Perception and measurement of the whiteness of papers with different gloss and FWA amount

Ludovic Gustafsson Coppel
Innventia AB
Drottning Kristinas väg 61, Box 5604
SE-114 86 Stockholm, Sweden
ludovic.coppel@innventia.com

Abstract
The effect of mean gloss level on perceived whiteness was evaluated by magnitude estimation and pair wise comparison of papers with varying shades and gloss levels. The samples were printed on substrates of different gloss to target $L^*a^*b^*$ values measured under a light booth illumination having a 5000 K correlated colour temperature. Observers were able to rate the whiteness of the samples with large mean gloss differences in the 5000 K illumination and in the same illumination with an additional UV lamp. The CIE whiteness equation predicted well the perceived whiteness in both illuminations and the mean gloss level had no significant effect on perceived whiteness. This means that the CIE whiteness equation can be used to compare the perceived whiteness of papers having very different mean gloss levels. However, due to the different amount of fluorescent whiteness agents (FWA) in the papers, the perceived whiteness prediction was only valid when the measurement was performed under the same illumination as for the visual evaluations. Typical indoor illumination with fluorescent light tubes contains much less UV than the D50 or C illuminants. The instrumental whiteness differences due to mean gloss and instrument geometry were negligible compared to the effect of the UV content of the illumination.

Keywords: Whiteness, Gloss, Perception, Fluorescence

1. Introduction
The gloss of an object has a significant effect on its perceived colour. Matte surfaces appear less intensely coloured with lower chroma and higher lightness than glossy surfaces, all else being equal, and the perceived or measured colour depends on how much of the light reflected at the surface is detected (Dalal, 1999). Makarenko and Shaykevich (2000) studied theoretically the dependence of measured paper whiteness on the surface roughness and proposed a numerical model taking the measurement geometry of a d/0° instrument and surface roughness into account. Although the effect is stronger for saturated colours, they found that measured whiteness depends on both surface roughness and instrument geometry. The appearance of unprinted paper or board is strongly influenced by the mean gloss level, especially if one looks at it from different angles. Two papers judged as equally white and tinted will still appear different if the gloss is different, which raises the question if the glossy appearance affects the perception and judgement of whiteness.

The aim of this study was to investigate the effect of mean gloss level on the perceived and measured whiteness of paper samples. For this purpose, it was required to produce samples with different mean gloss levels and with all other optical properties being equal. Dalal et al. proposed a model that describes the influence of gloss on the measured colours of xerographic prints and tested the model on prints fused to different gloss levels on an off-line fuser (Dalal, 1999). Ariño et al. applied the same model to injection-moulded pigmented plastics and varied the surface gloss by using plaques with different structures (Ariño, 2005). Unprinted paper surfaces are more difficult to control without changing the other optical properties such as bulk scattering. Coated surfaces can for instance be modified by the choice of pigment size and shape. Calendering can be used to make the surface smoother but the process will also change the inner structure of the paper and hence its scattering properties.

2. Method
By printing the paper samples, it was possible to produce samples with different gloss but with nearly identical colour in one instrument geometry. Commercial inkjet papers are available with a wide span in gloss and were suitable for this study. However, one problem in measuring the shade of high whiteness
paper was fluorescence. Commercial papers have different amounts of fluorescent whitening agents (FWA) and their appearance depends on the UV content of the illumination. By measuring the whiteness of the samples in the same illumination and approximately the same geometry as in the visual evaluations with real observers, the influence of mean gloss level on perceived whiteness could be evaluated. This was done by having a panel of observers perform a magnitude estimation of perceived whiteness in two illuminations with different UV content. The results were then compared to the instrumental whiteness of the samples in the specific illumination. The samples with different gloss levels were also evaluated pairwise by the observers.

In order to quantify the effect of measurement geometry on the measured whiteness, the samples were measured with three instruments: the spectroradiometer used as a reference measurement for the visual evaluations, a d/0° spectrophotometer, and a 45/0° densitometer. To avoid the effect of the different UV contents in the measuring devices, two substrates were modified so that their FWA content remained the same but their mean gloss varied.

2.1. Paper samples

Six paper samples were printed with a Canon BJC-8200 photo inkjet printer at 4 different L*a*b* target values measured with a Photo Research PR-650 spectroradiometer under an overhead illumination with a correlated colour temperature (CCT) of 5000 K. The instrument was calibrated with a non-fluorescing Spectralon white tile. The reflectance of the tile was measured with a calibrated Technidyne ColourTouch spectrophotometer. The samples consisted of three inkjet photo papers (matte, semi-glossy, and glossy) and a copy paper. The glossy inkjet photo paper was used with and without a taped transparent film on top of it and the copy paper was used with and without extra calendering.

