Effect of sleeve colour and background colour on change in colour assessments†

Coloration Technology

Society of Dyers and Colourists

S Westland,[a] C Graham,[b,‡] S Addison,[a] P Sharrott[c] and B Rigg[a]

[a] *Colour & Imaging Institute, Kingsway House, Derby University, Derby DE22 3HL, UK*

[b] *Society of Dyers and Colourists, Perkin House, 82 Grattan Road, Bradford BD1 2JB, UK*

[c] *Technicare Services Ltd, Bedford Road, South Wigston, Leicester LE18 4TS, UK*

The two standard grey scales, ISO 105-A02 and ISO 105-A03, used for assessing change in colour and staining, respectively, are both typically provided mounted in black material and within a black sleeve. However, some grey scales are now provided mounted in grey material and within a grey sleeve. This paper reports psychophysical data and shows that the change in colour grades obtained from grey scales are not affected by the background colour of the viewing cabinet. A small, but statistically significant, effect is found for the colour of the sleeve with higher grades being found when using the black sleeve compared to when using the grey sleeve. However, when the samples being assessed are masked by a sleeve of the same colour as that used with the grey scale, an extensive field trial shows that the colour of the sleeve does not affect the change in colour grades obtained under commercial conditions.

Introduction

The fastness characteristics of textiles are traditionally assessed by treating the material in a standard way and then visually comparing the colour difference between the treated and untreated materials with the difference between pairs of colour chips in a standard 'grey scale'. Although instrumental methods can also be used [1–3], visual methods are still extremely common and the existence and use of grey scales in industry is important.

In the textile industry there are two standard grey scales, ISO 105-A02 and ISO 105-A03, for assessing change in colour and staining, respectively. Both of these grey scales are typically provided mounted in black material and within a black sleeve as produced by the Society of Dyers and Colourists (SDC). However, some grey scales are now provided mounted in grey material and within a grey sleeve as provided by Deutsches Institut für Normung (DIN). This paper reports experiments to determine whether the assessments of change-in-colour that are obtained using grey scales are affected by the colour of the sleeve. Data are included from psychophysical experiments conducted in the laboratory and from comparative field trials of the grey scales at commercial test houses. Some supplementary data are provided on the effect of the lightness of the background of the viewing cabinet on change-in-colour grades.

† This publication is sponsored by the Society's Colour Measurement Committee

‡ Present address: BTTG-Wiratec, Wira House, West Park Ring Road, Leeds LS16 6QL, UK. Email: carol.graham@bttg.co.uk

Experimental

Physical grey scales

A set of nine pairs of achromatic chips are used to assess change in colour (ISO 105-A02:1993) providing fastness grades 1, 1/2, 2…5, where 5 represents no change in colour and 1 represents the maximum change in colour on the scale. For each pair, one of the chips is always a fixed grey and the colour difference between the pair representing grade 5 is nominally zero. The other chip becomes progressively lighter so that the maximum colour difference is evident in the pair representing grade 1 [1]. A separate set of grey chips are also used to assess staining (ISO 105-A03:1993).

The scales with a black sleeve will be referred to hereafter in this paper as the 'SDC scales' and the scales with the grey sleeve will be referred to as the 'DIN scales'.

Psychophysical study

A set of psychophysical experiments were conducted at Keele University to investigate the effect of sleeve colour and background colour on fastness grades obtained for a set of fabric samples.

A set of 18 pairs of coloured fabric was selected from a group of fabric samples provided by Dr Roger Wardman of Heriot-Watt University. The pairs were selected so that in each pair there were two samples of similar colour but the colour differences exhibited by the pairs represented approximately evenly the range of colour differences in the grey scale. The pairs were also selected to provide a good range of different colours. The samples (approx. 10 × 10 cm) were mounted on card in order to ensure a uniform and consistent presentation and for ease of handling. The card was provided for structural support only and was not visible during the experiments. A sample of 23 naive observers (male and female, aged 18–50) with good colour vision (tested using the Ishihara test plates) were selected

Table 1 Summary of conditions used in the field trial by the professional assessors A–J

