

VISUAL EXPERIMENT ON LED LIGHTING QUALITY WITH COLOR QUALITY SCALE COLORED SAMPLES

Nicolas Pousset, Gaël Obein, Annick Razet

Institut national de métrologie / Conservatoire national des arts et métiers
(LNE-INM/Cnam) La Plaine Saint-Denis, France

ABSTRACT

A psychophysical experiment developed to evaluate light quality of Light Emitting Diodes (LEDs) is described. It is based on coloured samples used in the "Color Quality Scale" (CQS) proposed by the NIST to replace the "Colour Rendering Index" (CRI). Predictions of the CQS and results from visual measurements are compared.

Keywords: Light Emitting Diode, quality of light, Color Rendering Index, Color Quality Scale, visual measurements, paired comparisons protocol, law of comparative judgment.

1. INTRODUCTION

The emergence of white high brightness LEDs, in the early 2000s, helped to renewed the debate on the evaluation of light quality through Colour Rendering Index (CRI). Several laboratories have launched research projects on this subject. It has been shown, by visual measurements with observers and by simulations, that CRI calculation according to the "Commission Internationale de l'Eclairage" (CIE) recommendation [CIE 95] were not satisfying for LED lightings [TAR 01] [NAR 02] [CIE 07] [BOI 07] [BOI 09].

From these observations, CIE recommended the development of a new calculation method in order to determine a new index (or a new set of indices) of colour rendering by a Division 1 Technical Committee. As there is a lot of possibility to define the colour rendering mathematically, it was decided that the evaluation of results based on visual experiments will always be necessary to promote one method over another [CIE 07].

During the CIE meeting in Ottawa in May 2006, the technical committee 1-69 (TC 1-69): "Colour rendering of white light sources" was created. Wendy DAVIS, from National Institute of Standards and Technology (NIST) has been appointed to head of this committee. The aim of the

TC 1-69 is "to investigate new methods for assessing the colour rendition properties of white-light sources used for illumination, including solid-state light sources, with the goal of recommending new assessment procedures". In September 2009, this committee were composed of about forty members and the LNE-INM/Cnam is one of them. In the last days of September 2009, ten laboratories participating to the TC 1-69 has issued reports on the results of their work. Several reports proposing new methods for assessing the quality of light have been submitted. The first to be published and formalized clearly was from the NIST. So, we decided to developed a psychophysical experiment based on this new method named "Color Quality Scale" (CQS) [DAV 05] [DAV 06] [DAV 07] [DAV 09].

In this context we have evaluated the sensation of quality of light for a set of coloured samples illuminated under different types of LED lightings using visual measurements. The experimental results obtained were compared to those obtained mathematically by the CQS. This study used one part of the coloured samples proposed by the CQS. These samples have been selected according to physical measurements.

2. PHYSICAL MEASUREMENTS

Our experimental bench, developed at the LNE-INM/Cnam, is dedicated to the characterization of high brightness LEDs. It is composed of a light booth with five embedded LED lightings (two cool white (denoted 1 and 2) and three warm white (denoted 3, 4 and 5)). Each lighting is composed of ten 1 W LEDs connected in series and mounted on a heat sink made of anodized aluminium. Each one is powered by a continuous current power supply whose relative stability is of the order of 10^{-5} .

Illuminance levels in the light booth were equalized by adjusting the current supply of each LED lighting compared to the nominal

current of 350 mA. The illuminance levels were measured at three point in the centre of the light booth. The average illuminance level is equal to 150 lx and the uniformity of illuminance in an area of 30 by 20 cm - corresponding to the size of coloured samples used - at the centre of the booth is 4%.

The illuminance absolute value in the booth is relatively low due to the size of the booth (1 m in length by 60 cm in depth and height) and the number of LEDs placed inside it. This aspect was not considered binding because observers were positioned in a completely dark environment where the only source of light come from the booth. Visual comfort was judged sufficient.

