How Well Can People Predict Subtractive Mixing?

P. M. Henry, S. Westland, S. M. Burkinshaw and T. L. V. Cheung, Centre for Colour Design Technology, University of Leeds, Leeds, United Kingdom

Abstract

This study is concerned with the design of effective colour tools to allow users to quickly and accurately select a given colour in a digital-display environment. It has been shown that the choice of colour space (for example, RGB colour space compared with a more perceptually relevant space) influences performance (speed and accuracy) in certain colour-related tasks.[1,2] We suggest that the nature of the colour-mixing model may also be a factor in certain tasks such as the selection of a target colour from a colour-selector tool. It is our hypothesis that users have a more accurate internal model for how subtractive colour mixing works than for additive colour mixing. The purpose of this work is to determine whether it is indeed the case that observers possess better internal models for subtractive colour mixing than for additive colour mixing. In Experiment 1 the variance in observers' abilities to predict the result of subtractive colour mixing is compared using real physical samples and using a computer monitor (CRT). Although the variance obtained on the CRT was greater than that obtained using the physical samples, the difference was not statistically significant. In Experiment 2, the abilities of observers to predict subtractive and additive mixing were directly compared using samples displayed on a CRT. Observers' abilities to predict additive mixtures were not as good as their abilities to predict subtractive mixtures ($p < 0.05$).

Introduction

This study is concerned with the design of effective colour tools that allow users to quickly and accurately select a given colour in a digital-display environment. It has been shown that when asked to match colours using software controls based on RGB, LCH and LRGYB (Lightness, red/green, yellow/blue), observers take the longest time when using the RGB controls.[1] The RGB control was also found to be the worst in terms of accuracy of the matches. We have found that some non-expert users find the RGB colour space to be non-intuitive and this has been confirmed formally elsewhere.[1] In some cases the level of expertise of the observer also influenced performance.[1] The choice of colour space on various visual tasks, such as colour naming, has been shown to be an important factor.[2] It is our hypothesis that observers find manipulation and selection in an RGB colour space to be difficult because they do not possess an appropriate internal model for additive colour mixing. On the other hand, observers may develop a useful internal model of subtractive colour mixing processes from a young age as they experiment with inks and paints. Many observers, particularly colour naives, would not be surprised to be informed that yellow and blue inks mixed together make green but may find it hard to believe that red and green lights can be added together to make yellow. The purpose of this work is to determine whether it is indeed the case that observers possess better internal models for subtractive colour mixing than for additive colour mixing.

The ability of observers to make predictions about subtractive and additive colour mixing has been evaluated in this study. Although accuracy has been measured, greater emphasis has been placed upon the consistency (or precision) of predictions by experts and naïve observers. The experimental paradigm chosen to investigate observers' ability makes use of colour matching.

Experimental

In each experiment 12 observers (6 designated as expert and 6 designated as naïve) were asked to predict the colour that would result from either a mixture of two physical paint specimens or two colour stimuli displayed on a computer monitor. The expert observers were either academics or students in the fields of textile, fashion or graphic design in the School of Design at Leeds University and therefore had experience in a colour-critical industry. The naïve observers had no professional interest in colour.

In the Experiment 1, observers were presented with pairs of samples (made from acrylic paint) and asked to select the colour that most closely matched their expectation of an equal (by weight) mixture of the two paints from a library of coloured samples. The library of colour samples was created using a HP8550 laser jet printer and the colours were specified according to Ned Graphics Printer Atlas. The CIE XYZ values (D65 illuminant; 1964 CIE standard observer) were measured for the selected matches using a Minolta CM2600 reflectance spectrophotometer. In order to be able to differentiate between variance caused by differences between observers' internal models for subtractive mixing and inherent variance caused by the task of selecting a match from an atlas to a sample on a different medium (painted rather than printed) observers were also asked to select a match in the printed atlas to certain paint samples.

A different set of observers were then asked to perform the mixture-prediction task where both the samples and the colour atlas were displayed on a computer monitor. The monitor was characterized using the GOG model[3] to enable the colours displayed to similar to the physical samples. The purpose of Experiment 1 was to ascertain whether the variance of predictions obtained with physical samples for the mixture-prediction task would be the same as the variance obtained using a monitor simulation of the physical samples.