Assessor	Conditions: Cabinet colour	Cabinet dimensions (mm)	Light source	Life history of tube (h)	Viewing geometry
A,B	Munsell N5 Grey	680 × 360 × 500	Thorn D65	200	Lecturn (45/0)
C,D	Smokey Pine	680 × 357 × 390	VeriVide	0	Lecturn (45/0)
E–G	Munsell N5 Grey	680 × 360 × 380	VeriVide	180	Lecturn (45/0)
H–J	Smokey Pine	710 × 575 × 405	VeriVide	70	Lecturn (45/0)

Table 2 Mean scores under the six conditions for each observer averaged over the 18 samples in the Keele psychophysical study

	SDC			DIN		
Observer	Black	White	Grey	Black	White	Grey
1	2.50	2.25	2.33	2.14	2.11	2.08
2	2.53	2.28	2.36	2.31	2.25	2.31
3	3.42	3.64	3.03	3.17	3.31	3.19
4	3.69	3.67	3.61	3.19	3.14	2.94
5	2.72	2.44	2.64	2.28	2.22	2.25
6	2.72	2.25	2.22	2.36	2.42	2.31
7	3.00	3.19	3.43	2.75	3.08	2.86
8	2.61	2.33	2.53	2.28	2.17	2.33
9	2.44	2.47	2.42	2.25	2.19	2.42
10	2.69	2.75	2.78	2.58	2.75	2.53
11	2.53	2.47	2.97	2.78	2.86	2.69
12	2.42	2.75	2.53	2.44	2.58	2.31
13	2.83	2.47	2.81	2.19	2.53	2.39
14	2.78	2.75	2.89	2.28	2.19	2.17
15	1.97	1.86	1.83	1.72	1.78	1.75
16	2.72	2.89	3.03	2.44	2.92	2.72
17	2.44	3.58	3.36	3.36	3.67	3.58
18	2.75	2.86	3.03	2.86	2.97	2.92
19	2.47	2.56	2.36	2.36	2.25	2.31
20	2.97	3.31	3.19	2.97	3.06	2.97
21	3.53	3.81	3.47	3.17	2.97	3.11
22	3.17	2.53	3.28	2.83	2.56	2.83
23	2.72	2.72	2.64	2.25	2.50	2.36
Mean	2.77	2.78	2.81	2.56	2.63	2.58

for the experiment. A Macbeth Spectralight II viewing booth illuminated with artificial daylight (D65) was fitted with removable viewing surrounds allowing for three general background conditions: white, black and grey. The SDC and DIN grey scales were used by the observers to report fastness grades for each pair of samples. The samples were arranged flat in the viewing booth such that the illumination was normal to the plane of the surface of the samples in the viewing booth and the observer viewing angle was approx. 45°. The experiments were conducted in a darkened room.

The sample pairs were presented in random order, against one of the background conditions (white, black or grey), and the observer was asked to assess the colour difference exhibited by each pair in relation to the grey scale (SDC or DIN). Observers reported fastness grades according to the grey scale, i.e. 5, 1/2, 3, etc. When all 18 samples had been assessed the procedure was repeated with a different grey scale but with the samples being presented in the same order as previously. This whole procedure was repeated until the samples had been assessed using both grey scales and all three backgrounds resulting in six assessments from each observer for each pair. Each observer was allowed a period of adjustment of 3 min whenever the background was changed so that they could be light adapted.

Field trial

A field trial of the SDC and DIN grey scales was carried out with the assistance of Technicare Services Ltd who provided and prepared the samples for assessment.

A set of 11 fabric samples were assembled and subjected to wash fastness tests. The original and washed samples were mounted on card as pairs for use in the change-in-colour assessments. A multi-fibre strip was included in the wash fastness tests and this was used to provide samples for staining assessments on acetate, cotton, nylon, polyester, acrylic and wool. (The data from the wash fastness trials are not reported here and will be presented in a subsequent publication.)