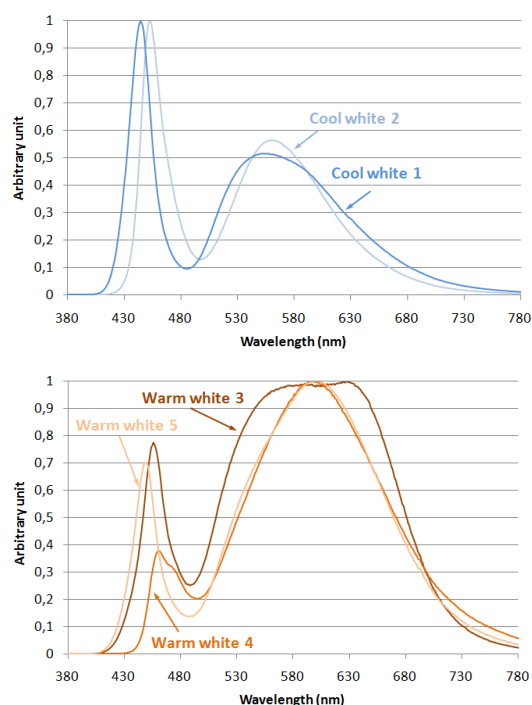


Figure 1. Top: spectra of cool white type LED lightings mounted in the LNE-INM/Cnam light booth. Bottom: spectra of warm white type LED lightings mounted in the LNE-INM/Cnam light booth.

Relative spectral measurements on these five LED lightings are performed by a spectroradiometer placed in front of the booth. The spectral distributions of sources 1 and 2 are shown in figure 1 (top) and those of sources 3, 4 and 5 are shown in figure 1 (bottom). The averages of correlated colour temperatures for the two systems 1 and 2, cool white type, are 5930 K and 6100 K, and those for the three systems 3, 4 and 5, warm

white type, are 3500 K, 2850 K and 3030 K, respectively.

For more details see POUSSET et al. [POU 09].

The special CQS indices, from Q_1 to Q_{15} , the general index Q_a , and the additional indices Q_f (colour fidelity scale) and Q_p (colour preference scale) were calculated, according to the spectral measurements of LED lightings, using the spreadsheet "NIST CQS Simulation 7.4" provided by W. DAVIS. Results (table 1) allowed us to select coloured samples. We have considered that the psychophysical experiment should not exceed forty five minutes. It forced us to select only eight samples. We took these which present the largest variations of index between different lightings. These samples were also chosen uniformly distributed along the Munsell hue circle. According to the indexes reported in table 1, we selected Q_1 , Q_3 , Q_4 , Q_8 , Q_{10} , Q_{12} , Q_{13} , Q_{14} .

These eight coloured samples, with a mate finish, were used for visual measurements with a panel of observers.

3. VISUAL MEASUREMENTS

The aim of our psychophysical experiment was to test the quality of LED lightings on different coloured samples from a visual point of view. This is done by establishing the ranking of the five lightings, from the best to the worst, for each selected sample according to observers' answers.

To make this experiment we didn't want to use incandescent artificial lighting or natural light as a reference so that the observer wasn't influenced by this kind of spectrum.

We used a paired comparisons protocol. The experiment is based on the ability, for an observer, to compare the colour of a sample observed successively under two different LED lightings. This comparison should allow the observer to make a judgment (good/not good) on the quality of the observed colour.

Before starting the experiment each observer was submitted to a Farnsworth D15 test to check their colour vision. Observers participating to the experiment were dressed with a white coat to hide their clothes. White gloves were also provided so they are not influenced by the colour of their hands. After that, each observer was placed in front of the booth in a dark room (fig. 2).

Table 1. Indices calculated by the CQS on the five LED lightings of the LNE-INM/Cnam light booth.

CQS indices	Cool white 1	Cool white 2	Warm white 3	Warm white 4	Warm white 5
Q_1	81	76	90	77	76
Q_2	90	97	98	91	97
Q_3	58	60	79	82	69
Q_4	54	40	75	70	66
Q_5	67	49	80	70	73
Q_6	72	54	82	68	72
Q_7	76	65	85	73	71
Q_8	87	79	91	84	80
Q_9	85	91	95	88	95
Q_{10}	67	72	89	81	79
Q_{11}	62	59	87	80	74
Q_{12}	65	59	88	80	74
Q_{13}	70	62	89	79	78
Q_{14}	63	51	84	69	73
Q_{15}	71	61	86	74	74
Q_a	69,3	61,9	85,2	76,6	75,2
Q_f	67,2	62,0	85,4	77,6	72,8
Q_p	76,5	64,9	85,5	75,2	82,3

