

# Repeatability of Colour Matching Tests using Psychophysical Methods.

J G J van der Westhuyzen<sup>1</sup>

Department of Electrical, Electronic & Computer Engineering, University of Pretoria,  
University Road, Pretoria 0002, South Africa

Received 10 October 2018, Accepted 02 March 2019

## Abstract

The researcher working in the field of illumination can use a number of measurement methods, depending on the requirement. Many practical evaluation methods employ varieties of a configuration where test stations are positioned next to or opposite each other for comparison purposes. Test stations can consist of test booths or even full-sized test rooms when using indoor evaluations. The number of test booths for indoor application may differ from one to three. Most colour and light perception studies thus depend on human observers to provide some feedback or input. It is the human observer who uses illumination devices and it is important that human response to these products be measured. Results achieved from these measurements can, in turn, be used by the lighting designer to optimize the design of illumination products. Many studies in the field of colour perception and/or matching rely on the method of “psychophysics” where humans are used. Reliability of results achieved when using human observers can be questioned. Results achieved with psychometric testing cannot be compared with an established benchmark and/or standard as such a benchmark was itself generated using human observers, endangering the researchers to become entangled in a closed circle of cause and effect where one depends on the other. This study investigated the reliability of human observers when studying colour perception. Repeatability of results achieved when using human observers was measured. A test was repeated using the same set of observers. Even this approach is not safe as human memory may serve the observer so well, that results can be selected, or at least influenced, on the basis of memory and not perception. The solution was to implement extended time between measurements, thereby hoping that sufficient time will lapse to erase details from the observer’s memory. A colour matching task had to be completed to test observers of two distinct age groups. One group was older than 50 years of age and the second group was younger than 40 years of age. The two groups were tasked to complete the same colour matching test at the beginning of a year and after one year. Test procedures and test equipment were identical before and after the year.

*Keywords: colour perception, test booth, LED, light source, SPD, CRI R<sub>a</sub>, chromaticity coordinates.*

## 1 Introduction

The purpose of artificial lighting is to enhance and support the performance of visual tasks. The exact influence of light on many human activity parameters is indeed difficult to measure. It has been established that humans will choose to work or sit next to a window with an outside view and natural lighting when given a choice. [1] Light influences all aspects of daily life and work for humans. Pechacek et al. [2] state that light affects the physical, physiological and psychological well-being of human beings. The designer of artificial light sources should thus take human factors into consideration.

The use of test booths as a research tool that allows light and vision parameters to be evaluated is well established. Various designs have been implemented for test set-ups in booth format. Single booths [3] [4] have been used, as well as triple booths. [5] [6] It appears, however, that the double-booth configuration is the more popular tool for researchers to investigate diverse parameters such as colour rendering [7] [8], brightness matching [9] [10], white perception [11] [12], actual light sources [13] and light level requirements.

<sup>1</sup> J.G.J van der Westhuyzen, E-mail address: koosvdw47@gmail.com

Although researchers use observers of different ages, the use of younger people is favored in many studies. [3] [5] [7] [13] It is common to use students younger than 30 years of age. [4] [10] [12] With so many researchers using the double-booth method for psychophysical evaluations, it is thus logical to investigate the repeatability of psychophysical evaluations regarding the specific double-booth test configuration.

## 2 Test Layout

### 2.1 Equipment, construction and field of view.

A conventional double-booth test setup was used for psychophysical evaluation studies. In double-booth design, one booth is typically fitted with a “reference” light source and the other with a light source called the “test source”. Some parameters of the test source can be adjusted and compared to similar parameters from the reference source. The double-booth design used in this study is similar to that used by Narendran et al [7], Sandor et al [8], Spaulding et al [9], Thompson et al [11], Wei et al [10], Fotios et al [13] and Houser et al [12].

Side-by-side test booths are often constructed in cubicle form, which may vary in size, ranging from 380 mm to 660 mm cubicles. [7], [10], [12], [13]. Similar to this study, the observer is positioned in front of the double booth and the distance is selected to ensure a particular field of view (FOV). Figure 1 shows a schematic drawing of the double booth used in this study.

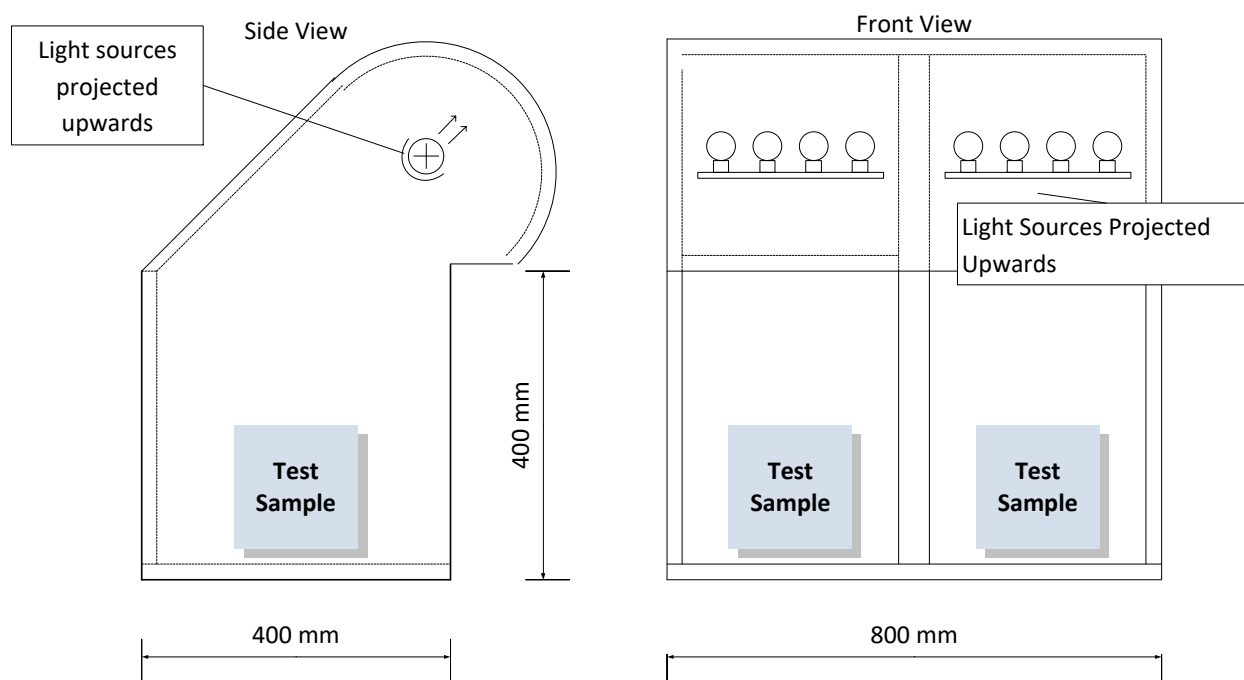


Figure 1. Double-booth design with domed roof. Light sources project towards roof. Both test and reference light sources transmit towards the dome, where diffuse reflections ensure constant luminance of ambient light reaching the test charts. (The two sources can be exchanged.)

Figure 2 shows the double-booth as seen by an observer.