Figure 1 illustrates how the on-screen matching was performed. Observers were asked to consider each pair of colours in turn (for example, the cyan and magenta pair shown in Fig. 1) and to select a colour from the colour atlas (right-hand side of Fig. 1) that represents the observer's imagined mixture of the two samples.

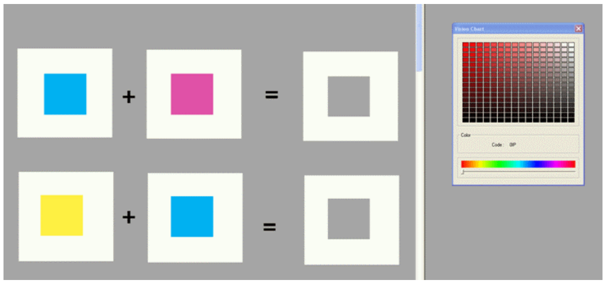

Figure 1. Partial screen-shot from CRT experiment of subtractive colour mixing (Experiment 1).

In Experiment 2 observers were presented with pairs of colour stimuli on a computer monitor (in a configuration identical to Fig. 1) but were instructed to predict the result of an additive mixture of the two stimuli. Prior to the experiment observers were presented with a graphical example of additive colour mixing and the concept was explained to them. For each set of matches (or trials) an ellipse was fitted to the data in the CIELAB a*-b* plane so that the major and minor axes of the ellipse was equal to the standard deviations along the major and minor directions, these directions being determined from a singular-value decomposition of the data (Fig. 2).

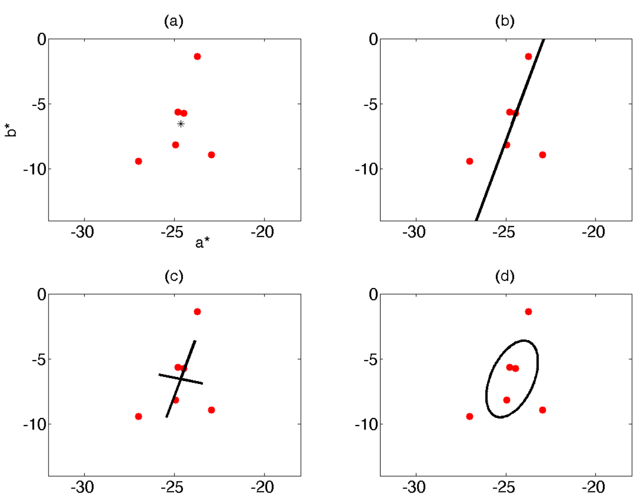

Figure 2. Example analysis showing (a) original matches made by six observers (the asterisk is the mean of the six matches); (b) the orientation of the first component derived from SVD; (c) lines oriented along the principal and orthogonal directions whose lengths represent the standard deviations of the points along those directions; (d) ellipse representing the distribution of points.

Results

A colour representation of the results for three of the trials from Experiment 1 is shown in Fig. 3.

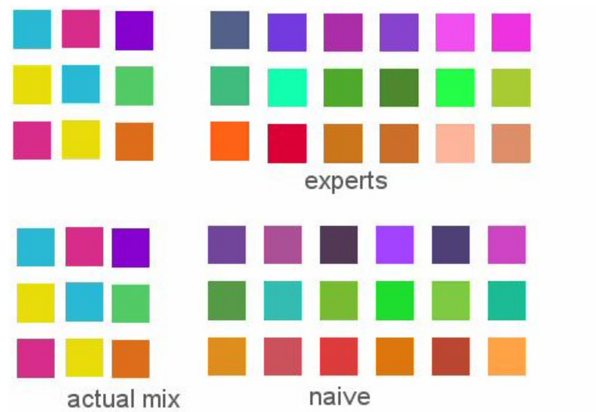

Figure 3. Colour representation of matches (Experiment 1) made for binary mixtures of cyan, magenta and yellow. Each column shows the match made by an observer (experts in top half; naives in bottom half) to the imagined mixture of the colours in 1st and 2nd columns. The colour in the third column is the actual physical mixture (which the observers did not have access to).