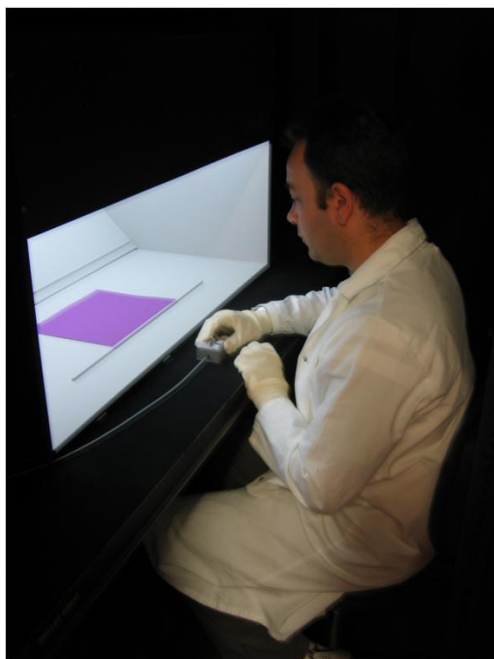


Figure 2. Observer placed in front of the light booth during visual measurements of the LNE-INM/Cnam psychophysical experiment.

We presented to an observer, for a given coloured sample used in the CQS, all the pairs of lightings that could be obtained with the five LED lightings (n) of the booth (excluding the pair formed twice from the same lighting). The observer is subjected to a series of ten $(n(n-1)/2)$ lighting pairs per coloured sample.

The choice of lighting pairs submitted to an observer was implemented by an operator, placed outside the dark room, which has a “control box” to enable or disable lightings. This choice was random and followed a uniform distribution. For each coloured sample a series of ten pairs of lightings was generated by computer (among the $10! = 3\,628\,800$ possible series), through a program developed with labVIEW. This allowed to overcome the effect of presentation of lighting pairs on results between observers.

For each pair of lightings the observer must answer to the question asked by the operator :

“For each pair of lightings, indicates which one gives the best quality of light on the coloured sample”

An “observer box” located on the left of the light booth (fig. 2), allowed to the observer to switch between the two proposed lightings, and to validate his choice by push a button. For each pair, the observer chose the LED lighting that provided on the coloured sample the best quality of lighting. Answers of the observer gave a system of ten inequalities.

This process was repeated for the eight coloured samples used in the CQS and for thirty five observers. The order of presentation of the coloured samples submitted to each observer was random and

followed an uniform distribution in order to overcome the possible effect of presentation on results.

Finally, 2 800 observations were collected (35 observers x 8 samples x 10 pairs of lightings) with an average forty minutes duration of the experiment for each observer.

4. ANALYSIS AND RESULTS

Two analysis were performed on experimental results : an analysis by averaging observers' answers and an analysis using the comparative judgment law [THU 27a] [THU 27b].

Concerning the first analysis we could establish a ranking of the LED lightings, from the best to the worst, according to observers' answers by solving system of ten inequalities. By a count of lightings sorted in first, then second, then third,... for all observers, and dividing the sum by the number of observers, it was possible to obtain the percentage of responses associated to the choice of a lighting in first, second, third,... position. Figure 3 gives a representative example of the distribution of observers' answers, as a percentage, for a dark blue colour sample (Q3). The horizontal axis of the graph correspond to the ranking of LED lightings from the best (1) to the worst (5).