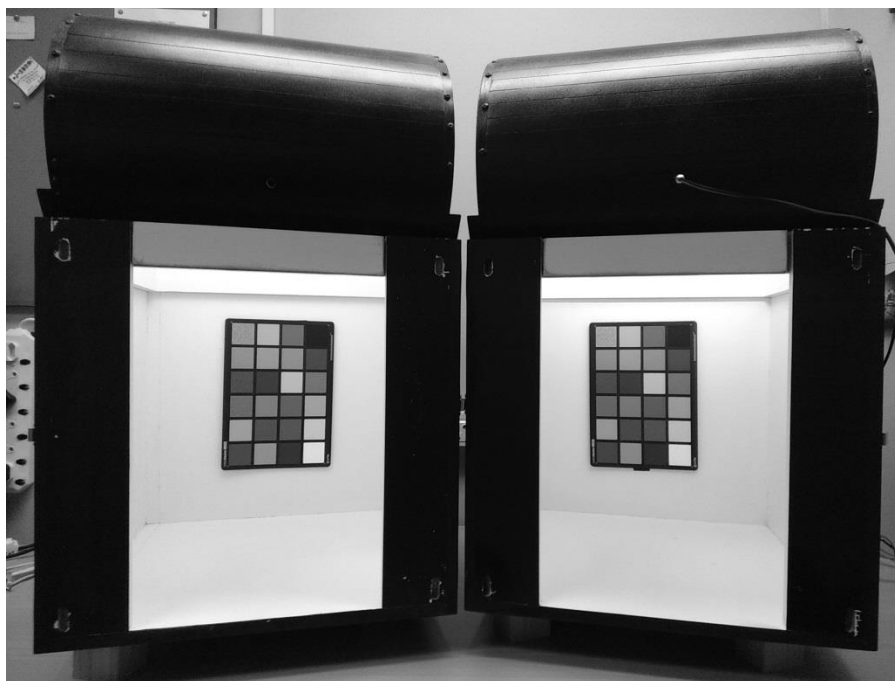


Figure 2: Double booths as seen by an observer. The MacBeth Colour Checker charts can be seen mounted on the rear vertical plane. Illumination is provided by light reflected from the rear-facing domes mounted on top.

A number of researchers chose booth dimensions and viewing distances such that a field of view (FOV) of around 40° is achieved. [7], [10], [12], [13]. The FOV used in the booth design, as repeated here, is calculated to be 43°.

The inside surface of the test booth is covered with white paint, which was measured to feature a reflectivity of better than 90% from 420 to 800 nm. It is necessary to ensure that reflectivity is spectrally constant and that surface characteristics remains unchanged for all measurements. The reflectivity graph is shown in Figure 3.

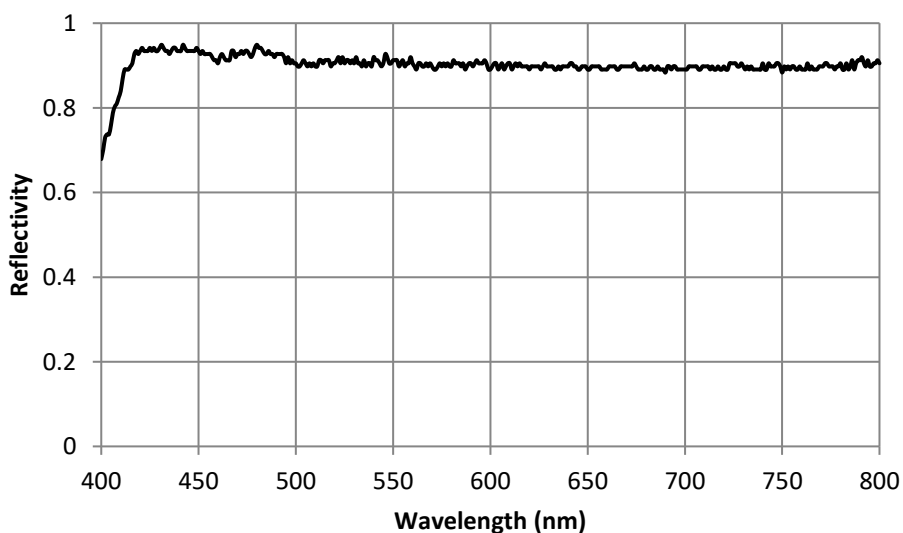


Figure 3. Inside test booth surface spectral reflectivity. (Reflectivity values are dimensionless.)

## 2.2 Evaluation objects

The CIE developed specific test samples in the form of coloured charts to be used in the evaluation of colour perception. Unfortunately, the original products are no longer available. Due to this fact, the CIE Technical Committee 1-33 recommended that the Macbeth Color Checker (MCC) be used for visual experiments [14]. For the purpose of this test, X-Rite Classic Color Checker charts were used as can be seen in Figure 2.

## 2.3 Light source

For the purpose of this evaluation, two types of light sources were used. A different type of source was fitted into each of the booth's chambers. A reference source was fitted in one chamber and a test source in the second chamber.

The reference source consisted of a conventional 12 V, 50 W halogen lamp. The driving current was controlled and could be adjusted from 2,4 A to 3,4 A. In this range, the average Colour Rendering Index (CRI  $R_a$ ) was measured to remain at 99. The correlated colour temperature (CCT) moved from 2300 K to 2500 K. The luminance level, measured at the colour sample, for this source was set at 150 lux. This level is similar to that specified by Moore et al for colour sample inspection and office illumination. [15] Moore et al presented results that showed that most workers prefer to work in an area where light levels range from 100 to 300 lux. This includes workstation and desk areas. Luminance levels used in similar double-booth studies are 200 lux [7], 300 lux [12] 260 lux [9], and 2- 15 lux. [14] Table 1 shows measurement results of the reference source (halogen lamp).

Table 1. Reference source parameters with changing current. Both the average colour rendering value  $R_a$  and the specific colour sample value  $R_9$  remain at 99. These values are thus not influenced by a change in driving current.

Current	$R_a$	$R_9$	CCT
3,4 A	99	99	2541 K
3,2 A	99	99	2494 K
3,0 A	99	99	2455 K
2,8 A	99	99	2413 K
2,6 A	99	99	2370 K
2,4 A	99	99	2319 K

The test source consists of four light-emitting diodes (LEDs). LED's were selected to cover the spectral region evenly and to feature a wide field ( $120^\circ$  inclusive angle minimum) of transmission in this application.

Selection of the tetrachromatic components of the test source is similar to that of other researchers. [16] For the selection of tetrachromatic source components, Zukauskas and Vaicekauskas describe optimization methods. [16] They suggest four coloured LED's with peak wavelengths in the following sectors: 410 – 490 nm, 490 – 540 nm, 540 – 610 nm and 610 – 680 nm. Table 2 shows the final selection of wavelengths together with my selection.

Figure 4 shows the three basic colours (RGB) human eye cone response curves. Tetrachromatic LED source responses, as used in this study, are overlaid on the cone response curves. LED wavelengths should be selected to be close to cone peak sensitivity. The yellow LED thus contributes well to the tetrachromatic wavelength spread as peak emission is close to the red cone sensitivity peak.

Table 2. Selection of LED wavelengths for tetrachromatic light source.