Although the sleeve of the DIN scale is grey it is surprisingly not a close match to Munsell N5. Munsell N5 card was obtained and special sleeves were made so that three sleeve colours could be used in this investigation: black (SDC), grey (DIN) and Munsell N5. Packs containing the fabric samples and grey scales with each of the three sleeves were assembled and distributed to eight commercial test houses, however only four of these returned any data. The packs included a protocol for the experiment and indicated that the samples being assessed should be masked using a sleeve of the same colour as that used with the grey scale. Assessors were instructed to perform one or two repeat assessments if possible. The pack was accompanied by a questionnaire to ascertain various parameters used by each test house including the colour of the viewing booth and the make and age of the lamps. Data from ten assessors representing four of the test houses were returned to Derby University where it was then analysed. Table 1 summarises the conditions that were used by these professional assessors. Two of the test houses used a viewing booth painted using Munsell N5 and two used a booth painted using Smokey Pine. However, the effect of booth colour was ignored in the field trial due to the results obtained from the Keele study (see next section). Assessors A–J were therefore treated as a single population.

Results and Discussion

Psychophysical experiments to assess change in colour

Table 2 shows a summary of the results obtained for the change in colour study carried out at Keele. For each condition, the grade for each of the 23 observers is reported as a numerical value (for statistical analyses grades were converted into numerical value; thus the grade 3/4 was converted to the numerical value 3.5) averaged over all samples.

The mean fastness grades in Table 2 appear to be consistent across all background colours but in general the grades obtained using the SDC scale are higher than those obtained using the DIN scale. A series of two-tailed paired t-tests were carried out on the data shown in Table 2 and these confirm that there are no significant effects of background colour but significant differences are found for sleeve colour (black, $p = 0.0063$; white, $p = 0.0144$; grey, $p = 0.0001$). The values of p indicate the probability that the difference could have arisen by chance and are always less than 0.05 for the effect of sleeve colour. Thus the grades obtained using the SDC (black) sleeve are significantly higher that those obtained using the DIN (grey) sleeve for all three background colours.

From Table 2, the differences between the means of the two scales for the black, white and grey backgrounds are 0.21, 0.15 and 0.23, respectively. Note that although these differences are statistically significant (in the case of the grey background significant at the 1% level) they represent differences of less than half of the step size used in the assessment scale. It is also important to note that these differences are small when compared with the difference between observers. So, for example, even if observer 15 (whose grades seem very low) is ignored there are intra-observer differences of more than two steps in the grey scale.

Table 3 Mean scores under the three conditions for each assessor averaged over the 11 samples in the field trial

	Sleeve		
Assessor	SDC	DIN	N5
A1	4.50	4.50	4.50
A2	4.50	4.50	4.50
B1	4.64	4.73	4.73
B2	4.64	4.73	4.73
B3	4.64	4.73	4.73
C1	4.23	4.23	4.23
C2	4.23	4.23	4.23
C3	4.23	4.23	4.23
D1	4.27	4.05	4.09
D2	4.32	4.09	4.18
D3	4.14	4.18	4.09
E1	3.73	3.64	3.64
F1	4.18	4.14	4.18
G1	4.14	4.09	4.09
H1	4.14	4.14	4.18
H2	4.18	4.14	4.14
H3	4.09	4.09	4.09
I1	4.18	4.14	4.18
I2	4.18	4.18	4.18
I3	4.23	4.18	4.18
J1	4.14	4.14	4.14
J2	4.14	4.14	4.14
J3	4.14	4.14	4.14
Mean	4.25	4.23	4.24

However, there are several reasons for exercising caution in accepting the finding that fastness grades depend upon sleeve colour (for the Keele data). Firstly, the observers were naive and examination by professional assessors using different sleeve colours may produce different results. Secondly, the conditions under which the Keele study was conducted were rigorously controlled but differ in certain respects from those that are commonly employed in industry. For example, the viewing condition was normal illumination and 45° viewing, whereas in industry it is normal to position the samples on a lectern at an angle of 45° and view at an angle of 90° to the plane of the surface of the samples. Thirdly, the history of the samples used in the Keele experiment was not fully known. This might be important; for example, on washing, samples will tend to lose dye and therefore exhibit mainly changes in L^* and C^* with very little change in hue. The Keele samples were not biased to reflect this and the average magnitude of differences in lightness, chroma and hue were approximately the same. Fourthly, although the Keele data used samples that covered the whole of the grey scale approximately uniformly this might not be typical in industry where, for example, more higher grades than lower grades might be expected. Finally, the samples being assessed were not covered with a sleeve whereas it is good practice in industry to view the samples through a sleeve in the same way that the grey scale is viewed.

For these reasons a field trial of the SDC and DIN grey scales was also conducted.

