Testing the Performance of a Modified Whiteness Formula for Dentistry

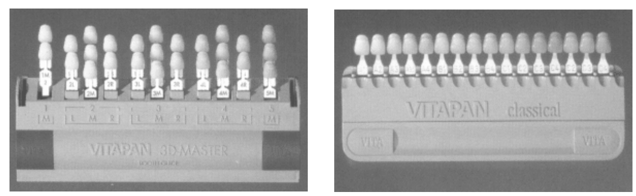
W. Luo andS. Westland, Centre for Colour Design Technology, University of Leeds, Leeds, United Kingdom;
I. Pretty and R. Ellwood, Colgate-Palmolive, Dental Health Unit, Manchester, United Kingdom

Abstract
In this study, various whiteness and yellowness indices were compared with regard to their ability to measure the perceived whiteness of human teeth. Psychophysical experiments were conducted by 80 observers on tooth whiteness perception under typical clinical test conditions. The Pearson correlation coefficient and 'per-cent-wrong decision' criterion were used to determine the best index for tooth whiteness measurement. The results confirmed the findings of a previous study that a modified form of the CIE Whiteness Index formula is appropriate for the prediction of tooth whiteness.

Introduction
Whiteness is an attribute of colours of high luminous reflectance and low purity situated in a relatively narrow region of the colour space along dominant wavelengths between 570 nm and 470 nm.[1] It is recognised that whiteness is a special colour attribute and many whiteness formulae have been developed and applied with the aim to predict perceptual whiteness. Currently, a number of whiteness formulae are in common use including the CIE Whiteness Index (WIC), the whiteness index according to the ASTM E-313-73 (WI), and the Z% index.

The increasing use of consumer-based tooth-whitening systems and the application of digital imaging for the assessment of tooth whiteness[2,3] have increased the need for methods for the quantification of whiteness for teeth. Although some researchers propose that changes in yellowness are the most important factors in the assessment of tooth whitening[4] there is evidence that a whiteness index may be applicable for the prediction of tooth whiteness,[5] but there is still some uncertainty of the applicability of such formulae since many were designated to be used only with samples whose colour coordinates are within a narrow range (the colour of teeth appear to be outside this range[6]). There is a clear need for psychophysical studies of tooth-whiteness perception in order to validate the performance of various metrics. Recent work has demonstrated that an optimized form of the CIE Whiteness Index (WIO) can most accurately predict the whiteness of teeth.[7]

In a previous study,[7] 9 observers were asked to rank 26 teeth samples from the Vita 3D shade guide (see Figure 1 (a)) in order of perceived whiteness. The Vita 3D-Master shade guide consists of five groups according to their lightness level where all the tabs in one group are designed to have the same lightness. The tabs go from lightest to darkest when moving from left to right in Figure 1. Within each group there are only differences in chroma and hue. These properties make the 3D shade guide highly suitable for use in teeth-bleaching research.[8] In the previous study, visual assessments were made on a neutral gray background under a light source approximating the CIE D65 illuminant. An optimized form of the CIE Whiteness Index (WIO) was found to give best agreement with the visual results when compared with other whiteness and yellowness formula and with the CIELAB b* value.[7]



(a) (b)

Figure 1. (a) The Vita 3D shade guide; (b) The Vita Classical Shade guide. They are both sets of porcelain teeth that are used in the dental community for shade assessment purposes.

In this study, in order to test the performance of the optimized whiteness index (WIO) on other teeth samples, the Vita Classical shade guide (see Figure 1(b)) was used as a testing set. Like the Vita 3D shade guide, the Vita Classical shade guide is commonly used in dental bleaching studies.[9] It consists of 16 teeth with a simpler arrangement in a general trend from lightest to darkest when moving from left to right (Figure 1). However, unlike the Vita 3D shade guide, measurements have shown that there are some reversals in Lightness in the standard order of the Vita Classical shade guide.[8]

In addition to observations made by 9 observers in a controlled laboratory, a large number of observers (drawn from dental patients) were asked to rank the 16 teeth samples (or tabs) from the Vita Classical shade guide under typical clinical lighting conditions. The purpose of the study was to investigate whether the WIO index would give good agreement with visual observations under typical clinical (non-laboratory) conditions made by non-colour-experts using a different set of teeth samples than those used in the original study.

