminous conditions created by real TC glazing transmitting daylight. The techniques such as virtual reality [63-65] could be used to explore dynamic change of colour under varying environmental conditions, the more complex luminous environments created in real building spaces, and potential impact on the experience of views to the outside. The results from tests such as these could inform the design of experiments conducted under field conditions. While this study measured visual performance and subjective assessments, additional physiological measurements could also be utilised (e.g. heart rate, skin conductance) to determine whether TC windows influence the health of building occupants[63, 64]. Additionally, the functions of windows are complex, including daylighting, ventilation, view outside, privacy protection, and characterization of the buildings. Therefore, the artificial window (i.e., LED lamps array with a diffuser) is necessary to be improved to mimic the multi-functions of a practical smart window in the further study, especially, enabling the evaluation on human response to view it provides.

The experimental procedure used in this study sought to counter a potential learning effect related to completion of the Landolt ring and d2 tests by presenting all subjects with the clear window condition at the beginning of the experiment and then randomising the presentation of the TC conditions. Another type of procedure that could be used in further studies is a between-subjects design, whereby subjects are randomly assigned to different groups and do not evaluate every experimental condition. This approach could be structured to still address the issue of learning effects while at the same time including the clear window condition within the analysis.

## 5. Conclusions

In this study, an artificial window was used to simulate clear glazing and glazing with four different TC conditions in a test cell designed to represent an office environment. Under an illuminance of around 300 lux, the translucent films used to simulate TC conditions produced light with a bronze (representing vanadium dioxide based TC in its unswitched and switched states denoted by $VO_2$_U and $VO_2$_S respectively) and blue (representing an ionic liquid based TC in its unswitched and switched states denoted by IL_U and IL_S respectively) appearance. Subjects were asked to evaluate the conditions and perform visual tasks under the different window conditions. This involved the use of questionnaires (subjective tests) and a coloured Landolt ring chart and the d2 test of attention (objective tests). To analyse the data, statistical analysis was performed and detected the following results:

1) When performing a colour naming test using the Landolt ring chart, subjects made more errors under the bronze ($VO_2$_U and $VO_2$_S) conditions than under the blue (IL_U and IL_S) conditions. The differences across the four TC window conditions were shown to be statistically significant.
2) No significant differences were found in the number of errors made between TC conditions when subjects used the Landolt ring chart to test visual acuity (gap detection). The same finding was obtained for the time it took subjects to complete this task.
3) Results from the d2 test of attention showed no significant differences in the number of errors made or task completion time between TC conditions, suggesting the different environments produced have no effect on concentration. While subjective assessment indicates that bronze-tinted TC window is relatively more conducive to concentration.
4) Practise effects were anticipated for both the Landolt ring and d2 tests and addressed by presenting subjects with the clear window condition at the start of their session. It was not confirmed whether the significant differences obtained between these results and those obtained under TC conditions were due to learning or the quality of the luminous environment presented.

5) The blue-tinted TC window conditions (IL_U and IL_S) were shown to increase subjective evaluations of alertness given by subjects, but created a luminous environment perceived as more artificial and unacceptable in a work environment as compared to the other conditions.

This study provides information useful in the development of TC materials applied to windows, as well as the application of TC glazing for building design. Besides aesthetics and energy-saving considerations, occupant visual perception and concentration are both significant issues to be considered for a space lit by daylight filtered through TC windows.

## Acknowledgements

This work was supported by the Faculty of Engineering at the University of Nottingham through a PhD studentship awarded to Runqi Liang, and project funded by China Postdoctoral Science Foundation project (2019M661617) and supported by Open Projects Fund of Key Laboratory of Ecology and Energy-saving Study of Dense Habitat (Tongji University), Ministry of Education (2019020106).

## References

1. M. Saeli, C. Piccirillo, et al., *Nano-composite thermochromic thin films and their application in energy-efficient glazing.* Solar Energy Materials and Solar Cells, 2010. **94**(2): p. 141-151.
2. S. Hoffmann, E. S. Lee, C. Clavero, *Examination of the technical potential of near-infrared switching thermochromic windows for commercial building applications.* Solar Energy Materials and Solar Cells, 2014. **123**: p. 65-80.
3. M. Saeli, C. Piccirillo, et al., *Energy modelling studies of thermochromic glazing.* Energy and Buildings, 2010. **42**(10): p. 1666-1673.

