Measuring the Uncertainty of Colour Matching

W. Luo, M. R. Luo*, S. Westland, A. Tarrant* and A. Robertson**
Centre for Colour Design Technology, University of Leeds, UK
**Department of Colour and Polymer Chemistry, University of Leeds, UK*
***Institute for National Measurement Standards, Canada*
Corresponding author: W. Luo (texwl@leeds.ac.uk)

ABSTRACT

This study investigated various uncertainties in colour matching, including matching stimulus and measuring instruments (such as random fluctuation, drift over time, spatial uniformity, temperature), calibration (photometric, wavelength and additivity) and observer (intra- and inter agreements). From the research results, it was found that the observer error was the major source of errors in colour matching, which was about 10 times of instrumental uncertainties. While among the instrumental uncertainties, the major source of error was spatial non-uniformity of beams and filters. The occurrence of uncertainties could influence the derivation of colour matching functions. Another task for the project was to reduce uncertainties in colour matching as much as possible. The uncertainties of instruments could be reduced to a minimum by adjusting the equipment and choosing better methods of measurement. It was found the uncertainty of observers could be further reduced by choosing a better way of arranging experiments.

1. INTRODUCTION

CIE standard colorimetric observers, the basis of colorimetry, have been widely used in industry[1]. The colorimetric data based on these standard observers are used for calculating colour differences and colour-appearance attributes. However, there has been insufficient information about the uncertainty in the standard observer data[2]. A better quantitative understanding of the uncertainties will provide important insights into the accuracy of the colorimetric data reported. Furthermore, breakdowns of the CIE standard observer can also be studied relative to experimental uncertainties. Therefore measuring uncertainty in colour-matching experiments is essential.

There are many sources of uncertainty in colour-matching experiments. According to their sources, they can be categorised into two main types: uncertainties due to instruments and due to observers. The uncertainties due to instruments also consist of two types: uncertainty of stimulus and uncertainty of measuring equipment. The uncertainties due to observers include intra- and inter-observer variability. The intra-observer error is the error within one observer, which can be considered as 'observer repeatability' whilst the inter-observer error is the error between observers, which is called 'observer accuracy'. In this study an analysis of the experimental results was used to determine the major sources of error in colour matching with the aim to reduce the sources of uncertainty as much as possible.

2. METHOD

The Tarrant Visual Colorimeter[3] was used for the experimental investigation of uncertainties in colour matching with a panel of observers. The set-up of the matching experiment is shown in Figure 1. The intensity of the R, G and B channels can be adjusted to match the test colour (T channel) on the left-side of the viewing field. A built-in photometer and a spectroradiometer were used as the measuring equipment. Four colours, white, cyan, yellow and orange-red were chosen as the test stimuli in the colour-matching experiments. These colours were generated from a tungsten source using narrowband filters to produce the cyan (500nm), yellow (580nm) and orange-red (610nm) narrowband stimuli and a daylight filter to produce the white stimulus.

The uncertainties due to instruments were first investigated with seven individual tests to evaluate temporal stability, instrumental agreement between photometer and spectroradiometer, stray light, repeatability, uniformity of illuminance over matching fields, photometer additivity and channel interaction.

The next stage was to measure the uncertainties due to observers and this included intra- and inter-observer error in matching the four coloured stimuli using an additive mixture of three primaries. Six observers (three male, three female; aged 23 – 38) performed the colour-matching experiments. Two methods for arranging colour-matching experiments were investigated. In the Individual Colour-Matching method each observer matched only one stimulus ten times in one day and matched a different stimulus each day. In the Mixed Colour-Matching method each observer matched each of the four stimuli once each day continuously for ten days. Therefore each observer was asked to do 20 matches for each test colour (80 colour matches in total).

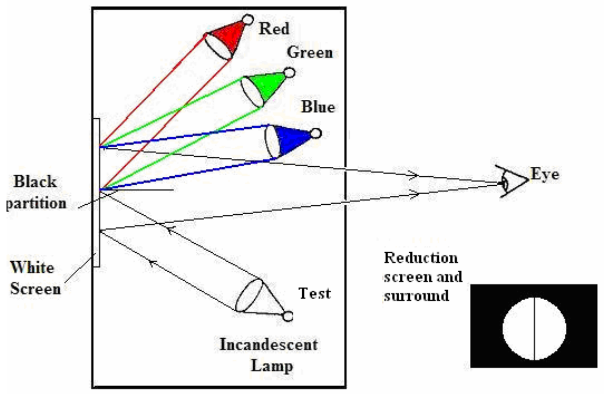

Figure 1 Typical Set-up for Visual Colour Matching Experiment

Table 1: The Experimental strategy of Colour Matching

Individual Colour-Matching sessions	
Day 1	10 matches of white
2	10 matches of cyan
3	10 matches of yellow
4	10 matches of orange-red
Mixed Colour-Matching sessions	
5-14	1 match for each of four test colours

3. RESULTS

Instrumental uncertainties

Seven tests were carried out to investigate the instrumental uncertainties. In the warming-up time test, the illuminance level of the lights was found to be unstable until at least 2 minutes after turn-on the instrument and stable thereafter.