Figure 4 shows the ellipse representation for the matches illustrated in Fig. 3 from which it appears that the performance of the two groups of observers are broadly similar. However, the variance in matching was quite large. In order to ascertain the cause of this variance observers were also asked to select samples from the colour atlas that were a match to specific physical samples. The variance of these matches was analysed in the same was as the variance of the predicted subtractive mixtures and the resulting ellipses are shown in Fig. 5. The size of the ellipses in Fig. 5 illustrates the error inherent in the matching paradigm itself. The difference in size between the ellipses in Fig. 4 and those in Fig. 5 can be attributed to variance in or between observers' internal models for subtractive colour mixing.

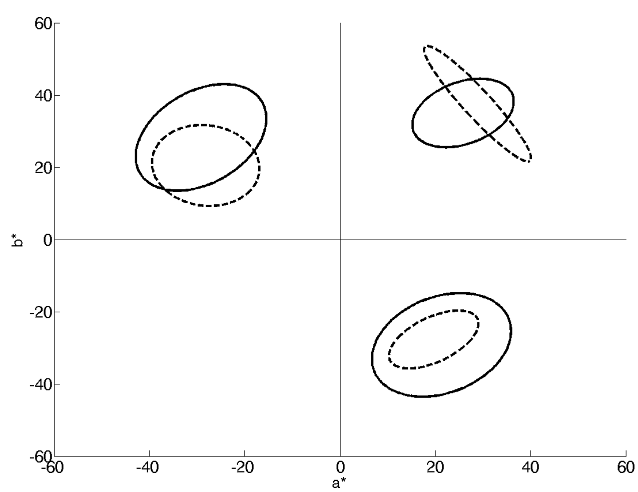

Figure 4. Distribution of matches made by expert (solid lines) and naïve (dashed lines) observers for mixture of subtractive primaries (Experiment 1).

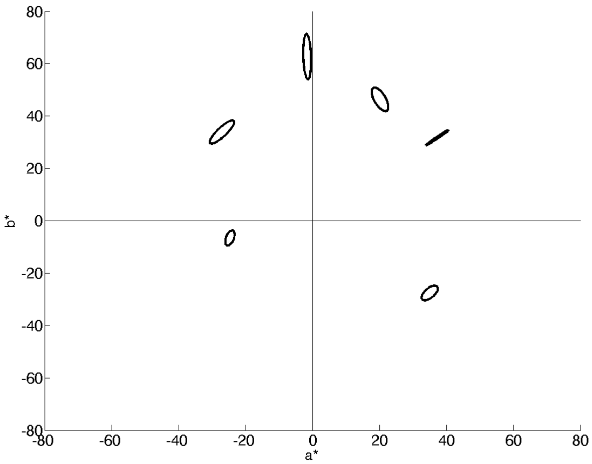

Figure 5. Distribution of matches made by expert observers to individual physical samples.

To enable a more quantitative measure of the variance in predicting the subtractive mixture, for each trial the CIEDE2000(2:1) colour difference was computed between each observer's match and the mean of all observers' matches for that trial. The variance in matching for each trial is then represented by the mean of those colour differences. The statistics of these mean colour differences for the three trials illustrated in Fig. 4 and for all the trials in Experiment 1 are shown in Table 1. It is evident from Table 1 that whereas for the case where three binary mixtures of cyan, magenta and yellow are concerned the performance of naïve and expert observers was broadly similar. When all nine mixtures were considered, however, the maximum variance for the naïve observers was almost twice that of the expert observers. Nevertheless, a Wilcoxon Matched-Pairs Signed-Ranks test[4] revealed no statistically significant differences between the mean scores for experts and naïves ($p > 0.05$).

Table 1: CIEDE2000(2:1) Colour Differences for Experiment 1 for Expert Observers (Naïve Observers in Parentheses)

	mean	median	max
Experiment 1 (only primary mixtures)	9.80 (10.49)	8.79 (11.15)	12.28 (11.29)
Experiment 1 (all mixtures)	8.54 (11.65)	8.39 (11.15)	12.28 (21.79)

In Table 2 the variance in matches is illustrated in terms of ΔE distributions for Experiments 1 and 2. The variance for the subtractive-mixture prediction was greater for samples simulated on a CRT display than for physical samples. However, the difference between the mean values was not statistically significant matching ($p > 0.05$). This would seem to indicate that the validity of asking observers to predict subtractive mixtures of samples displayed on a CRT.