LED source colour	Zukauskas & Vaicekauskas [16]	My selection of wavelengths.
Blue	452 nm	455 nm
Green	523 nm	528 nm
Yellow	589 nm	590 nm
Red	637 nm	657 nm

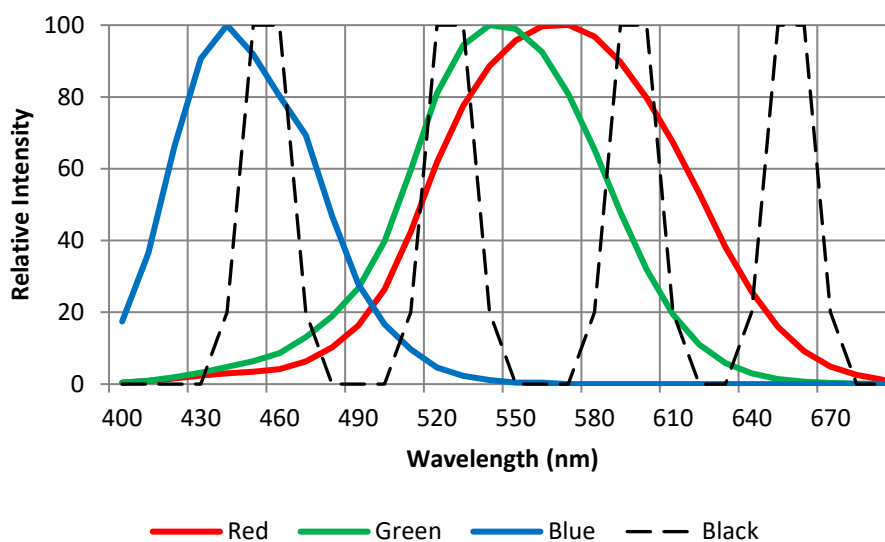


Figure 4: Eye response curves for red, green and blue. The black dotted line shows the wavelength areas covered by the tetrachromatic source in this study. These curves do not show relative cone sensitivity but rather spectral areas where cone colour sensitivity reside. (Colour cone curves were calculated using the procedure as described by DJ Heeger. [17])

Doing two sets of identical experiments one year apart means that variables must be minimized. The exact test instruments were used as well as the same set of LED's which formed part of the tetrachromatic illumination source. Stability of light output was ensured by mounting the LEDs on heat sinks as prescribed by the manufacturer. [18] (Heat sink temperature was monitored.) Manufacturer data regarding the stability of light output over time shows that a change in LED light output can be expected from 20 000 hours of usage and further. [19]. Total LED on-time for the complete set of experiments is not expected to be more than 50 hours. Continuous high currents affect the time that an LED can deliver stable light output negatively. During this study, peak currents never exceeded 50% of the specified maximum current. Because observer groups are tested concurrently, any possible degradation in LED performance will affect both groups in the same way and no group will thus be at an advantage/disadvantage. LED intensity and wavelength characteristics were measured before and after the test.

#### 2.4 Observers

All observers were subjected to an Ishihara coloured plate screening test in order to ensure basic colour differentiation ability. Only observers who scored 100% in the Ishihara test were used in this study. (The number of observers subjected to the screening test is thus more than the number of observers actually used.)

Persons younger than 16 years of age were also not used.

Israel [20] states that the sample sizes of similar studies should be considered to provide guidance regarding sample size numbers. Some similar double-booth psychophysical studies that can be compared are those of Narendran et al (two groups consists of 20 and 10 observers) [7], Sandor and Schanda (one group consisting of 10 observers) [8], Spaulding et al (one group with 13 observers) [9], Thompson et al (one group consisting of 10

observers) [11], Wei et al (one group consisting of 48 observers) [10], Fotios and Cheal (one group with 21 observers) [13] and Houser et al (one group with 40 observers) [12].

This study used 41 observers, which were divided into two age groups. The first group consisted of people 40 years old and younger (18 people with mean age 30,4 and a standard deviation of 5,04). The second group consisted of people 50 years and older (23 people with mean age 57,3 and a standard deviation of 4,6). It is desirable to use observers spanning the complete human age range. This is logistically different and may be one of the reasons why many researchers in the field use younger people as observers. Examples of these are Wei et al [10], Houser et al [12] and Rea and Freyssinier [4]. When older observers are included, a number of researchers used a younger group with a mean age of 55-57 years of age, for example Dangol et al [5], Islam et al [6], Elliot et al [21] and Scheffrin and Werner [22].

Observers used in this study are experienced in the fields of electro-optics, radiometry and basic photometry. They are not experienced in the field of colorimetry. Some recognized the 1931 CIE chromaticity diagram but did not know what it means. The name "Munsell" was unfamiliar to all.

### 3 Methodology

#### 3.1 Method

Observers were informed about the procedure and applicable safety aspects. All participants signed a consent form. The research was approved by the Faculty Committee for Research Ethics and Integrity of the Faculty of Engineering, Built Environment and Information Technology of the University of Pretoria.

The tetrachromatic LED (test source) and the tungsten (reference source) illumination sources were allowed to stabilize for 10 minutes prior to commencement of the test procedure. The observer looks at the MCC (Figure 2) which consists of 24 coloured blocks. The test booth containing the test source is positioned on the right-hand side of the observer's view, while the test booth containing the reference source is positioned on the left-hand side of the observer's view.

The test procedure is as follows:

- (a) Test source LED's are switched on and set at minimum currents of 20mA each. The test source illuminates an MCC inside the test illumination chamber. (Figure 2)
- (b) The reference source is adjusted to a level of 150 lux. (Usually at 3 A) The reference source illuminates an MCC inside the reference illumination chamber. (Figure 2)
- (c) The room is darkened and the observer is allowed 10 minutes to adjust to the visual environment. The illuminance level in the dark room is between 1 and 1,2 lux.
- (d) The test commences with the observer looking into the reference chamber and considers the coloured blocks of the MCC.  
The X-Rite Macbeth Color Checker (MCC) is mounted vertically against the test chamber wall.
- (e) The observer looks into the chamber with the test source and considers the coloured blocks of the MCC.
- (f) The observer adjusts the currents through the four coloured LED sources inside the test booth until the colours on the reference MCC is perceived to match the colours on the test MCC.. The observer thus adjusts the four LED sources (red, green, blue, yellow-orange) until the reflection from the reference MCC matches the reflection of the test MCC. The observer may re-adjust LED source currents as many times as required. No time limit is enforced.
- (g) The facilitator notes the test source inputs and measures the following:
  - i. Light level (lux) as measured at the MCC,
  - ii. Forward current through LED tetrachromatic test source.
- (h) The researcher mounts the tetrachromatic LED source in an integrating sphere as shown in Figure 5. The test method used for the measurement of photometric as well as electrical characteristics is based on the "Approved Method: Electrical and Photometric Measurements of Solid-State Lighting Products" by the Illuminating Engineering Society (IES) (Illuminating Engineering Society 2008). The IES method (paragraph 9.1, page 4) prescribes the use of a single integrating sphere and an internal baffle is used to screen direct illumination from the entrance of the radiometer. The method used in this study uses a dual integrating sphere where the source is positioned inside the primary sphere and the entrance of the

radiometer is aimed at the secondary sphere. The primary and secondary spheres are constructed as a unit, as shown in Figure .

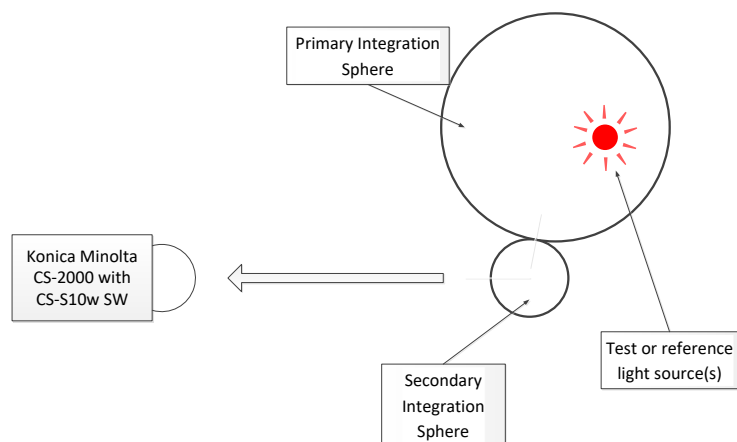


Figure 5. Test setup used to measure photometric data of the source selected by the observer.

- (i) The following parameters are measured and recorded.
- i. Correlated Colour Temperature (CCT),
  - ii. General Colour Rendering Index (CRI) value ( $R_a$ ),
  - iii. CRI values ( $R_1$  to  $R_{15}$ ),
  - iv. Values  $x$  and  $y$ ,
  - v. Spectral Power Distribution (SPD), and
  - vi. Duv values.