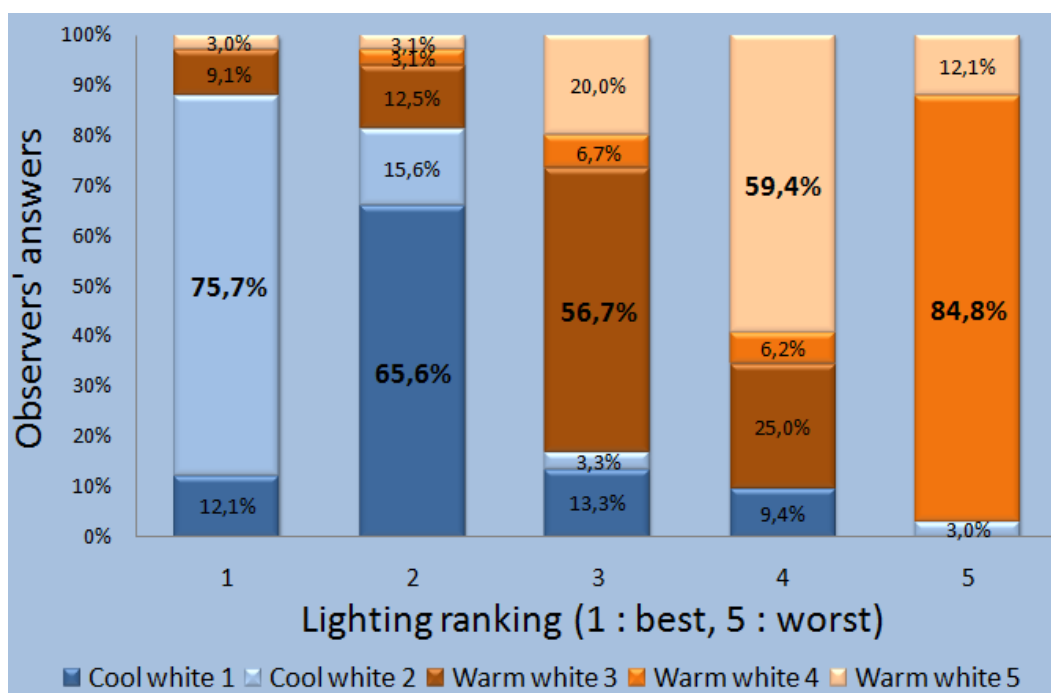


Figure 3. Results obtained with the analysis by averaging observers' answers for the Q3 sample (dark blue).

The average ranking of LED lightings was obtained for each sample, considering lighting which had a majority of answers from observers. For example with Figure 3, "Cool white 2" was judged the best lighting by approximately 76% of observers and "Warm white 4" was judged the worst by 85% of observers.

This first analysis method provides à LED lighting ranking from the best to the worst but it does not allow to determine a quantification scale of differences between each lightings. To solve this aspect we applied, in a second analysis, the law of comparative judgment.

Concerning this second analysis, we quantified the differences between each lighting according to the probability of judgement observers which express the proportion from a lighting is judged better than other. From the lighting ranking obtained for each sample, it was possible to compare results issued from visual measurements and those obtained from CQS calculation method. Figure 4 shows a representative example. Uncertainties were associated to the average values of normal deviate (M(z)) calculated from the law of comparative judgment. Uncertainties are represented by error bars in figure 4.

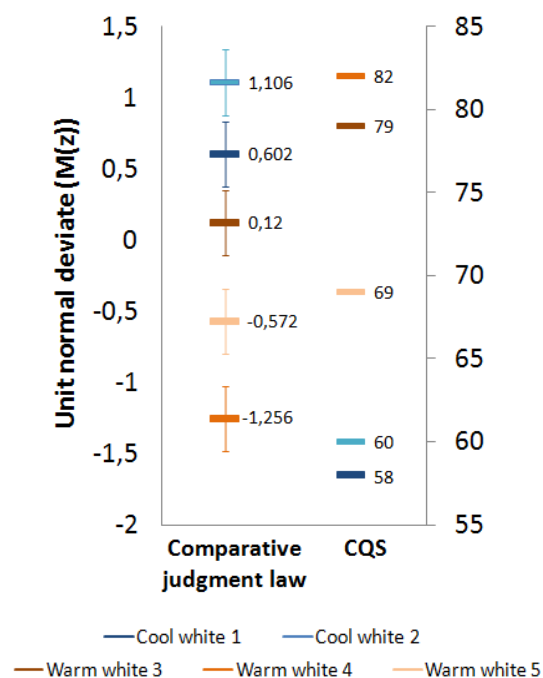


Figure 4. Experimental results obtained with the comparative judgment law (left) compared to calculation results obtained with the CQS (right) for the Q3 sample (dark blue).