Table 2: CIEDE2000(2:1) Colour Differences for Experiments 1 and 2 for Expert Observers

	mean	median	max
Experiment 1 (physical samples)	8.54	8.39	12.28
Experiment 1 (CRT simulation)	12.31	13.21	20.54
Experiment 2 (additive mixing)	15.02	14.80	22.81

The final row in Table 2 shows the variance in matches when observers were asked to predict the result of additive mixture of pairs of samples displayed on a CRT (Experiment 2) and should be compared with the subtractive data in the previous row. Our hypothesis was that the variance of predictions for additive mixing would be greater than those for subtractive mixing if observers possess a more accurate internal model for subtractive mixing than they do for additive mixing. The data in Table 2 support our hypothesis and the difference between the variance for additive and subtractive mixing was found to be statistically significant ($p < 0.05$).

Conclusions

The variance in matches between observers who are asked to predict the results of subtractive mixing were much larger than the variance between observers who were asked to simply match a sample. This is not surprising when one considers the complexity of subtractive mixing. The subtractive-mixture prediction task is, in fact, an example of a task with a non-unique solution. Two physically different yellow paints, for example, may look visually identical when applied as an opaque layer but then produce quite different results when mixed with another colour paint. This lack of uniqueness is not the case for additive colour mixing however. Grassman's second law of colorimetry dictates that if $A \equiv B$ and $C \equiv D$ (where the symbol $\equiv$ indicates a visual match) then the additive mixtures $A + C$ and $B + D$ will be a visual match.[5] Therefore, it could be argued that observers would require a more complex internal model of subtractive mixing than of additive mixing if the performance (expressed in terms of this work in terms of variance between observers' matches) for the two tasks was identical.

In this study the variance between observers' predictions of additive mixtures was greater than that for predictions of subtractive mixtures. This would appear to support our hypothesis that observers possess a more accurate model of subtractive mixing than additive mixing (presumably based on their greater practical experience of the former task). Thus, even though the subtractive task is more difficult, observers perform it better.

There are two caveats to the finding that the data support our hypothesis. The first is that, strictly, we did not find that observers possess a more accurate internal model of subtractive mixing than they do of additive mixing. This is because we did not find that observers were more accurate at predicting subtractive mixture.

Rather, what we found was that observers were more consistent between themselves at predicting subtractive mixtures. The accuracy of the matching is still being analysed but is problematic because of the non-unique nature of the subtractive-mixing process.

The second caveat is a concern about whether the observers in Experiment 2 actually made additive predictions at all. Indeed, an analysis of the L* values of the predicted mixtures from the subtractive task and the additive task showed that, on average, the L* values from the additive tasks were not statistically higher than those from the subtractive task. Further work is required in this area to confirm our hypothesis.

The importance of choosing the most appropriate colour space for an application has been understood.[1,2,6] This work is motivated by this understanding and is concerned with the development of efficient tools for the selection of colour in computer software. However, we further suggest that the choice of colour mixing model may also be important when designing such software. Thus, allowing users to select and explore colours by mixing together virtual paints on screen may be more effective that the traditional use of additive RGB systems.

References

1. H. Zhang and E.D. Montag, "How well can people use different color attributes", Proceedings of IS&T/SID 12th Color Imaging Conference, 10-17, (2004).
2. H. Yaguchi, "Color categories in various color spaces", Proceedings of IS&T/SID 9th Color Imaging Conference, 6-8, (2001).
3. R.S. Berns, "Methods for characterizing CRT displays", *Displays*, **16**, 173-182, (1996).
4. G. Upton and I. Cook, Understanding Statistics, Oxford University Press, (1996).
5. H.G. Grassman, On the theory of compound colours, Philosophical Magazine, **4** (7), 254-264, (1854).
6. J.M. Taylor, "Color spaces: Language and framework for color", Proceedings of IS&T/SID 1st Color Imaging Conference, 6-10, (1993).

Author Biography

Phil Henry is a Teaching Fellow at the University of Leeds. He is currently working towards a PhD in the Centre for Colour Design Technology at Leeds University.