### 3.2 Measurement equipment

All photometric measurements were completed using a Konica Minolta CS-2000 photo radiometer with custom CS-S10W software. A calibration certificate was valid at the time of testing.

## 4 Results

The specifications of an illumination source selected by two age groups to match colours are presented. Observers considered colour reflection from the MCC and adjusted the composite tetrachromatic light source to match the colours in the “test” booth with those in the “reference” booth.

### 4.1 Photometric data

Detailed information regarding the spectral composition of the two different sources selected by two age groups can be extracted by considering some photometric parameters. The CIE specified test colour samples which were Munsell colour samples. These values are the approximate Munsell values for calculating CRI although only the first eight values are used for calculating the general colour rendering index  $R_a$ . This section expands on results achieved.

The Spectral Power Distribution (SPD) curves as colour matched, for the under 40 group, before and after a year is shown in Figure 6 and Figure 7.

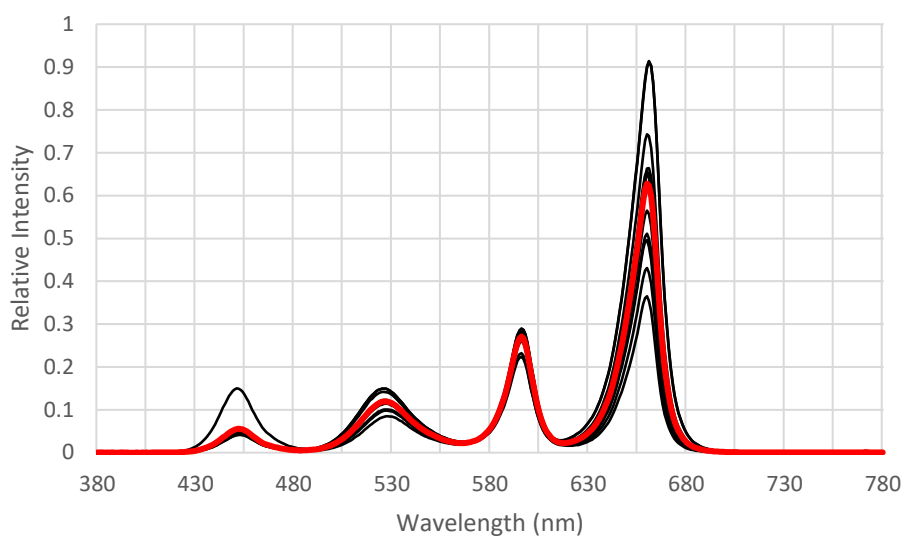


Figure 6. SPD for a matched colour source for observer group younger than 40 and at the beginning of the test, before the year started.

The red line shows the average value.

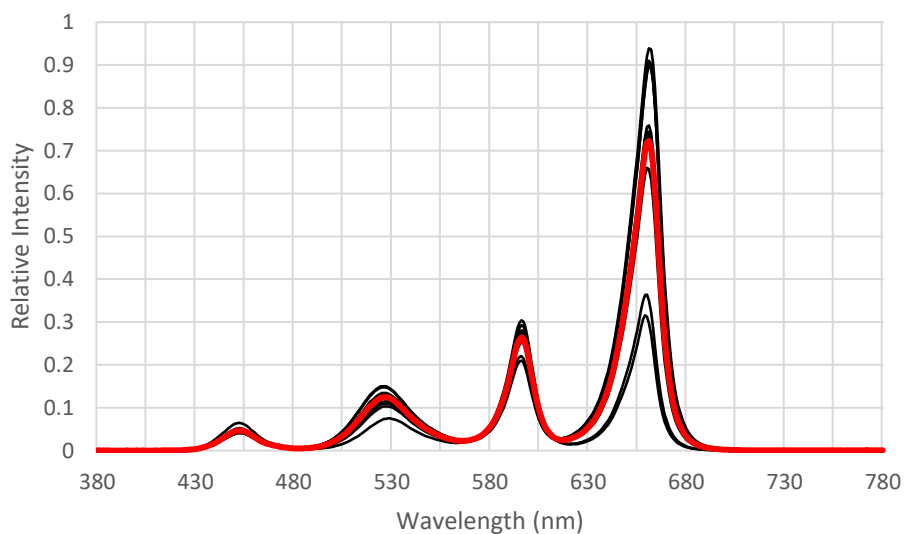


Figure 7. SPD for a matched colour source for observer group younger than 40 and after completion of the year.

The red line shows the average value.

The Spectral Power Distribution (SPD) curves as colour matched, for the over 50 group, before and after a year is shown in Figure 8 and Figure 9.



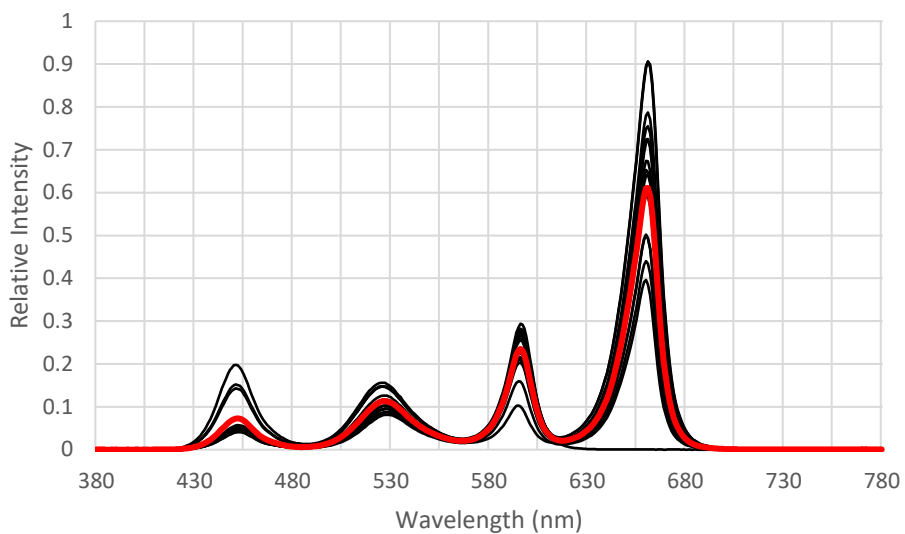


Figure 8. SPD for a matched colour source for observer group older than 50 and at the beginning of the test, before the year started.

The red line shows the average value.

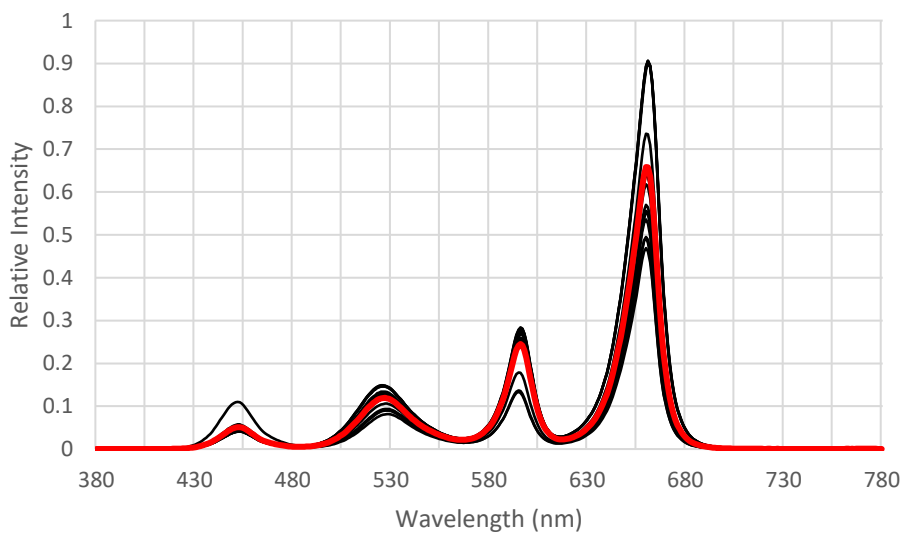


Figure 9. SPD for a matched colour source for observer group older than 50 and after completion of the year.