From these two complementary analysis we can give the following results on coloured samples:

Blue samples

For Q3 and Q4 samples (Figure 3, Figure 4 and Figure 5) discrimination between lightings is very good. Rankings issued from observers agree with a majority of answers (> 50%). For these samples it appears that the light quality on this kind of colour is clearly defined by the observers in the same way. The cool white lightings are ranked in first and the warm white lightings in last. These rankings are not consistent with those obtained by the CQS.

Orange-red samples

For Q12, Q13 and Q14 samples (Figure 5) the observers' answers are much more dispersed. There is a less good discrimination between lightings however it seems that warm white lightings are ranked in first and cool white lightings in last. This is consistent with the CQS predictions. However, the term light quality seems to be problematic. There is obviously no consensus between observers on this issue contrary to what seems to happen for bluish samples. According to the uncertainties on

visual measurements, predictions of the CQS seem acceptable.

Other coloured samples

For the Q1 sample (Figure 5) observers' answers are scattered, except for the warm white 4 which is ranked last with 91% of answers. The CQS predicts that warm white 4, cool white 2 and warm white 5 are very close. According to observers' answers, warm white 4 seems to be really less good than the two others lightings. Furthermore, it appears that cool white lightings are first ranked for this sample.

For the Q8 sample (Figure 5), discrimination between lighting is good. Cool white lightings are in majority first ranked. Observers' answers for the warm white match to more than 50%.

For the Q10 sample (Figure 5), observers' answers give a bad discrimination between lightings so cool white lighting is also first ranked but with a low majority. As for Q1 sample, warm white 4 is ranked last with 65% of answers then it is ranked in second by the CQS.

5. CONCLUSIONS

The concept of quality of light appears to be correlated with the correlated colour temperature of the lighting and the colour of the sample. For instance, for bluish samples, the observers select preferentially a lighting with a high colour temperature to obtain the best quality of illumination. For orange/red samples, observers' answers are fuzzier that for bluish samples but warm white lightings seem to be favoured.

We think that if the observers prefer a cool white lighting over a blue sample and a warm white lighting on an orange-red sample, it is because the notion of quality in the context where we tested it, refers mainly for the observers to a sensation of saturation on the samples.

The term quality forwards to multiple subjectivity criteria, even sometimes contradictory according to observers. An anthropological and linguistic approach seems necessary. It is already possible to say that some observers prefer lighting that produces a saturation effect on the sample while others reject this saturation effect to the benefit of the sharpness of the surface

where we can distinguish defaults of uniformity on the sample.

According to the multitude of concepts behind the term quality, we can question the relevance of a single indicator to quantify the quality of light. The definition of several indices referencing to more explicit concepts seems more appropriate than a single user-friendly commercial index, inappropriate for expert user.

Finally, from the results obtained within the conditions of our experiment, the predictions of the current CQS does not seem satisfactory.

So we propose some improvements to the CQS:

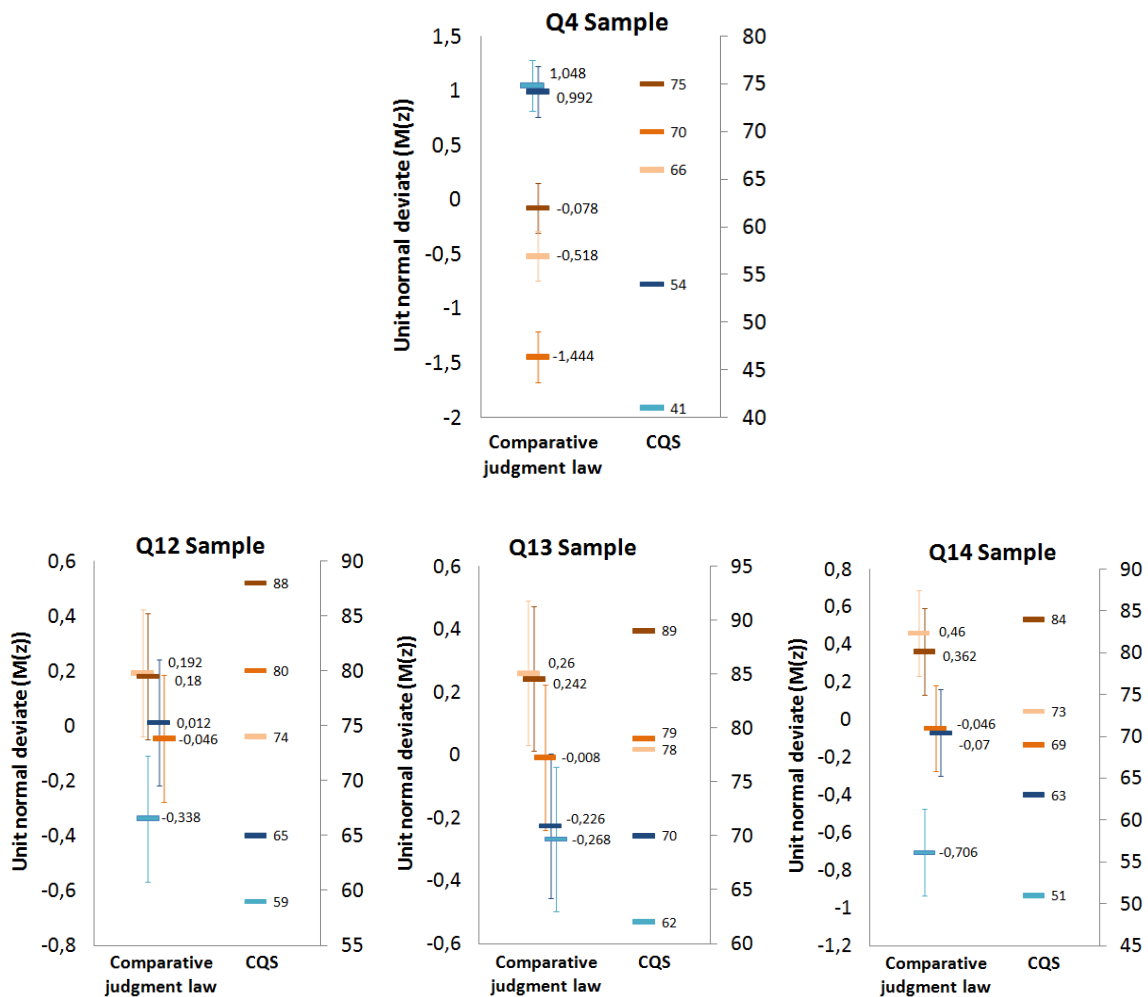
- We think it would be interesting to add a parameter that takes into account the general colour of the scene and promote lightings that enhance the saturation of colour, in the calculation of special indices Q_i , particularly for bluish samples. Observers' results seem to be unambiguous on this point.

- It appears, for several coloured samples (especially Q1, Q3, Q4), that warm white 4 lighting (correlated colour temperature equal to 2850 K), is often ranked in last position by the observers. This is because this lighting illuminates samples with a very yellowish colour. It seems interesting to review the calculation of "M_{CCT} factor" [DAV 09] introduced in the CQS to further penalize sources of low correlated colour temperature with these samples.

- We also propose that the additional indices Q_p and Q_f , developed for expert users, are calculated for each sample to complete the special indices Q_i , and not merely supplementing the general index Q_a .

ACKNOWLEDGEMENTS

The authors would like to thanks Stéphane Aubert of the LNE-INM/Cnam for his important contribution to this project in the field of electronics.