Field trial analysis of sleeve colour on fastness grades for change in colour

Table 3 shows the mean grades for the change in colour returned by each of the assessors A–J in the field trial averaged over the 11 samples using each of the three sleeve conditions. Repeat assessments are shown for assessors where appropriate. For example, the rows headed A1 and A2 refer to two separate assessments made by assessor A on two separate days. The total number of separate assessments (including repeats by the same assessor) made on each of the samples was 23.

In general assessors were reliable when making repeat assessments and there is generally good agreement between assessors, although assessor E appears to grade consistently lower than the other assessors. The mean scores averaged over all assessors are very similar (SDC, 4.25; DIN, 4.23; N5, 4.24) for all three sleeve colours and no significant effects of sleeve colour are found even at the 10% level. The removal of observer E from the analysis did not generate any significant effects.

Several possible reasons could be postulated for the apparent discrepancy between the two findings including (a) the experience of the observers, (b) the viewing conditions or (c) the nature of the samples. We suggest that the viewing conditions are primarily responsible for the differences obtained in the two related studies. In the Keele trial the samples were viewed without any masking whereas the grey scale was viewed through the sleeve. There was therefore a size difference between the samples being assessed and the grey scale chips. Furthermore, it is known that the colour of the immediate background can affect the size of the colour difference apparent between two samples. The field trial study was conducted using viewing conditions appropriate for the way in which samples are assessed professionally. In the field trial the samples were always masked using a sleeve of the same colour as that used with the grey scale. Under these conditions the effect of sleeve colour completely disappeared.

Conclusions

The psychophysical experiments conducted with naive observers and the field trial involving professional assessors provide somewhat contradictory findings. The Keele experiment suggests that the use of a black sleeve results in significantly higher grades being obtained than when the grey sleeve is used. However, the difference in grades obtained is less than half a step on the grey scale and is small compared with the difference found between observers. Indeed, even the professional assessors exhibit intra-observer differences of almost two steps in the grey scale. The data from the field trial suggest no significant effect of sleeve colour whatsoever.

The data from Keele are convincing that there is a small effect of sleeve colour in the grades obtained from visual assessments of change in colour. However, this difference is small, even for the large colour differences used, and of no practical significance when typical samples are assessed by experienced assessors under commercial test conditions. Further work is required to consolidate this finding since, although the lack of masking of the samples caused the effect of sleeve colour in the Keele study, there were in fact several differences between the Keele study and the field trial and the significance of these differences requires examination. There seems little reason, given this study, to distinguish between grey and black sleeves when conducting commercial assessments of change in colour. It does seem important, however, to mask the samples being assessed with a grey sleeve of identical colour to that used with the grey scale whether this be grey or black.

It is important to note that this study does not suggest that the colour appearance of a sample, or even the perceived colour difference between two samples, is invariant to the background colour. Indeed, there is clear evidence to the contrary. For example, it has been established that the colour appearance of a sample depends upon the background [4] and this effect is accounted for in recent colour-appearance models such CIECAM97 [5]. Also, it has been long known that the perceived lightness difference between two samples is greatest when the lightness of the samples is close to the lightness of the background [6]. This phenomenon is known as the crispening effect [7] and can also occur for chromatic stimuli [8].

* * *

This work was supported by the Colour Measurement Committee (CMC) of the SDC. The authors would like to thank members of the CMC for their helpful advice during the experimental and analysis stages of the research.

References

1. B Rigg, *J.S.D.C.*, **107** (1991) 244.
2. T Sato, N Takada, N Ueda, T Nakamura and M R Luo, *J.S.D.C.*, **113** (1997) 17.
3. T Sato, N Ueda, T Nakamura and M R Luo, *J.S.D.C.*, **113** (1997) 356.
4. M R Luo, X W Gao and S A R Scrivener, *Col. Res. Appl.*, **20** (1995) 18.
5. M R Luo, *Rev. Prog. Color.*, **30** (2000) 77.
6. C C Semmelroth, *J. Opt. Soc. Am.*, **60** (1970) 1685.
7. H Takasaki, *J. Opt. Soc. Am.*, **56** (1966) 504.
8. G Cui, *Colour-difference evaluation using CRT displays*, PhD Thesis, Derby University (2000).