The red line shows the average value.

Average SPD graphs for the younger than 40 group before and after a year is shown in Figure 10.

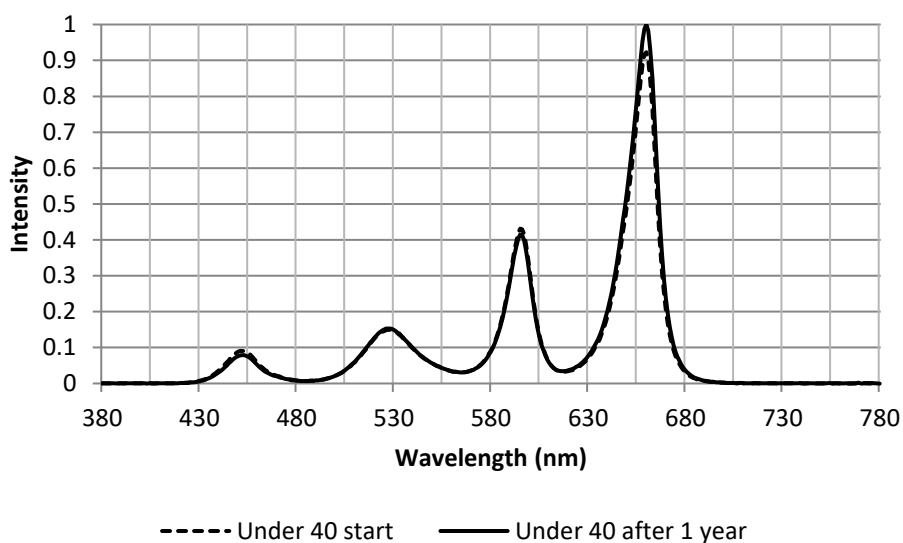


Figure 10. SPD of the reference source as selected by the younger than 40 group before and after a year.

(Average values) Differences at tetra-chromatic source wavelengths are 12.1% at 461 nm, 2% at 537 nm, 3.4% at 590 nm and 7% at 645 nm. (Intensity values are normalized and not absolute.)

The SPD graph of the older than 50 group before and after a year is shown in Figure 11.

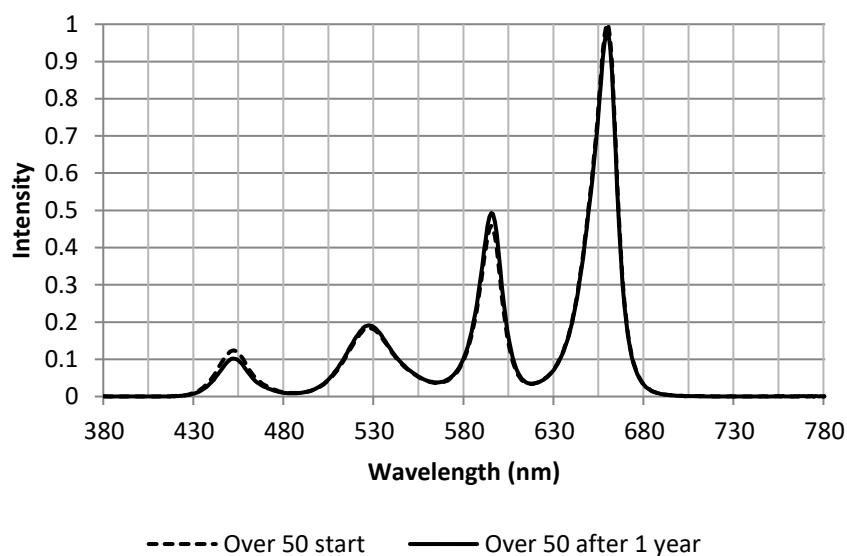


Figure 11. SPD of the reference source as selected by the older than 50 group before and after a year.

(Average values) Differences at tetra-chromatic source wavelengths are 15.3% at 461 nm, 1.9% at 537 nm, 5.7% at 590 nm and 1.9% at 645 nm. (Intensity values are normalized and not absolute.)

Detailed information on the spectral composition of the different sources selected by the two age groups can be extracted by considering photometric parameters. The CIE [23] specified test colour samples that are Munsell colour samples. These values are the approximate Munsell values for calculating Colour Rendering Index (CRI), although only the first eight values are used for calculating general CRI,  $R_a$ . Table 3 presents detailed results.

Table 3. Parameters of illumination sources as selected by two different age groups to match colours in the “test”booth with those in the “reference” booth.

Parameter	Younger than 40		Older than 50	
	Year start	After 1 year	Year start	After 1 year
R <sub>a</sub> (General colour rendering index)	92	90	91	93
CCT (Correlated Colour Temperature)	2331 K	2300 K	2563 K	2545 K
X coordinate	0.477	0.4845	0.4563	0.467
Y coordinate	0.3922	0.3984	0.3884	0.4032
Duv value	-0.005	-0.0027	-0.007	-0.002

The x and y coordinates of the two sources selected for matching can be presented on a CIE chromaticity diagram. Figure 12 and Figure 13 shows CIE 1931 chromaticity diagrams with MacAdam ellipses as completed with ColorCalculator software. [24]

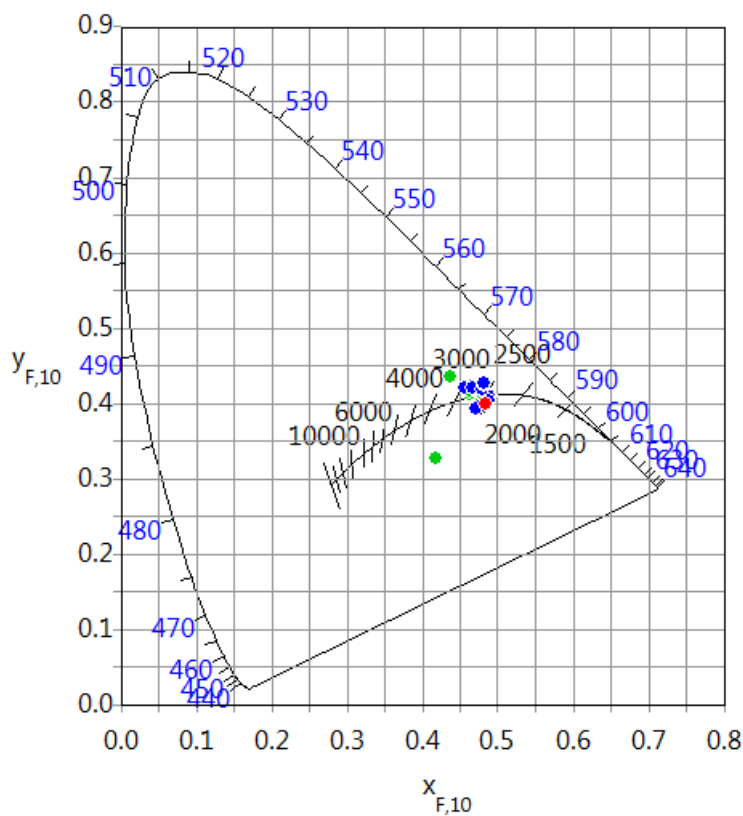


Figure 12: CIE 1931 chromaticity diagram for the under 40 observer group. The green dots show the distribution at the beginning of the year. The blue dots show the distribution at the end of the year with the red dot showing the average value.