Experimental
Measurement of Samples
It is difficult to measure the spectral reflectance of real teeth and of shade guide samples because the surfaces of real and porcelain teeth are generally not flat and the samples are not completely opaque. Since a reflectance spectrophotometer requires the instrument to be in physical contact with the samples, the curved surface of teeth may cause non-uniform reflection of light to the

detector and the effect of translucency of teeth may be another problem especially when the instrument flash is particularly intense. Another characteristic of teeth is fluorescence, whereby light in the UV spectrum is absorbed and re-emitted in the short-wavelength region of the visible spectrum. Some of the shade-guide samples may also fluoresce and the degree of fluorescence of these samples and of real teeth will naturally vary properties of the light source. The light source used in the visual assessments in the laboratory was a D65 simulator. However, although, in principle, a spectrophotometer with a pulsed-xenon light source that is properly filtered to daylight D65[10] has the ability to calibrate the UV component to fluorescent standard, there could be small differences amount of UV component between the visual assessments and the measurements. The non-flat and translucent nature of the teeth samples, and the concern about variation in UV content between the light sources used for the visual assessments and in the spectrophotometer, all contributed to the decision to use a non-contact tele-spectroradiometer (TSR) for the measurements.

A Minolta CS1000a TSR was used in this study to measure the spectral reflectance since it can measure the colour of teeth in a way that matches the visual assessments (the same light source can be used for visual assessments and for measurements and the optical geometry can be kept constant). The set-up for teeth measurement using the TSR is shown in Figure 2. A viewing cabinet was used as an illuminant D65 simulator and the teeth were measured on a chromatically neutral background. A triangular holder was constructed to avoid specular reflection from the glossy surfaces of the samples.

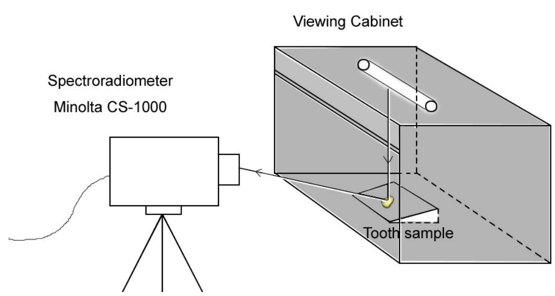


Figure 2. The set-up for measuring teeth using spectroradiometer Minolta CS1000a.

For each sample, the recorded spectral power at each wavelength interval was divided by the spectral power of the light source (obtained from measurements of a white tile, corrected for the spectral reflectance factors of the tile) to yield the spectral reflectance factors.

Whiteness/Yellowness Formulae

In order to calculate the whiteness and yellowness indices, the CIE tristimulus values of the teeth samples were obtained from measurements of the spectral reflectance of the teeth using TSR. The CIE XYZ values for each of the teeth were computed (D65 illuminant; 1931 CIE standard observer) from the reflectance factors. The candidate whiteness metrics considered were four whiteness indices (WIC, WIO, Z%, WI), the CIELAB b* value (b*), and two yellowness indices (E313, D1925).

The CIE whiteness formula (WIC) is given as Equation 1, where (x, y) and (x_n, y_n) are the chromaticity coordinates of the sample and the reference white respectively, thus,

$$WIC = Y + 800(x_n - x) + 1700(y_n - y) \quad (1)$$

The Z% whiteness index is computed from the CIE Z value, where Z_n is the CIE Z value of the reference white, thus,

$$Z\% = 100Z/Z_n \quad (2)$$

A whiteness index (WI) that combines Z% with luminance factor was also evaluated thus,

$$WI = 4Z\% - 3Y \quad (3)$$

The E313 and D1925 yellowness indices were also evaluated.

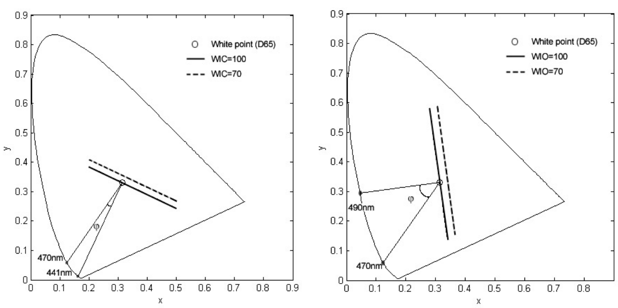
$$YIE313 = 100\,(1 - 0.847Z/Y) \quad (4)$$

$$YID1925 = 100\,(1.275X - 1.057Z)/Y \quad (5)$$

In the previous study,[7] 9 observers were asked to rank the 26 teeth from the Vita 3D shade guide in order of perceived whiteness when viewed on a neutral gray background under a light source approximating the CIE D65 illuminant. The ordinal-scale rank data were converted into interval-scale z scores (the higher the value of the z score, the greater the perceived whiteness). It was found that the CIE whiteness formula (WIC) gave the best result among the five previously published metrics that were tested. However, in order to provide an even better fit to the visual data, a modified version of the CIE Whiteness Index (WIO) was developed (Equation 6). The original linear form as suggested by Ganz was retained but the coefficients of the equation were optimized to achieve the best fit between the z scores from the visual assessments and the values from the index. WIO (Equation 6) was found to give good performance for the Vita 3D shade guide.[7]