4. H. Ye, L. Long, *Smart or not? A theoretical discussion on the smart regulation capacity of vanadium dioxide glazing.* Solar Energy Materials and Solar Cells, 2014. **120**: p. 669-674.

5. H. Ye, X. Meng, B. Xu, *Theoretical discussions of perfect window, ideal near infrared solar spectrum regulating window and current thermochromic window.* Energy and Buildings, 2012. **49**: p. 164-172.

6. R. Liang, Y. Sun, et al., *Evaluation of the thermal and optical performance of thermochromic windows for office buildings in China.* Energy and Buildings, 2018. **176**: p. 216-231.

7. Y. Gao, H. Luo, et al., *Nanoceramic VO2 thermochromic smart glass: A review on progress in solution processing.* Nano Energy, 2012. **1**(2): p. 221-246.

8. C. S. Blackman, C. Piccirillo, et al., *Atmospheric pressure chemical vapour deposition of thermochromic tungsten doped vanadium dioxide thin films for use in architectural glazing.* Thin Solid Films, 2009. **517**(16): p. 4565-4570.

9. X. Wei, L. Yu, et al., *Solar-thermochromism of Pseudocrystalline Nanodroplets of Ionic Liquid-Ni Complexes immobilized inside translucent Microporous PVDF Films.* Advanced Materials, 2009. **21**: p. 776-780.

10. X. J. Wei, L. P. Yu, et al., *Thermo-solvatochromism of chloro-nickel complexes in 1-hydroxyalkyl-3-methyl-imidazolium cation based ionic liquids.* Green Chemistry, 2008. **10**(3): p. 296-305.

11. Pleotint. *Suntuitive Self-tinting Glass Technical Information*. 2016 [cited 2018 http://www.suntuitive.com/uploads/1/2/4/1/12414735/suntuitive_technical_brochure_063017.pdf].

12. R. Liang, Y. Sun, et al., *An exploration of the combined effects of NIR and VIS spectrally selective thermochromic materials on building performance.* Energy and Buildings, 2019. **201**: p. 149-162.

13. M. Aburas, V. Soebarto, et al., *Thermochromic smart window technologies for building application: A review.* Applied Energy, 2019. **255**.

14. R. Binions, C. Piccirillo, et al., *Hybrid aerosol assisted and atmospheric pressure CVD of gold doped vanadium dioxide.* Chemical Vapor Deposition, 2008. **14(12)**: p. 33-39.

15. Z. Huang, C. Chen, et al., *Tungsten-doped vanadium dioxide thin films on borosilicate glass for smart window application.* Journal of Alloys and Compounds, 2013. **564**: p. 158-161.

16. A. D. Galasiu, J. A. Veitch, *Occupant preferences and satisfaction with the luminous environment and control systems in daylit offices: a literature review.* Energy and Buildings, 2006. **38**(7): p. 728-742.

17. F. M. Vossen, M. P. J. Aarts, M. G. Debije, *Visual performance of red luminescent solar concentrating windows in an office environment.* Energy and Buildings, 2016. **113**: p. 123-132.

18. J. Y. Shin, G. Y. Yun, J. T. Kim, *View types and luminance effects on discomfort glare assessment from windows.* Energy and Buildings, 2012. **46**: p. 139-145.

19. S. A. Fotios, C. Cheal, *Lighting for subsidiary streets investigation of lamps of different SPD. Part1—Visual Performance.* Lighting Research and Technology, 2007. **39**(3): p. 215-232.

20. B. Manav, *An experimental study on the appraisal of the visual environment at offices in relation to colour temperature and illuminance.* Building and Environment, 2007. **42**(2): p. 979-983.