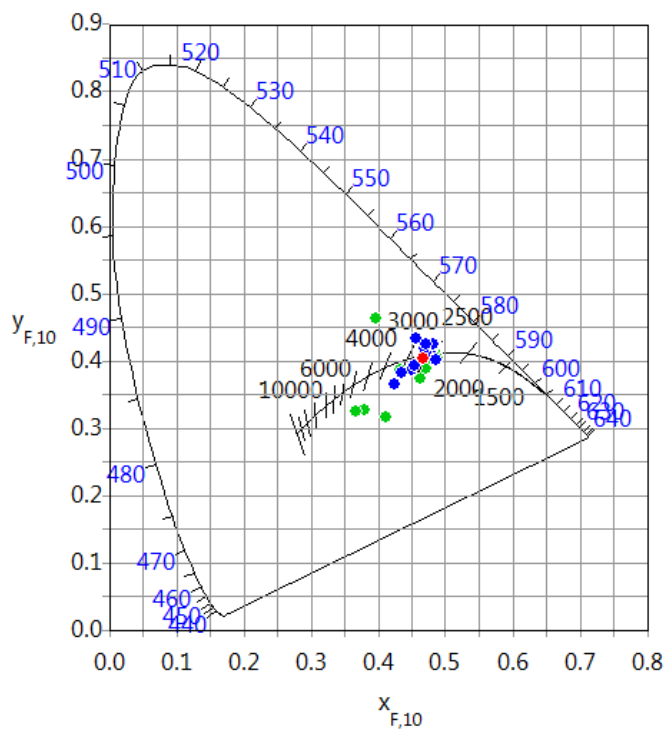


Figure 13: CIE 1931 chromaticity diagram for the older than 50 observer group. The green dots show the distribution at the beginning of the year. The blue dots show the distribution at the end of the year with the red dot showing the average value.

Figure 14 shows the average values of the x and y coordinates of the matched sources selected by the group younger than 40 and the group older than 50 before and after a year.

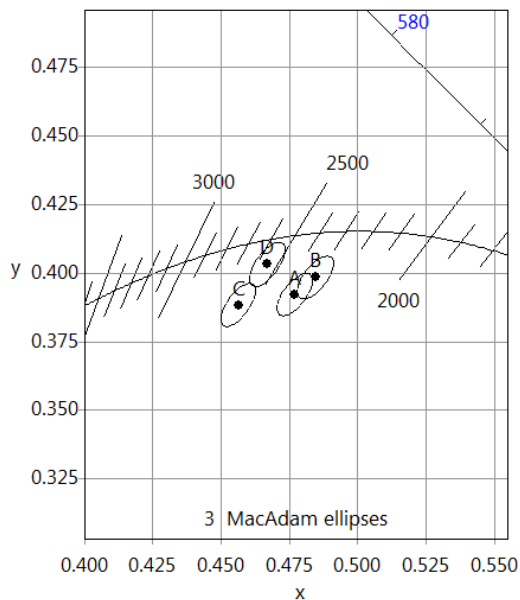


Figure 14. Sectional 1931 CIE chromaticity diagram showing enlarged x and y average coordinate values with three-step MacAdam ellipses.

(A=under 40 year start, B=under 40 year end, C=Over 50 year start, D=over 50 year-end)

## 4.2 Statistical analysis

Nonparametric tests were completed using the Independent-Sample Mann-Whitney U test.

The following parameters were statistically evaluated:

- The general colour rendering index number  $R_a$  between the beginning and after one year for each age group.
- The x coordinates of sources as selected by two age groups at the beginning of the test year.
- The y coordinates of sources as selected by two age groups at the completion of the test year.
- The Colour Rendering Index (CRI) number  $R_9$  (strong red) before and after the test year for both age groups.
- The CRI number  $R_{10}$  (Strong yellow) before and after the test year for both age groups.
- The CRI number  $R_{11}$  (Strong green) before and after the test year for both age groups.
- The CRI number  $R_{12}$  (Strong blue) before and after the test year for both age groups.

The following hypotheses were tested:

- H01: The distribution of General Colour Rendering Index value  $R_a$  is the same before and after the test year for the observer group younger than 40 years of age.
- H02: The distribution of the 1931 CIE chromaticity diagram x coordinate is the same before and after the test year for the observer group younger than 40 years of age.
- H03: The distribution of the 1931 CIE chromaticity diagram y coordinate is the same before and after the test year for the observer group younger than 40 years of age.
- H04: The distribution of CRI colour strong red ( $R_9$ ) is the same before and after the test year for the age group younger than 40.
- H05: The distribution of CRI colour strong yellow ( $R_{10}$ ) is the same before and after the test year for the age group younger than 40.
- H06: The distribution of CRI colour strong green ( $R_{11}$ ) is the same before and after the test year for the age group younger than 40.
- H07: The distribution of CRI colour strong blue ( $R_{12}$ ) is the same before and after the test year for the age group younger than 40.
- H08: The distribution of General Colour Rendering Index value  $R_a$  is the same before and after the test year for the observer group older than 50 years of age.
- H09: The distribution of the 1931 CIE chromaticity diagram x coordinate is the same before and after the test year for the observer group older than 50 years of age.
- H10: The distribution of the 1931 CIE chromaticity diagram y coordinate is the same before and after the test year for the observer group older than 50 years of age.
- H11: The distribution of CRI colour strong red ( $R_9$ ) is the same before and after the test year for the age group older than 50.
- H12: The distribution of CRI colour strong yellow ( $R_{10}$ ) is the same before and after the test year for the age group older than 50.
- H13: The distribution of CRI colour strong green ( $R_{11}$ ) is the same before and after the test year for the age group older than 50.
- H14: The distribution of CRI colour strong blue ( $R_{12}$ ) is the same before and after the test year for the age group older than 50.
- (All references to CRI numbers regarding colours refers to CRI test sample numbers [25])

Table 4 presents a summary of statistical results.

Table 4. Summary of statistical data. (Statistical calculator used was on-line available from Statistics Kingdom [26])

<b>Null Hypothesis</b>	<b>Mann-Whitney U-value</b>	<b>Exact significance (2-sided test, p-value)</b>	<b>Decision</b>
H <sub>01</sub>	105.5	0.0518	Retain the null-hypothesis.
H <sub>02</sub>	52	0.2656	Retain the null-hypothesis.
H <sub>03</sub>	68.5	0.8428	Retain the null-hypothesis
H <sub>04</sub>	70	0.0077	Reject the null-hypothesis
H <sub>05</sub>	57	0.1614	Retain the null-hypothesis
H <sub>06</sub>	56	0.1902	Retain the null-hypothesis
H <sub>07</sub>	53	0.2973	Retain the null-hypothesis
H <sub>08</sub>	97.5	0.5392	Retain the null-hypothesis
H <sub>09</sub>	96.5	0.5124	Retain the null-hypothesis
H <sub>10</sub>	76	0.1369	Retain the null-hypothesis
H <sub>11</sub>	91.5	0.3892	Retain the null-hypothesis
H <sub>12</sub>	85	0.2670	Retain the null-hypothesis
H <sub>13</sub>	124	0.6528	Retain the null-hypothesis
H <sub>14</sub>	80	0.1872	Retain the null-hypothesis

The null-hypothesis is retained in all cases except for H<sub>04</sub> where the p-value is 0.007. The p value of 0.007 is smaller than 0.050 which means that the difference in R<sub>9</sub> value before and after the test year is significant for the under 40 age group.

4.3 Complementary statistical analysis

The Student t probability mass function (PMF) is shown in Figure 15 for the group younger than 40 years of age and in Figure 16 for the group older than 50 years of age.