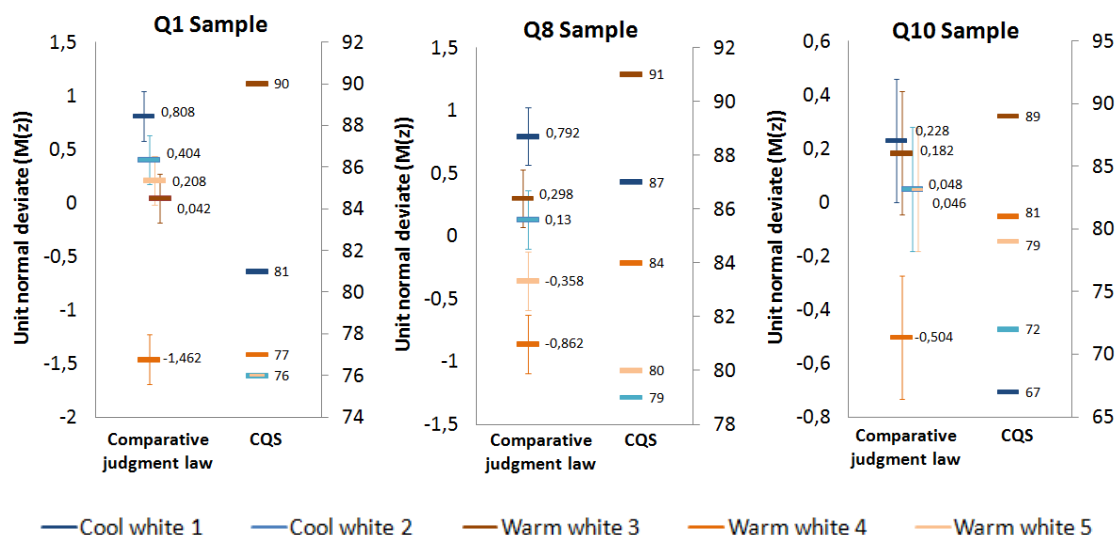


Figure 5. Experimental results obtained with the comparative judgment law (left) compared to those obtained with the CQS (right) for the selected CQS coloured samples.

REFERENCES

- [BOI 07] S. Boissard, M. Fontoynt, P. Avouac, Optimisation of mixing of LED-based light for object presentation, Proceedings of the 26th session of the CIE, 2007.
- [BOI 09] S. Boissard, M. Fontoynt, Optimisation of LED-based Light Blendings for Object Presentation, Color research and application, vol. 34, n°4, 2009.
- [CIE 95] Commission internationale de l'éclairage, Method of Measuring and Specifying Colour Rendering Properties of Light Sources, 3rd edition, CIE 13.3 – 1995.
- [CIE 07] Commission internationale de l'éclairage, Colour rendering of white LED light sources, CIE 177 – 2007.
- [DAV 05] W. Davis, Y. Ohno, Toward and improved color rendering metric, Fifth International Conference on Solid State Lighting, vol. 5941, p. 59411G.1-59411G.8, 2005.
- [DAV 06] W. Davis, Y. Ohno, Development of a Color Quality Scale, Sixth International Lighting Research Symposium on Light and Color, 2006.
- [DAV 07] W. Davis, Y. Ohno, Measuring color quality of light sources, Eleventh International Symposium on Science and Technology of Light Sources, p. 459-468, 2007.
- [DAV 09] W. Davis, Y. Ohno, The Color Quality Scale, TC1-69 report, 2009 (only for TC members).
- [NAR 02] N. Narendran, L. Deng, Colour Rendering Properties of LED Light Sources, Solid State Lighting II: Proceedings of International Society for Optical Engineering, 2002.
- [POU 09] N. Pousset, B. Rougié, A. Razet, Uncertainty evaluation for measurement of LED colour by Monte Carlo simulations, Metrologia, vol. 46, n°6, p. 704-718, 2009. Corrigendum, vol. 47, n°1, p. 112-112, 2010.
- [TAR 01] T. Tarczali, P. Bodrogi, J. Schanda, Colour rendering properties of LED sources, CIE Expert Symposium on LED measurement standard methods for specifying and measuring LED and LED cluster Characteristics, 2001.
- [THU 27a] L. L. Thurstone, Psychophysical analysis, American journal of psychology, 38, p. 368-389, 1927.
- [THU 27b] L. L. Thurstone, A law of comparative judgment, Psychological Review, 34, p. 273-286, 1927.

AUTHORS

Nicolas Pousset, Gaël Obein, Annick Razet

Institut national de métrologie / Conservatoire national des arts et métiers (LNE-INM/Cnam), 61 rue du Landy, 93210 La Plaine Saint-Denis.
Tèl : + 33 1 58 80 89 03
Fax: + 33 1 58 80 89 00
e-mail : nicolas.pousset@cnam.fr