The spectral radiance factor of the samples was measured under the 5000 K illumination with and without an additional UV lamp. The spectral radiance factor was also measured with a TechKon spectro-densitometer (with and without polarisation filter), and with a Technidyne ColourTouch 2 spectrophotometer with the UV settings UV(D65), UV(C), and UVex. The mean gloss level was measured with a Zehntner ZLR 1050M gloss meter at 75° angle of incidence (Table 1).

<table>
<thead>
<tr>
<th>Substrate Type</th>
<th>Gloss 75° (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Copy Paper</td>
<td>11</td>
</tr>
<tr>
<td>2 Calendered Copy</td>
<td>18</td>
</tr>
<tr>
<td>3 Canon SG-201</td>
<td>50</td>
</tr>
<tr>
<td>4 Canon GP-501</td>
<td>79</td>
</tr>
<tr>
<td>5 Epson S041256</td>
<td>3</td>
</tr>
<tr>
<td>6 Canon GP-501 + transparent film</td>
<td>95</td>
</tr>
</tbody>
</table>

2.2 Visual evaluations

Four experiments were performed by a total of 43 observers with no previous experience in psychophysical scaling or paper related issues. All the observers had an average colour discrimination in the Farnsworth Munsell 100 hue test (Farnsworth, 1943).

1) The 24 samples (4 shades and 6 gloss levels) were presented one by one to 12 observers. The samples were attached to a grey background and laid on the grey table (Figure 1, left). The observers were instructed to report their perceived whiteness of the sample in relation to a reference sample, which was always present on the table. The reference was one of the printed Epson Matte samples. If the observers perceived the sample to be half as white as the reference they had to report the value 50, if they perceived it to be twice as white they should report 200. Each observer gave three individual estimates of their perceived whiteness. The individual response was taken as the geometric mean of the three observations.
The mean rating was calculated as the geometric mean of the individual mean responses. The samples were evaluated in a perception laboratory with homogeneous overhead illumination having a correlated colour temperature about 5000 K. The observers were standing at about 50 cm from the samples, away from specular reflexes from the samples.

2) The next 12 observers performed the same experiment holding the samples in their hands, free to look in any direction.

3) All the previous 24 observers participated also in a pair-wise comparison experiment. Samples from substrates 3, 4, and 5 were evaluated pair-wise. A pair of samples attached to a grey background were presented on top of a grey table (Figure 1, right). The observers were asked which sample they experienced as the whitest. In order to reduce the number of observer evaluations, the samples were not compared to themselves (same gloss and shade) and each pair was evaluated only once at random, but constant for all observers, left/right positions. This resulted in 66 evaluations from each observer.

4) 19 observers who did not participate in the previous experiments performed the same evaluation as in experiment 1 but under illumination with an additional UV lamp.

3. Results

3.1. Magnitude estimation

The rating of the 24 samples on the table is plotted versus the CIE whiteness measured in the overhead 5000 K illumination, \( W_{CIE}^* \), in Figure 2 (left) and the rating of the held samples is plotted in Figure 2 (right). The rating correlated highly with the measured CIE whiteness, irrespective of the mean gloss level of the samples, both when the samples were laid on the table and when the observers had the sample to be rated in their hands. The perceived whiteness of the samples having 50 % mean gloss level was higher than the perceived whiteness of the copy papers having 11 % mean gloss level, when the samples were evaluated on the table. However, samples with a 2 % mean gloss level were judged as white as glossier samples, and some of the glossiest samples (95 %) were judged less white than matter samples.

The rating of the 24 samples under an overhead 5000 K illumination with additional UV illumination versus CIE whiteness applied to that illumination is shown in Figure 3. Since the illumination was different, the reference sample’s appearance changed when the UV lamp was turned off. Therefore, the rating of the samples with additional UV should not be compared to the rating obtained in the illumination without additional UV.

Figure 1. Samples in the visual evaluations. Left: Magnitude estimation of the sample to the left compared to the right reference white (Experiments 1 and 4). In experiment 2, the observers were holding the sample to be evaluated in their hands and compared to the reference on the table. Right: Pair wise comparison in experiment 3.
However, since the matte sample used as a reference white was compared to itself, it is expected that observers should report 100 for that sample. The mean rating for that sample was 99 (Figure 3) in the illumination with additional UV and 98 in the illumination without additional UV (Figure 2, left). On the other hand, the measured whiteness was 125 with additional UV but only 103 without additional UV.

A high correlation was obtained between the rating in the illumination with additional UV and the measured whiteness in that illumination ($R^2 = 0.87$). All the samples but the printed Canon SG-201 inkjet papers (gloss 50%) had a significantly larger instrumental whiteness with the UV lamp on. The SG-201 samples exhibited very little fluorescence and their instrumental whiteness was not much affected by the UV content of the illumination. These samples were also judged less white than the other samples under illumination with additional UV.