21. M.-C. Dubois, *Effect of Glazing Types on Daylight Quality in Interiors: Conclusions from Three Scale Model Studies*, in *EXPERIENCING LIGHT 2009 International Conference on the Effects of Light on Wellbeing*, Y.A.W.d. Kort, et al., Editors. 2009: Eindhoven, the Netherlands. p. 86-97.

22. CIBSE, *Guide A: Environmental design*. 2006, CIBSE Publications: London.

23. M. Wei, K. W. Houser, et al., *Field study of office worker responses to fluorescent lighting of different CCT and lumen output.* Journal of Environmental Psychology, 2014. **39**: p. 62-76.

24. E. Öztürk, S. Yılmazer, S. E. Ural, *The effects of achromatic and chromatic color schemes on participants' task performance in and appraisals of an office environment.* Color Research & Application, 2012. **37**(5): p. 359-366.

25. K. Yildirim, A. Akalin-Baskaya, M. L. Hidayetoglu, *Effects of indoor color on mood and cognitive performance.* Building and Environment, 2007. **42**(9): p. 3233-3240.

26. P. R. Boyce, *Human factors in lighting* 2014: CRC Press.

27. R.E.Harrington, *Effect of Color Temperature on Apparent Brightness.* Journal of The Optical Society of America, 1954. **44**(2).

28. S.Fotios, G. J. Levermore, *perception of electric light sources of different colour properties.* Lighting Research and Technology, 1997. **29**(3): p. 161-171.

29. K. C. H. J. Smolders, Y. A. W. de Kort, *Investigating daytime effects of correlated colour temperature on experiences, performance, and arousal.* Journal of Environmental Psychology, 2017. **50**: p. 80-93.

30. Revantino, R. A. Mangkuto, et al., *The effects of illuminance, colour temperature, and colour rendering of various existing light-emitting diode lamps on subjective preference and performance in Indonesia.* Journal of Building Engineering, 2018. **19**: p. 334-341.

31. A. Brambilla, W. Hu, et al., *How correlated colour temperature manipulates human thermal perception and comfort.* Building and Environment, 2020. **177**.

32. I. Golasi, F. Salata, et al., *Influence of lighting colour temperature on indoor thermal perception: A strategy to save energy from the HVAC installations.* Energy and Buildings, 2019. **185**: p. 112-122.

33. P. F. Boray, R. Gifford, L. Rosenblood, *Effects of warm white cool white and full spectrum fluorescent lighting on simple cognitive performance, mood and ratings of others.* Journal of Environmental Psychology, 1989. **9**(4): p. 297-307.

34. P. L. Vrabel, C. A. Bernecker, R. G. Mistrick, *Visual performance and visual clarity under electric light sources: Part II-visual clarity.* Journal of the Illuminating Engineering Society, 1998. **27**(1): p. 107-129.

35. E. Rautkyla, M. Puolakka, et al., *Effects of correlated colour temperature and timing of light exposure on daytime alertness in lecture environments.* . Journal of Light and Visual Environment, 2010. **34**(2): p. 59-68.

36. P. J. C. Sleegers, N. M. Moolenaar, et al., *Lighting affects students' concentration positively: findings from three Dutch studies.* Lighting Research and Technology, 2013. **45**: p. 159-179.

37. B. M. T. Shamsul, C. C. Sia, et al., *Effects of Light's Colour Temperatures on Visual Comfort Level, Task Performances, and Alertness among Students.* American Journal of Public Health Research, 2013. **1**(7): p. 159-165.

38. R. H. Huang, L. Lee, et al., *Effects of correlated color temperature on focused and sustained attention under white LED desk lighting.* Inc. Col Res Appl, 2015. **40**: p. 281-286.

39. H. Yu, T. Akita, *The effect of illuminance and correlated colour temperature on perceived comfort according to reading behaviour in a capsule hotel.* Building and Environment, 2019. **148**: p. 384-393.

40. M. Bodart, A. Deneyer, *A guide for the building of daylight scale models*, in *The 23rd Conference on Passive and Low Energy Architecture*. 2006: Geneva, Switzerland.