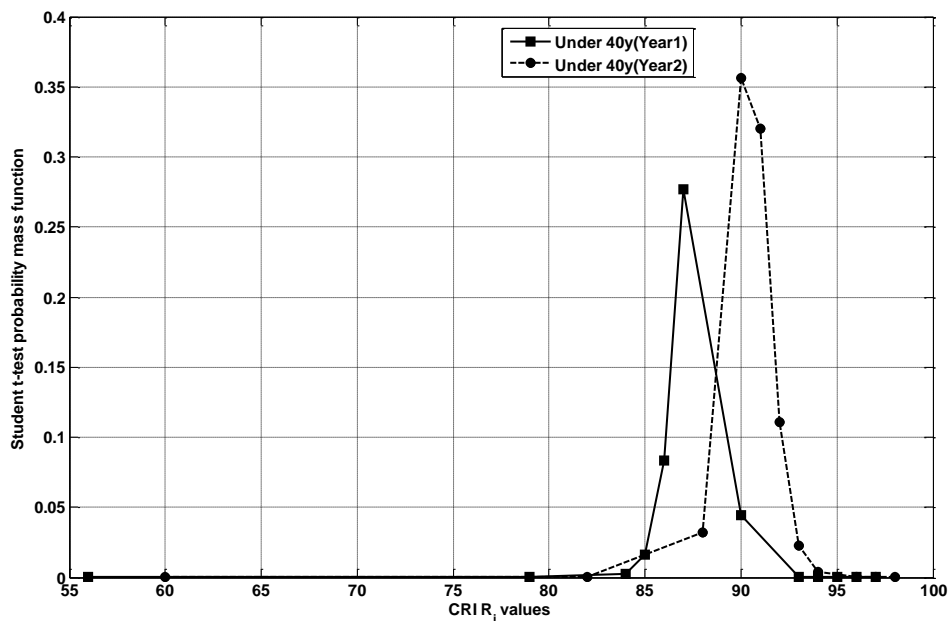


Figure 15: Student t probability mass function as a function of sampled CRI  $R_i$  values for the age group younger than 40 and evaluated one year apart.

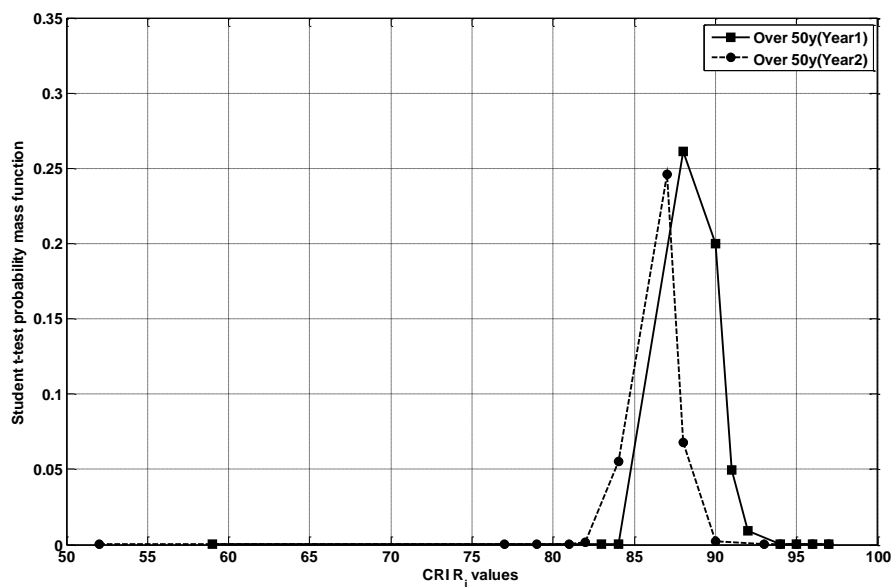


Figure 16: Student t probability mass function as a function of sampled CRI  $R_i$  values for the age group older than 50 and evaluated one year apart.

Figure 15 and Figure 16 display the Student t PMF as a function of the sampled CRI  $R_i$  values for both age groups and with observer group evaluations one year apart. The x-axis shows CRI  $R_i$  values. These values are the detail CRI values, which consist of a set of values  $R_i$ , which ranges from  $R_1$  to  $R_{15}$ . The sources selected by the two groups of observers were measured using a spectrophotometer. Some of the results measured by the spectrophotometer are a set of results that yields the CRI  $R_i$  values. The values can range from 0 to 100, although negative values are a possibility for specific colors. These values are drawn in Figure 15 and Figure 16 for each group of observers.

Figure 15. The solid line shows the results for the under-40 group at the beginning of the year. A variation of  $R_i$  value from 51 to 91 was measured. The dotted line shows the results after one year with a variation from 52 to 97. These are not  $R_a$  values, as only the first eight  $R_i$  values are used to calculate  $R_a$ . The two  $R_a$  values were measured at 92 and 90 for the beginning and end of year measurements. When a specific  $R_i$  value is chosen repeatedly, the value on the y-axis increases.

Figure 16. The solid line shows the results for the over-50 group at the beginning of the year. A variation of  $R_i$  value from 79 to 97 was measured. The dotted line shows the results after one year with a variation from 60 to 98. These are not  $R_a$  values, as only the first eight  $R_i$  values are used to calculate  $R_a$ . The two  $R_a$  values were measured at 91 and 93 for the beginning and end of year measurements. When a specific  $R_i$  value is chosen repeatedly, the value on the y-axis increases.

These graphs are known as the PMF. (The probability density function is related but uses continuous instead of discrete values) The graphs thus show the point values on the  $R_i$  scale where a specific group of observers is likely to select a source with values in that area. The probability of a specific observer group selecting a source with specific  $R_i$  values can be read from Figure 15 and Figure 16. When considering Figure 15 and reading the "Year 2" graph, it can be established that the probability of the group younger than 40 to select an  $R_i$  value of 90 is 0.35 (or 35%).

## 5 Discussion

Two observer groups selected similar light sources to match colours when reflectance from an MCC test chart is considered before and after one year.

Houser et al [27] state that differences in CRI  $R_a$  values of less than 5 points are mostly not noticeable. The difference in  $R_a$  values for both observer age groups after one year is 2, which is thus not significant. When considering specific colour rendering values  $R_9$ ,  $R_{10}$ ,  $R_{11}$  and  $R_{12}$  no statistical significant values were measured except where the  $R_9$  value difference before and after one year for the younger than 40 age group can be classified as significant. CCT values for both age groups remain similar after one year.

When considering MacAdam ellipses, lamp manufacturers are advised by ANSI to stay within the boundary of a four-step ellipse [28]. The outline of the three-step ellipses overlap for the younger age group and just touches for the older age group. The average observer's vision will perceive the two sources as chromatically similar as measured before and after the "test" year, for both age groups.

Duv values is important as this is applicable to light source quality. [29]. Matching source values for both age groups are below the Planckian locus (before and after one year). Padfield states that Duv values with a value larger than 0.006 are not ideal as "white light". The ideal Duv value for "white light" should be less than 0.001 [30]. Considering these parameters, sources selected by the two observer groups can thus not be classified as being ideally "white", neither before nor after one year. The difference in Duv values for the older group was 0.005 after one year, whereas that for the younger group was 0.0023 after one year.

Dorcus [31], Ou et al [32] and Taylor et al [33] investigated human colour perception with regard to age. These studies considered the possible change of colour perception with regard to age without reporting the repeatability of data when using psychometric colour perception experiments. Dorcus found that different age groups have the same colour perception whereas Ou et al came to the conclusion that colour emotion may indeed change as people grow older. It is thus possible that the one year time laps between experiments, as applied in this study, may not be sufficient to measure a change in relative colour perception, using the identical group of observers.



## 6 Conclusion

Repeatability of psychophysical colour matching tests was investigated over a period of one year. Two age groups were used as observers. The investigation showed that human observers have a remarkable ability to achieve similar and consistent results after one year which makes psychometric test procedures a reliable option for colour perception/measurement procedures.