3.2. Pair-wise comparison

In experiment 3, the observers were asked which sample they experienced as the whitest in each pair of samples. The value 0 was assigned to pairs for which an observer judged the left sample as the whitest and the value 1 was assigned to pairs for which an observer judged the right sample as the whitest. The
The mean response of the 20 observers was then computed to give the proportion, \( p \), of whiteness judgement for each pair and the Just-Noticeable Difference (JND) was calculated with (ISO 20462-3, 2003)

\[
\text{JNDs} = \frac{12}{\pi} \sin^{-1}\left(\sqrt{p}\right) - 3. \quad [1]
\]

The JNDs are plotted versus the difference in CIE whiteness measured under the overhead illumination with a correlated colour temperature of about 5000 K in Figure 4. The JNDs are highly correlated to the CIE whiteness \( R^2 = 0.84 \), despite the fact that the samples had different mean gloss levels. This confirms the results from the magnitude estimation experiments and suggests that the mean gloss level has a negligible effect on the perceived whiteness. The CIE whiteness equation is suitable for comparing these samples when applied to tristimulus values measured in the illumination used for the visual evaluation. From the figure, one JND can be determined to be 4 whiteness values.

![Figure 4. JNDs versus the difference in CIE whiteness measured in the real illumination for the samples in the pair-wise comparison. The line shows the linear fit.](image)

### 3.3. Instrumental prediction of perceived whiteness

The CIE recommends two illuminants and standard observers for the determination of whiteness, D65/10° (outdoor whiteness), and C/2° (indoor whiteness). The whiteness is calculated from the tristimulus values in the defined illuminant. Measurement instruments on the other hand use real illuminations designed to fit the standard illuminant. Spectrophotometers with a d/0° geometry used in the paper industry are UV calibrated to get the same amount of UV as in the D65 or C illuminant but the power spectral distribution of the illumination is not well defined. Densitometers used in the graphical industry lack UV calibration methods although they report values as being measured with e.g. a D50 illuminant.

Moreover, real office illuminations such as the fluorescence light tubes in the overhead illumination used in the visual evaluations have spectral power distributions that differ much from the C or D65 illuminants. The typical spiky spectral power distribution of the light tubes with a 5000 K correlated colour temperature is compared to standard illuminants in Figure 5. Besides their spiky nature, common 5000 K or 6500 K fluorescent light tubes illumination have much lower UV content than the D65 or C illuminants. It is important to note that the C illuminant used for predicting indoor whiteness contains much more UV than typical indoor illuminations (outdoor through-window light excluded). The addition of a UV fluorescent light tube increased the UV content to a level close to that of C or D50 illuminants.

![Figure 5](image)
Because the samples in this study had different amounts of fluorescent whitening agents (FWA), the differences in their measured whiteness was strongly dependent on the UV amount of the illumination. Figure 6 shows the perceived whiteness rating of the samples under the overhead illuminations with and without additional lamp versus CIE indoor whiteness. There are clearly three groups of samples, with low, medium, and high FWA content. Compared to Figure 2 (left), where the whiteness was measured under the real illumination, the indoor whiteness fails totally in predicting the perceived whiteness under 5000 K CCT light tube illumination. Clearly, the higher UV amount in the measurement overestimates the perceived whiteness of the samples under the illumination without additional UV lamp. On the other hand, the correlation between indoor whiteness and perceived whiteness was high ($R^2 = 0.85$) when the visual evaluations were made under 5000 K CCT illumination with additional UV lamp (Figure 6, right).

The results obtained with other instrument/illuminations are summarised in Table 2. Non surprisingly, the outdoor whiteness performs even worse than the indoor whiteness for the visual evaluation without additional UV lamp, due to the higher UV content in the measuring illumination. The densitometer, with unknown UV content in the illumination, gave similar results, meaning that the UV content in the densitometer is much higher than that of common overhead light tube illumination.

Figure 6. Rating of the samples compared to the reference versus indoor CIE whiteness measured with a ColourTouch spectrophotometer. Left: Visual evaluation made under the 5000 K overhead illumination. Right: Visual evaluation made under the 5000 K overhead illumination with additional UV lamp.
Table 2. Correlation coefficients ($R^2$) between visual evaluation of perceived whiteness under the two real illuminations and the measured whiteness from different instruments and illumination/standard observer conditions.