$$WIO = Y + 1075.012(x_n - x) + 145.516(y_n - y) \quad (6)$$

In order to investigate why the optimized formula (WIO) performs better than the standard whiteness index (WIC), loci of iso-whiteness were plotted in the CIE chromaticity diagram. Figure 3 shows iso-whiteness lines at two levels of whiteness for both the WIO and WIC formulae. In this diagram each line represents samples of constant whiteness but of different hue. It was found that these lines are all parallel and shift along their perpendicular line from higher to lower whiteness values. Since Ganz found that variations in whiteness of neutral tint caused by changes in colorimetric saturation are observed in range of dominant wavelengths of approximately 465-475nm, d = 470 nm was selected as the reference dominant wavelength for neutral whites.[1] The angle between the dominant wavelength line and the perpendicular to the iso-whiteness lines is the hue preference angle. The angle of WIC is about 9.6^o and the dominant wavelength of its perpendicular is about 441nm (the intersection point of the perpendicular and the spectral locus) which indicates that bluish colours are considered to be whiter than yellowish colours in the WIC formula system. However, for the WIO equation, the hue preference angle is about -45°; the dominant wavelength of the perpendicular is about 490nm, which is a cyan colour; this indicates that, for teeth, green-bluish whites appear whiter than reddish whites. Therefore the effect of optimizing the coefficients in the WIO equation was to rotate the lines of iso-whiteness compared with the original CIE whiteness index in chromaticity space. This indicates that the hue preference of observers when assessing the visual whiteness of teeth is different from that of observers when viewing other white objects for which WIC has been found to be quite satisfactory.



(a) (b)

Figure 3. The comparison of (a) the CIE Whiteness formula (WIC) and (b) the optimized version (WIO) in the CIE chromaticity diagram with illuminant D65 and CIE 1931 2o observer. Based on a dominant wavelength for neutral whites of 470nm the WIC formula can be shown to have a hue preference φ of 9.6° whereas the optimized WIO formula has a hue preference of -45°.

Performance Testing

In this study, the Vita Classical shade guide (16 teeth) was used to test the performance of the optimised whiteness formula by comparison with visual assessments made by a large number of observers in a clinical environment.

As in the previous study, 9 observers were asked to rank the 16 teeth in the viewing cabinet under a CIE D65 illuminant. Furthermore, a total of 80 observers (drawn from dental patients who attended the Colgate Dental Health Unit in Manchester) were asked to rank the 16 teeth samples under typical clinical lighting conditions.

The Pearson correlation coefficients between the linear z scores and the values of the whiteness or yellowness indices were computed as a measure of goodness of fit between the visual data and the metrics. However, an additional analysis was carried out based upon the "% wrong decisions" criterion. Using this approach the consensus of the 80 visual rankings was deemed to be the correct rank order and this allowed the comparison of any pair of the 16 samples to be made and a consensus decision of which sample was whiter to be determined. Pair-wise comparisons were also extracted from each of the 80 observers' rankings and the number of wrong decisions made by each observer determined. The proportion of wrong decisions made on average by the observers was compared with the proportion of wrong decisions made by the various metrics.

Results

Table 1 lists the correlation coefficients between the visual rankings (expressed in z scores) pooled for the 80 observers and the various metrics for the 16 tabs of the Classical Vita shade guide.

It is noted that even though the WIO formula was optimized using the visual assessments of 9 observers for the 26 tabs of the Vita 3D shade guide, the WIO formula gives good agreement with the new data based on 80 observations of the 16 samples of the Vita Classical shade guide made under typical clinical conditions and also for the 9 observers who ranked these samples under laboratory conditions. However, although the correlation coefficients for all of the formulae are quite high, it is difficult to infer whether the use of any of the whiteness formulae would be adequate in practical terms. We define adequate performance in the context of this work to mean that a formula would result in no more errors than an average observer. In order to evaluate the formulae in this context, an additional analysis based upon the statistical method called '% wrong decisions' was therefore used in this study. Given the 16 samples of the Classical Vita shade guide there are 120 possible paired comparisons. Table 2 shows the results obtained when the all possible paired comparisons were considered. In addition, the 15 most colour-critical paired comparisons (each sample being compared to its neighbour in the canonical or consensus ranking) were separately considered.

Table 1: Pearson correlation coefficients (first row) between the various indices and z scores from visual assessment from this study using 80 (row 1) and 9 (row 2) observers and the 16 tabs from the Classical Vita shade guide. For comparison, the correlations from the previous study[7] are shown using 9 observers and 26 tabs from the Vita 3D shade guide (row 3).