41. HSE, *L24: Workplace health, safety and welfare (2nd edition)* 2013, Health and Safety Exective.

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

43. S. Li, *VO2-based Thermochromic and Nanothermochromic Materials for Energy-Efficient Windows*, in *Science and technology, Uppsala University*. 2013. p. 142.

44. I. Ashdown, *The Kruithof Curve: A Pleasing Solution.* https://www.researchgate.net/publication/273763550, 2015.

45. S. Fotios, *A Revised Kruithof Graph Based on Empirical Data.* Leukos, 2016. **13**(1): p. 3-17.

46. F. Viénot, M.-L. Durand, E. Mahler, *Kruithof's rule revisited using LED illumination.* Journal of Modern Optics, 2009. **56**(13): p. 1433-1446.

47. R. Brickenkamp, E. Zillmer, *The d2 Test of Attention*. 1998, Seattle: Washington: Hogrefe & Huber Publishers.

48. R. M. Ross, *The D2 Test of Attention: An Examination of Age, Gender, and Cross-cultural Indices*. 2005: Argosy University.

49. C. J. Ferguson, *An effect size primer: A guide for clinicians and researchers.* Professional Psychology: Research and Practice, 2009. **40**(5): p. 532-538.

50. S. Odabasioglu, N. Olgunturk, *Effects of coloured lighting on the perception of interior spaces.* Percept Mot Skills, 2015. **120**(1): p. 183-201.

51. A. Borisuit, *Effects of realistic office daylighting and electric lighting conditions on visual comfort, alertness and mood.* Lighting Research and Technology, 2015. **47**: p. 192-209.

52. J. Scharfen, J. M. Peters, H. Holling, *Retest effects in cognitive ability tests: A meta-analysis.* Intelligence, 2018. **67**: p. 44-66.

53. M. Friedman, *The use of ranks to avoid the assumption of normality implicit in the analysis of variance.* American Statistical Association, 1937. **32**(200): p. 675-701.

54. M. Friedman, *A comparison of alternative tests of significance for the problem of m rankings.* . The Annals of Mathematical Statistics, 1940. **11**(1): p. 86-92.

55. I. E. Allen, C. A. Seaman, *Likert scales and data analysis.* Quality Progress, 2007. **40(7)**: p. 64-65.

56. S. Holm, *A Simple Sequentially Rejective Multiple Test Procedure.* Scandinavian Journal of Statistics, 1979. **6**: p. 65-70.

57. E. Whitley, J. Ball, *Review: Statistics review 6: Nonparametric methods.* Critical Care, 2002. **6**(6): p. 509.

58. M. G. Kent, S. Altomonte, et al., *Temporal effects on glare response from daylight.* Building and Environment, 2017. **113**: p. 49-64.

59. M. G. Kent, T. Cheung, et al., *A Bayesian method of evaluating discomfort due to glare: The effect of order bias from a large glare source.* Building and Environment, 2018. **146**: p. 258-267.

60. M. E. Bates, E. P. Lemay, *The d2 Test of Attention: Construct validity and extensions in scoring techniques.* Journal of the International Neuropsychological Society, 2004. **10**: p. 392-400.

61. I. Iskra-Golec, A. Wazna, L. Smith, *Effects of blue-enriched light on the daily course of mood, sleepiness and light perception: A field experiment.* Lighting Res. Technol., 2012. **44**: p. 506-513.

62. S. Han, *The effect of illuminance, CCT and decor on the perception of lighting*. 2002, Rensselaer Polytechnic Institude: Troy, NY.

63. F. Abd-Alhamid, M. Kent, et al., *Evaluating the impact of viewing location on view perception using a virtual environment.* Building and Environment, 2020: p. 106932.

64. F. Abd-Alhamid, M. Kent, et al., *Developing an Innovative Method for Visual Perception Evaluation in a Physical-Based Virtual Environment.* Building and Environment, 2019. **162**: p. 106278.

65. J. Li, Y. Jin, et al., *Building environment information and human perceptual feedback collected through a combined virtual reality (VR) and electroencephalogram (EEG) method.* Energy and Buildings, 2020. **224**.