<table>
<thead>
<tr>
<th>Instrument, illumination, and standard observer</th>
<th>$R^2$ Rating under 5000K</th>
<th>$R^2$ Rating under 5000K + UV</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spectroradiometer 5000K/10° UV(5000K CCT tube)</td>
<td>0.84</td>
<td>0.54</td>
</tr>
<tr>
<td>Spectroradiometer 5000K/10° UV(5000K CCT tube) + UV (UV tube)</td>
<td>0.12</td>
<td>0.87</td>
</tr>
<tr>
<td>ColourTouch, outdoor whiteness (ISO 11475, 2004) D65/10° UV(D65)</td>
<td>0.05</td>
<td>0.80</td>
</tr>
<tr>
<td>ColourTouch D50/2° UV(C)</td>
<td>0.33</td>
<td>0.84</td>
</tr>
<tr>
<td>ColourTouch, indoor whiteness (ISO 11476, 2000) C/2° UV(C)</td>
<td>0.31</td>
<td>0.85</td>
</tr>
<tr>
<td>ColourTouch C/2° UV$_{ex}$ (420nm)</td>
<td>0.83</td>
<td>0.14</td>
</tr>
<tr>
<td>TechKon Densitometer D50/2° no pol. filter</td>
<td>0.53</td>
<td>0.80</td>
</tr>
<tr>
<td>TechKon Densitometer D50/2° with pol. filter</td>
<td>0.29</td>
<td>0.72</td>
</tr>
</tbody>
</table>

4. Conclusions

Commercial paper samples with mean gloss levels ranging from 11 to 95% were printed to target $L^*a^*b^*$ values in order to study the effect of mean gloss level on perceived whiteness. The D65/10° CIE whiteness of the samples varied from 120 to 155. The samples varied not only in gloss but also in their amount of fluorescence whitening agents (FWA). Thus, the instrumental whiteness of the samples, and their relative whiteness difference varied greatly depending on the measurement illumination.

The copy paper substrate was calendered to raise the mean gloss level while keeping the FWA amount constant. The gloss increased only from 11% to 18%. Another way of increasing gloss was to laminate an inkjet paper with a transparent film but UV absorption in the transparent film prevented a thorough analysis of the effect of measurement geometry and gloss trap on the measurement of whiteness from samples with different mean gloss levels. It is very difficult to produce samples with different gloss levels and keep all other optical properties constant. Nevertheless, this study suggests that instrumental whiteness is not significantly affected by the measurement geometry and sample gloss, unlike for saturated colours where larger effects would be expected.

Densitometers and spectrophotometers with different UV contents in the illumination will on the other hand rank samples with different FWA amounts differently. The CIE whiteness stipulates a D65 (outdoor) or C (indoor) illuminant. These theoretical illuminations are difficult to generate in measuring instruments, and impossible to produce in a viewing environment large enough for visual evaluation of large sample sets. Moreover, paper and prints are often inspected and used in an office environment under light tube illuminations far from the theoretical C or D65 illuminants. Common overhead illumination with 5000 K correlated colour temperature was revealed to contain much less UV than the D50 illuminant and most office environments are likely to have less UV than the “indoor” C illuminant. Defining the right choice of illuminant for the measurement of whiteness is out of the scope of this report but it is crucial to use the same (real) illumination in the measurement and visual evaluation when comparing the perceived whiteness of samples with different amounts of FWA. Previous ranking of commercial copy papers was not affected by the illumination (Coppel, 2007) because all samples had a relatively high FWA content. Predicting the perceived whiteness from samples with different amounts of FWA requires the illumination to be the same in both prediction (measurement) and visual evaluation (from real observers).
Using the same illumination in both measurement and visual evaluation, a high correlation was obtained between predicted and perceived whiteness ($R^2 = 0.84$ with 5000 K CCT illumination and $R^2 = 0.87$ with 5000 K CCT and additional UV lamp). The CIE whiteness equation applied to the real illumination predicted well the perceived whiteness of samples having a large difference in mean gloss level, both when the samples were laid on the table and when the observers had the samples in their hands and were free to look at them in any direction. This reveals that the mean gloss level does not have a significant impact on the perceived whiteness. This result was confirmed by a pair-wise experiment, from which the whiteness just-noticeable difference (JND) of a group of observers was determined to be 4 whiteness values with whiteness ranging from 90 to 110 under the 5000 K CCT illumination.

This work shows that the CIE whiteness equation can predict well the perceived whiteness of samples with different amount of FWA when the same illumination is used in the measurement and the visual evaluation. Observers were able to rate the whiteness of papers with very different mean gloss levels and their response was well predicted by the CIE equation.

Acknowledgements

This work is part of a consortium project on the appearance of paper and on print quality. Financial support from Holmen, M-real, Norske Skog, and Stora Enso is gratefully acknowledged. Siv Lindberg, Innventia, is acknowledged for numerous discussions. Per Edström, Mid Sweden University, and Caisa Andersson, Karlstad University, are thanked for comments on the manuscript.

References


ISO 11475, (2004), Paper and board - Determination of CIE whiteness, D65/10° (outdoor daylight).