Indices	WIO	WIC	Z%	Y(D1925)	Y(E313)	WI	b*
r^2(80 obs) 16 tabs	0.85	0.79	0.79	0.79	0.72	0.59	0.57
r^2(9 obs) 16 tabs	0.92	0.87	0.90	0.84	0.81	0.66	0.66
r^2(9 obs) 26 tabs	0.93	0.87	0.85	0.85	0.81	0.56	0.62

Table 2: Percent-wrong decisions for the various indices based on the full set of 120 paired comparisons and the 15 colour–critical paired comparisons of adjacent samples.

Indices	WIO	WIC	Z%	Y(D1925)	Y(E313)	WI	b*
120 pairs	6.67	8.33	9.17	10.83	9.17	14.17	15.00
15 pairs	26.67	33.33	46.67	33.33	40.00	40.00	40.00

The average per-cent-wrong decisions for the 80 observers was 6.20 and 16.44 for the 120- and 15-pair comparisons. (For comparison the equivalent scores for the 9 observers were 4.63 and 17.04 respectively.) Only the WIO formula gives performance that is comparable to that of the average visual performance and we therefore conclude that only this formula can be said to be acceptable in the context of practical use. The low values for all the indices in the first numerical row of Table 2 are caused by the fact that many of the pairs included samples that were very different in terms of whiteness. The results in Tables 1 and 2 all indicate worse performance than in the previous study during which a smaller number of observers made visual judgments under carefully controlled conditions[7]. However, the relative performance of the metrics was similar for the two studies.

Conclusion

Various whiteness and yellowness indices were compared for their ability to predict the perceptual whiteness of teeth. A psychophysical study involving 80 observers under typical clinical viewing conditions was carried out and the data from this study were best predicted using an optimized version of the CIE Whiteness Index (WIO) developed in a previous study[7]. Although the WIO index was optimized using observations made under laboratory conditions of a different set of porcelain teeth samples, it gave good performance when used to predict perceptual whiteness. When used to assess which of a pair of samples was whitest, four of the metrics tested (including the CIELAB b*) gave wrong decisions in 40% or more of cases. Further work in the study is concerned with visual assessments of teeth in vivo and the development of a digital-camera system to reliably record the colorimetric data that is needed as input to the optimized whiteness index.

References

1. E. Ganz, "Whiteness: Photometric Specification and Colorimetric Evaluation", Applied Optics, Vol. 15, No.9, pp.2039-2058, (1976).
2. S. Westland, "Review of the CIE system of colorimetry and its use in dentistry", Journal of Esthetic and Restorative Dentistry, S1, 5-12 (2003).
3. R.W. Gerlach, J.M. Dunavent, R.D. Gibb, M.F. Gabbard, H.L. Tucker and C.E. Martinez, "Comparative clinical effectiveness of two dual-phase whitening dentrifices", Conference for International Association for Dental Research, #1804 (2005).
4. A.P. Barlow, K. Brennan and A.P. Desai, "Clinical Comparison of Two Self-directed At-home Bleaching Systems", The 81st General Session of the International Association for Dental Research, 0900,Goteborg, Sweden, (2003).
5. J.A. Garcia, M. Rubino, J. Romero and E. Hita, "Measuring the Whiteness of Human Teeth", Color research and application, 18 (5), 349-352 (1993).
6. M. Rubino, J.A. Garcia, L. Jimenez del Barco and J. Romero, "Colour Measurement of Human Teeth and Evaluation of a Colour Guide", Color research and application, 19 (1), 19-22 (1994).
7. W. Luo, S. Westland, R. Ellwood and I. Pretty, "Evaluation of whiteness formulae for teeth", Proceedings of the 10th Congress of the International Color Association, 2005.
8. W.D. Browning, "Use of Shade Guide for Color Measurement in Tooth- Bleaching Studies", Esthetic and Restorative Dentistry, Vol. 15, S13-S20 (2003).
9. G.A. Gescheider, "Psychophysics: The Fundamentals", 3rd edition, Lawrence Erlbraum Associates (1997).
10. M. Burns, "Instrumental Measurement of Optically Brightened Textiles", Minolta Corporation, Instrument System Division (2005).
11. Axiphos Gmbh, "On whiteness formulas", Axiphos GmbH, Arend-Braye Str. 42, D-79540 Loerrach GERMANY, (2002)

Author Biography

Wen Luo completed her BSc degree (First class) in Computer Science and Communication at Southwest Jiaotong University in 2002, and then obtained her MSc degree (with Distinction) in Colour Science from the Colour Imaging Institute at the University of Derby in 2003. She is currently working on her PhD project 'Assessment of Tooth Whiteness' (sponsored by Colgate) under the supervision of Stephen Westland at the University of Leeds. Her current research interests include colour characterisation of digital cameras, imaging methodologies and psychophysical methods.