In the inter-instruments test, it was found that the two colour-measuring instruments agreed well with each other; i.e. the correlation coefficient between two sets of readings was about 0.99. However, it was also found that the spectroradiometer (Photo Research PR 650) cannot measure below certain luminance levels (different channels have different limits) whereas the built-in photometer was capable of measuring the full range of voltage values.

In the stray light test, the intensities of the two sides of the viewing filed were measured when only one side (test side) was illuminated. The ratio of the other side and the test side of the viewing filed was less than 0.3%. This shows that the influence of the stray light was quite small. The largest stray light occurred in the G and B channels, while the T channel gave the smallest stray-light ratio.

In the repeatability test, the short-term (10 minutes) error was less than 0.04% and the medium-term (in 3 hours) error and long-term (in 3 days) error were each less than 0.5%. It was concluded that the repeatability error of the colorimeter was acceptable.

The uniformity of illuminance over the whole area of the matching fields was investigated. The uniformity of beams (T, R, G and B channels) and filters (four test colours) were both measured. The built-in photometer and PR-650 were used to measure 12 points in arc and vertical positions. It was found that the percentage errors (CV) for the R, G, and T channels were less than 1%, whereas the B channel had the largest non-uniformity with a CV about 3%. For filters, their uniformity was a combination of the uniformity of the filter itself and the uniformity of the Test beam. Experimental results indicated that the error in uniformity of the filters was less than 1%.

In the colour-matching experiments, the intensities of the R, G and B channels are measured by the built-in photometer and therefore its additivity performance was examined by measuring a stimulus comprised of R, G and B channels, and then measuring each channel separately. The additivity error for the photometer was found to be about 0.1%.

The last test of instrumental uncertainty was to check the interaction between the four channels (T, R, G and B). During colour matching the test colour must be kept constant during the time when the observers adjust the three primaries but, due to a limitation of the electronic design of potentiometers, the four channels might interfere with each other. The Test beam was set at a fixed voltage level whilst changing the voltage input of the other three channels (R, G and B) from 0 to 1000 at step intervals of 100 (these three channels were put outside of the viewing field). The interaction error was found to be less than 0.05%.

Observer uncertainties

For measuring the uncertainties of the observers, the data obtained from the colour-matching experiments were illuminance readings of the R, G and B channels. Because the colorimeter is based on the law of additive colour mixing, it was able to calculate the tristimulus values from the visual colorimeter using transform matrices (Equation 1),

$$\begin{bmatrix} X' \\ Y' \\ Z' \end{bmatrix} = \begin{bmatrix} x_r/y_r & x_g/y_g & x_b/y_b \\ 1 & 1 & 1 \\ z_r/y_r & z_g/y_g & z_b/y_b \end{bmatrix} \begin{bmatrix} l & 0 & 0 \\ 0 & \mathrm{m} & 0 \\ 0 & 0 & \mathrm{n} \end{bmatrix} \begin{bmatrix} R_L' \\ G_L' \\ B_L' \end{bmatrix} \quad (1)$$

where *X', Y'* and *Z'* are the normalised tristimulus values, and R_L', G_L' and B_L' are the normalised illuminance reading. The first transform matrix is generated from the chromaticity coordinates for each primary used to match an equal-energy illuminant (SE) measured by a TSR PR-650. The unknown coefficients *l, m, n* were calculated for the normalized illuminance (R_L',=G_L'= B_L'=100), the measured chromaticity coordinates, and normalized tristimulus *Xw, Yw, Zw* for SE. The equation was solved to give: *l*=0.16749, *m*=0.72887, *n*=0.10384.

Observer repeatability is the closeness of the agreement between the results of successive observations for the same observer carried out under the same conditions of measurement. Each observer was asked to measure each test colour 10 times and observer repeatability was expressed quantitatively in terms of CIELAB colour difference between each observation and the mean of the 10 observations. Since there were two methods of organizing the colour-matching experiments (Individual Colour-Matching and Mixed Colour-Matching) the observer repeatability was calculated using both data sets for comparison.

Table 2: Observer repeatability in Colour Matching: 'Individual' represents the Individual Colour-Matching sessions; 'Mixed' represents the Mixed Colour-Matching sessions.