## Appendix

**(A)** Wilcoxon signed-rank test of responses to questions in questionnaire Part I with significant results produced under five studied lighting conditions

| Questions | Hypothesis (M1 vs M2) | $M_{1dn}$ | $M_{2dn}$ | *p*-value | Positive | Negative | Ties | *Z* | Effect Size (*r*) |
|---|---|---|---|---|---|---|---|---|---|
| Q5 | IL_S vs Clear | 2 | 3 | 0.004** | 5 | 17 | 9 | -2.910 | -0.370 |
| | IL_S vs IL_U | 2 | 2 | 0.007** | 4 | 14 | 13 | -2.686 | -0.341 |
| | IL_S vs VO2_U | 2 | 3 | 0.001*** | 3 | 20 | 8 | -3.217 | -0.409 |
| | VO2_S vs IL_S | 3 | 2 | 0.002** | 16 | 3 | 12 | -3.159 | -0.401 |
| Q7 | VO2_U vs Clear | 3 | 4 | 0.001*** | 3 | 20 | 8 | -2.288 | -0.291 |
| | VO2_S vs Clear | 3 | 4 | 0.006** | 3 | 18 | 10 | -2.287 | -0.290 |
| | IL_U vs Clear | 3 | 4 | 0.022* | 4 | 14 | 13 | -3.318 | -0.421 |
| | VO2_U vs Clear | 3 | 4 | 0.022* | 3 | 13 | 15 | -2.748 | -0.349 |
| | IL_S vs IL_U | 2 | 3 | 0.030* | 5 | 16 | 10 | -2.164 | -0.275 |
| Q9 | IL_U vs Clear | 2 | 2 | 0.027* | 5 | 17 | 9 | -2.215 | -0.281 |
| | VO2_U vs Clear | 3 | 2 | 0.000*** | 22 | 0 | 9 | -4.193 | -0.533 |
| | IL_S vs Clear | 1 | 2 | 0.000*** | 4 | 25 | 2 | -3.906 | -0.496 |
| | VO2_S vs Clear | 4 | 2 | 0.000*** | 26 | 0 | 5 | -4.517 | -0.574 |
| | VO2_U vs IL_U | 3 | 2 | 0.000*** | 28 | 0 | 3 | -4.697 | -0.597 |
| | IL_S vs IL_U | 1 | 2 | 0.003** | 1 | 12 | 18 | -3.000 | -0.381 |
| | VO2_S vs IL_U | 4 | 2 | 0.000*** | 30 | 1 | 0 | -4.853 | -0.616 |
| | IL_S vs VO2_U | 1 | 3 | 0.000*** | 0 | 31 | 0 | -4.934 | -0.627 |
| | VO2_S vs VO2_U | 4 | 3 | 0.001*** | 16 | 2 | 13 | -3.252 | -0.413 |
| | VO2_S vs IL_S | 4 | 1 | 0.000*** | 30 | 0 | 1 | -4.871 | -0.619 |
| Q10 | IL_S vs IL_U | 3 | 3 | 0.038* | 2 | 8 | 21 | -2.070 | -0.263 |
| | VO2_S vs IL_U | 3 | 3 | 0.009** | 4 | 18 | 9 | -2.628 | -0.334 |
| | VO2_S vs VO2_U | 3 | 3 | 0.011* | 2 | 14 | 15 | -2.540 | -0.323 |
| Q12 | IL_S vs VO2_U | 2 | 3 | 0.000*** | 2 | 20 | 9 | -3.765 | -0.478 |
| | VO2_S vs IL_S | 3 | 2 | 0.001*** | 18 | 4 | 9 | -3.286 | -0.417 |
| | IL_S vs Clear | 2 | 3 | 0.003** | 3 | 16 | 12 | -2.982 | -0.379 |
| | VO2_S vs Clear | 3 | 3 | 0.036* | 13 | 4 | 14 | -2.101 | -0.267 |
| | VO2_U vs IL_U | 3 | 3 | 0.006** | 15 | 2 | 14 | -2.751 | -0.349 |
| | IL_S vs IL_U | 2 | 3 | 0.037* | 4 | 14 | 13 | - | -0.264 |

| | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | | 2.082 | |
| | VO2_S vs IL_U | 3 | 3 | 0.033* | 14 | 4 | 13 | -2.130 | -0.271 |
| | VO2_U vs Clear | 3 | 3 | 0.010* | 14 | 2 | 15 | -2.588 | -0.329 |
| Q13 | IL_S vs IL_U | 2 | 2 | 0.002** | 2 | 15 | 14 | -3.13 | -0.398 |
| | IL_S vs VO2_U | 2 | 3 | 0.003** | 5 | 12 | 5 | -2.983 | -0.379 |
| | VO2_S vs VO2_U | 3 | 3 | 0.029* | 2 | 9 | 20 | -2.179 | -0.277 |
| | VO2_S vs IL_S | 3 | 2 | 0.027* | 16 | 6 | 9 | -2.217 | -0.282 |
| | IL_S vs Clear | 2 | 3 | 0.009** | 3 | 16 | 12 | -2.622 | -0.333 |
| Q14 | VO2_U vs Clear | 4 | 3 | 0.005** | 18 | 5 | 8 | -2.811 | -0.357 |
| | IL_S vs Clear | 2 | 3 | 0.004** | 3 | 17 | 11 | -2.899 | -0.368 |
| | IL_S vs IL_U | 2 | 3 | 0.003** | 2 | 16 | 13 | -2.982 | -0.379 |
| | IL_S vs VO2_U | 2 | 4 | 0.000*** | 3 | 24 | 4 | -4.139 | -0.526 |
| | VO2_S vs VO2_U | 3 | 4 | 0.003** | 3 | 16 | 12 | -3.013 | -0.383 |
| | VO2_S vs IL_S | 3 | 2 | 0.005** | 20 | 5 | 6 | -2.819 | -0.358 |
| Q15 | IL_S vs Clear | 2 | 3 | 0.001*** | 4 | 20 | 7 | -3.312 | -0.421 |
| | IL_S vs IL_U | 2 | 3 | 0.000*** | 1 | 20 | 10 | -3.908 | -0.496 |
| | IL_S vs VO2_U | 2 | 4 | 0.000*** | 5 | 23 | 3 | -3.741 | -0.475 |
| | VO2_S vs VO2_U | 3 | 4 | 0.009** | 3 | 15 | 13 | -2.6 | -0.330 |
| | VO2_S vs IL_S | 3 | 2 | 0.014* | 18 | 7 | 6 | -2.455 | -0.312 |

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

**(B)** Wilcoxon signed-rank test of responses to questions in questionnaire Part II with significant results produced under the five studied lighting conditions

| Questions | Hypothesis ($M_1$ vs $M_2$) | $M_{1dn}$ | $M_{2dn}$ | *p*-value | Positive | Negative | Ties | *Z* | Effect Size (*r*) |
|---|---|---|---|---|---|---|---|---|---|
| Q16 | IL_S vs IL_U | 3 | 3 | 0.009** | 3 | 16 | 12 | -2.622 | -0.333 |
| | IL_S vs $VO_2$_U | 3 | 3 | 0.052 | 7 | 15 | 9 | -1.941 | -0.247 |
| Q20 | $VO_2$_U vs Clear | 3 | 3 | 0.004** | 15 | 3 | 13 | -2.853 | -0.362 |
| | $VO_2$_S vs Clear | 4 | 3 | 0.018 | 18 | 6 | 7 | -2.373 | -0.301 |
| | IL_S vs IL_U | 3 | 4 | 0.000*** | 1 | 16 | 14 | -3.578 | -0.454 |
| | IL_S vs $VO_2$_U | 3 | 3 | 0.003** | 4 | 17 | 10 | -2.992 | -0.380 |
| | $VO_2$_S vs IL_S | 4 | 3 | 0.003** | 15 | 4 | 12 | -2.941 | -0.374 |

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

**(C)** Plots showing the evaluations made by subjects for the questions found in Table 6.