## References

- [1] Aries, M. B., Aarts, M. P., & van Hoof, J. (2015). Daylight and health: A review of the evidence and consequences for the built environment. *Lighting Research & Technology*, 47(1), 6-27.
- [2] Pechacek, C. S., Andersen, M., & Lockley, S. W. (2008). Preliminary method for prospective analysis of the circadian efficacy of (day) light with applications to healthcare architecture. *Leukos*, 5(1), 1-26.
- [3] Ohno, Y., & Fein, M. (2013) Vision Experiment on White Light Chromaticity for Lighting. [Online]. Retrieved February 23, 2015. <http://cltc.ucdavis.edu/sites/default/files/files/publication/2-yoshi-ohno-mira-fein-white-light-chromaticity-vision-experiment.pdf>
- [4] Rea, M. S., & Freyssinier, J. P. (2010). Color rendering: Beyond pride and prejudice. *Color Research & Application*, 35(6), 401-409.
- [5] Dangol, R., Islam, M., LiSc, M. H., Bhusal, P., Puolakka, M., & Halonen, L. (2013). Subjective preferences and colour quality metrics of LED light sources. *Lighting Research & Technology*, 45(6), 666-688.
- [6] Islam, M. S., Dangol, R., Hyvärinen, M., Bhusal, P., Puolakka, M., & Halonen, L. (2013). User preferences for LED lighting in terms of light spectrum. *Lighting Research & Technology*, 45(6), 641-665..
- [7] Narendran, N., & Deng, L. (2002, November). Color rendering properties of LED light sources. In *Solid State Lighting II* (Vol. 4776, pp. 61-68). International Society for Optics and Photonics.
- [8] Sándor, N., & Schanda, J. (2006). Visual colour rendering based on colour difference evaluations. *Lighting Research & Technology*, 38(3), 225-239.
- [9] Spaulding, J. M., Thompson, M. R., & Levin, R. E. (2011, February). Human preference in tunable solid state lighting. In *Light-Emitting Diodes: Materials, Devices, and Applications for Solid State Lighting XV* (Vol. 7954, p. 795403). International Society for Optics and Photonics.
- [10] Wei, M., Houser, K. W., David, A., & Krames, M. R. (2015). Perceptual responses to LED illumination with colour rendering indices of 85 and 97. *Lighting Research & Technology*, 47(7), 810-827.
- [11] Thompson, M., O'Reilly, U. M., & Levin, R. (2007, September). Psychophysical evaluations of various color rendering from LED-based architectural lighting. In *Seventh International Conference on Solid State Lighting* (Vol. 6669, p. 66690Y). International Society for Optics and Photonics.
- [12] Houser, K. W., Wei, M., David, A., & Krames, M. R. (2014). Whiteness perception under LED illumination. *Leukos*, 10(3), 165-180.
- [13] Fotios, S. A., & Cheal, C. (2007). Evidence for response contraction bias in side-by-side matching tasks. *Lighting Research & Technology*, 39(2), 159-169.
- [14] McCamy, C. S., Marcus, H., & Davidson, J. G. (1976). A color-rendition chart. *J. App. Photog. Eng*, 2(3), 95-99.
- [15] Moore, T., Carter, D. J., & Slater, A. I. (2002). A field study of occupant controlled lighting in offices. *Lighting Research & Technology*, 34(3), 191-202.
- [16] ŽUKAUSKAS, A., & VAICEKAUSKAS, R. (2011). LEDs in lighting with tailored color quality. *International Journal of High Speed Electronics and Systems*, 20(02), 287-301.
- [17] Heeger, D. J. (2018, August). color-tutorial.xlsx-CNS@NYU. [Online]. Retrieved August 30, 2018. [www.cns.nyu.edu/~david/courses/perceptionAdvanced/Tutorials/color-tutorial.xlsx](http://www.cns.nyu.edu/~david/courses/perceptionAdvanced/Tutorials/color-tutorial.xlsx)
- [18] OSRAM. (2013). Thermal Management of Golden Dragon LED. [Online]. Retrieved July 18, 2017. [https://www.osram-os.com/Graphics/XPic8/00165240\\_0.pdf/Thermal%20Management%20of%20Golden%20DRAGON%20LED.pdf](https://www.osram-os.com/Graphics/XPic8/00165240_0.pdf/Thermal%20Management%20of%20Golden%20DRAGON%20LED.pdf)
- [19] OSRAM. (2014). Reliability of the DRAGON Product Family. [Online]. Retrieved July 18, 2017. [https://www.osram-os.com/Graphics/XPic8/00165204\\_0.pdf/Reliability%20of%20the%20DRAGON%20Product%20Family.pdf](https://www.osram-os.com/Graphics/XPic8/00165204_0.pdf/Reliability%20of%20the%20DRAGON%20Product%20Family.pdf)
- [20] Israel, G.D. (2013) University of Florida. [Online]. Retrieved November 27, 2017. <https://edis.ifas.ufl.edu/pdffiles/PD/PD00600.pdf>
- [21] Elliott, S. L., Hardy, J. L., Webster, M. A., & Werner, J. S. (2007). Aging and blur adaptation. *Journal of vision*, 7(6), 8-8.

- [22] Scheffrin, B. E., & Werner, J. S. (1993). Age-related changes in the color appearance of broadband surfaces. *Color Research & Application*, 18(6), 380-389.
- [23] Commission Internationale de l'Eclairage. (1995). Method of Measuring and Specifying Colour Rendering Properties of Light Sources: Technical Report: CIE 13.3-1995. CIE.
- [24] OSRAM Tools & Resources. [Online]. Retrieved May 28, 2018. [www.osram-americas.com/en-us/tools-and-resources/Pages/led-color-calculator.aspx](http://www.osram-americas.com/en-us/tools-and-resources/Pages/led-color-calculator.aspx)
- [25] Waveformlighting. (2018) [Online]. Retrieved September 2018 and April 4, 2019 <https://www.waveformlighting.com/tech/cri-ra-test-color-samples-tes>
- [26] Statistics Kingdom. (2018) [Online]. Retrieved September 21, 2018. [http://www.statskingdom.com/170median\\_mann\\_whitney.html](http://www.statskingdom.com/170median_mann_whitney.html)
- [27] Houser, K., Mossman, M., Smet, K., & Whitehead, L. (2015). Tutorial: Color Rendering and its Applications in Lighting. *Leukos: Journal of the Illuminating Engineering Society of North America*. 00, 1 - 20.
- [28] Sylvania. (2018) [Online]. Retrieved September 25, 2018. [http://www.usailighting.com/stuff/contentmgr/files/1/f386756b8bcf5da65a8d3c26a56b82e7/misc/macadam\\_faqs.pdf](http://www.usailighting.com/stuff/contentmgr/files/1/f386756b8bcf5da65a8d3c26a56b82e7/misc/macadam_faqs.pdf)
- [29] Ohno, Y. (2011, May). Calculation of CCT and Duv and practical conversion formulae. In *CORM Conference, May* (pp. 3-5).
- [30] Padfield, J.(2018). The National Gallery. [Online]. Retrieved May 8, 2018. <http://research.ng-london.org.uk/scientific/spd/?page=info>
- [31] Dorcus, R. M. (1926). Color preferences and color associations. *The Pedagogical Seminary and Journal of Genetic Psychology*, 33(3), 399-434.
- [32] Ou, L. C., Luo, M. R., Sun, P. L., Hu, N. C., & Chen, H. S. (2012). Age effects on colour emotion, preference, and harmony. *Color Research & Application*, 37(2), 92-105.
- [33] Taylor, C., Schloss, K., Palmer, S. E., & Franklin, A. (2013). Color preferences in infants and adults are different. *Psychonomic bulletin & review*, 20(5), 916-922.