Obs. Repeatability	White	500nm	580nm	610nm
ΔE^*_{ab} (Individual)	1.14	10.90	1.63	2.72
ΔE^*_{ab} (Mixed)	1.92	11.45	1.88	3.00

As can be seen in Table 2, for both methods, white had the best observer repeatability, while the cyan was the worst of the four colours tested. The reason may be that the human visual system is more sensitive to changes in a neutral colour. For cyan, the matching may be difficult because at 500nm (the dominant wavelength of cyan) the G and B colour-matching functions intersect whilst the R colour-matching function is negative. The 'Individual' method had smaller colour differences than the 'Mixed' method. The reason may be that in Mixed Colour-Matching sessions, observers were asked to match four test colours once in a short time (approximately thirty minutes). For each session, the order of the four colours was randomly presented and it is likely that some incomplete adaptation of the visual system will occur between two matches making it more difficult to achieve repeatable match between sessions. Consequently, considering the observer repeatability, the Individual Colour-Matching method should be recommended for organising experiments.

Observer accuracy is the closeness of the agreement between the result of each individual observer and a true value of the measurement. In this study, the mean of all the observations for each colour was regarded as the true value, so the observer accuracy can be called as the observer variation. The inter-observer variations are shown in Table 3 where the colour differences are calculated between the mean of each observer and the mean of all observations. Again, the cyan colour had the largest colour difference and the white colour had the smallest colour difference. Unlike the observer repeatability, the observer accuracy of the two experimental arranging methods were quite similar.

Table 3: Observe accuracy in Colour Matching: 'Individual' represents the Individual Colour-Matching sessions; 'Mixed' represents the Mixed Colour-Matching sessions.

Obs. Accuracy	White	500nm	580nm	610nm
ΔE^*_{ab} (Individual)	1.34	10.34	2.42	3.09
ΔE^*_{ab} (Mixed)	1.40	11.27	2.26	2.56

Since the investigation of instrumental uncertainty was based on an illuminance meter, percentage error (CV) in Lux units was used to show the degree of uncertainties. A zero CV means a perfect agreement. Since measuring the uncertainty of observers involved psychophysical experiments, the results are best represented by colour differences, say ΔE^*_{ab}. Although there are advantages in using different units for the two types of experiments, it is difficult to compare the results. Fortunately, the visual results were measured by both the illuminance meter and TSR. Hence, they can also be represented in Lux. Therefore, $\Delta E'$ was introduced to quantify all the uncertainties in colour matching as given in Equation 2,

$$\Delta E' = \sqrt{\Delta Lux_R^2 + \Delta Lux_G^2 + \Delta Lux_B^2} \qquad (2)$$

where ΔLux_R , ΔLux_G , and ΔLux_B are the differences in Lux values of the R, G and B primaries.

Table 4 summarises the overall uncertainty results in $\Delta E'$ units. It can be clearly seen that the observer errors were the major source of error in colour matching, i.e. the errors of observer variability were about 10 times larger than the instrumental errors. Among all instrumental uncertainties, the major source was found to be the lack of spatial uniformity of the stimuli.

Table 4: The overall uncertainties in colour matching in ΔE' unit.

Uncertainty sources		ΔE'
Instrumental Uncertainty	Stray Light	0.0100
	Non-uniformity	0.0356
	Repeatability	0.0159
	Additivity	0.0087
	Interaction	0.0043
Observer Uncertainty	Obs. Repeatability	0.2555
	Obs. Accuracy	0.3926

4. CONCLUSIONS

The present results showed that the observer error was the major source of error in colour matching. For the instrumental measurements, the major source stems from the spatial non-uniformity of the stimuli. Besides determining the major source of uncertainties in colour matching, another task in this study was to reduce the sources of error as much as possible. The uncertainty of instruments can be reduced to a minimum by correct adjustments of the equipment and by designing better methods of measurement. The uncertainty of observers can be reduced by an appropriate experimental design. From the results of the psychophysical experiments, it was found that the Individual Colour-Matching method performed better than the Mixed Colour-Matching method in terms of inter- and intra-observer error. The former method should be used in future experiments.

References

1. A. R. Robertson, "Overview of sixty years of CIE colorimetry", in Proc. CIE Symp. On Advanced Colorimetry, Vienna, pp. 3 - 6, CIE Central Bureau Publ. CIE x007 (1993).
2. A. R. Robertson, "Review of experiments leading to and properties of the 1931 and 1964 standard observer", CIE Central Bureau Publ. CIE x007, 46-47(1993).
3. A. Tarrant, "The Colour Tutor: Visual Color Matching Equipment for Teaching and Research", in Proceedings of the 9th Congress of the International Colour Association, Rochester, NY, 406-407 (2001).