Double-masked placebo-controlled trial of precision spectral filters in children who use coloured overlays

A. J. Wilkins\(^1\), B. J. W. Evans\(^2\), J. A. Brown\(^1\), A. E. Busby\(^1\), A. E. Wingfield\(^5\), R. J. Jeanes\(^1\) and J. Bald\(^6\)

\(^1\)MRC Applied Psychology Unit, 15 Chaucer Road, Cambridge CB2 2EF; \(^2\)The Institute of Optometry, London; \(^3\)Dollond and Aitchison, Leeds; \(^4\)Upbury Manor School, Gillingham; \(^5\)Dyslexia Institute, Leeds; and \(^6\)11 Oxford Road, Colchester, UK

(Received 9 December 1993, in revised form 21 June 1994)

We selected 68 children who reported benefit from individually chosen sheets of coloured plastic placed upon the page when reading, and who used these regularly without prompting. These children viewed text illuminated by coloured light in an apparatus that allowed the separate manipulation of hue (colour) and saturation (depth of colour), at constant luminance. Many of the children reported improvements in perception when the light had a chromaticity within a limited range, which was different for each individual. A pair of plastic spectacle lenses ('experimental' lenses) was dyed so as to provide the appropriate chromaticity under conventional white (F3) fluorescent light. An additional pair was prepared having very similar colour but with a chromaticity outside the range in which perception was reported to improve ('control' lenses). Each pair was provided for 1 month in random order. The children kept diaries (36 completed) recording symptoms of eye-strain and headache. The children and those responsible for their assessment were unable reliably to distinguish 'experimental' from 'control' lenses. Nevertheless, symptoms were less frequent on days when the 'experimental' lenses were worn ($P < 0.003$).


Helen Irlen reported that some people experience perceptual distortions (e.g. movement of letters, changes in their spacing, and blurring) and symptoms of discomfort when reading, and that the distortions may abate when the text has a particular colour, different for each individual\(^1\). She has developed a system of coloured overlays and coloured lenses as treatment, available at the Irlen Institutes she has founded. The efficacy of this treatment remains controversial\(^1\), although anecdotal reports of success are now numerous\(^4\).

Errors of binocular coordination have been proposed as explanations for the beneficial effects of coloured lenses\(^5\). Nevertheless it is possible that a hypersensitivity to light (photophobia) is responsible for the perceptual distortion, and is related to headaches and eye-strain in a variety of ways\(^7\).

A system for precision ophthalmic tinting has recently been described\(^8\). The system uses an Intuitive Colorimeter\(^9\) that enables subjects to choose any coloured illuminant within a large range (gamut) and then provides coloured glasses that match the selected chromaticity. In open trials using the system, more than 80% of subjects have reported continued use of their tinted lenses when followed up one year later\(^1\).

The present study used a double-masked, placebo-controlled cross-over design to address the following questions: is the benefit solely attributable to a placebo effect; and does the tint need to be individually determined with precision in order to be maximally beneficial. The colorimeter was used to provide two chromaticities, one at which perceptual distortions abated, and another similar setting at which distortions were just apparent. Subjects were given two sets of lenses, one of which matched the optimal colorimeter setting; symptoms of eye-strain and headache were recorded.

The purpose of the study was simply to determine whether any category of subjects can show a benefit from such lenses over and above that from a placebo. The population of people who are likely to benefit from individually prescribed tinted lenses has not been identified. The literature suggests (but without double-masked placebo control) that a subset of children from many different diagnostic groups may sometimes show a benefit. These groups include children with reading difficulties\(^2\),\(^10\), children with ashenopia\(^1\) and children with migraine in the family\(^7\). It was therefore important to include children from these
groups. The proportion of children in these groups who show a benefit from colour is uncertain. Given this uncertainty it was necessary to increase the likelihood of selecting children who benefited from colour and thereby improve the statistical power. To this end we selected only children who benefited from the use of coloured overlays in the classroom. A behavioural index of benefit was adopted: the children were admitted to the study only if they had used a coloured overlay consistently without prompting for at least 3 weeks.

Procedure

Selection of subjects

Children were recruited by teachers in two state comprehensive schools in Essex, a secondary high school in Kent, a private boys' school in Leeds and the Dyslexia Institute in Leeds. The teachers examined children individually. Each child was questioned concerning any perceptual distortion of text, avoiding leading questions. The child was then shown a page of random letters arranged to resemble words in a paragraph of closely-spaced text. The page was covered in turn by each of 10 theatre filters: Rose (Rosco 603); Orange (Rosco 09); Yellow (Rosco 96); Green (Lee 244); Turquoise (Rosco 683); Aquamarine (Rosco 66); Blue (Lee 202); Purple (Rosco 642); Lavender (Rosco 38); Neutral grey (Lee 209). By a process of successive elimination during which any beneficial filters were compared side-by-side the child chose a filter that made the text comfortable to look at and best reduced the perceptual distortions (e.g. movement of letters, changes in their spacing, blurring). The assessment was performed under lighting conditions similar to those normally experienced by the child when reading. All children who reported benefits from an overlay were offered one to use. Only those using it consistently without prompting for a period of at least 3 weeks were eligible for entry to the study.

In three schools the children receiving overlays were selected from those who were failing in reading. In the remaining school the children were selected from those who reported eye-strain, headaches or perceptual distortions of print in a questionnaire circulated to all Year 7 pupils (11–12 years old): 68 children entered the study between January and July 1992.

The children were referred for a full optometric assessment (including assessment of binocular function) at the Institute of Optometry in London (by B.E.) or at Dollond and Aitchison in Leeds (by J.B.). Of the 37 subjects in the final analysis of symptoms, none needed a change in refractive correction and none had changed their correction within 2 months of entry to the study.

Following the optometric investigation, the child was examined using an Intuitive Colorimeter®, an apparatus that allowed the independent control of hue (colour) and saturation (depth of colour) without any associated change in luminance (brightness)³.

Before receiving the first pair of coloured glasses all entrants underwent an initial examination of reading using the Diagnostic Tutor form of the Neale Analysis of Reading (Revised British Edition). The lenses and optometric assessments were provided free of charge. The participants were informed that they were free to choose whether to wear the glasses and when to do so. This was for ethical reasons.

Selection of experimental and control lenses

The selection of experimental lenses closely resembled the 'selection of appropriate chromaticity' described in detail elsewhere¹. Random letters arranged to resemble words in a paragraph of text were illuminated by coloured light in the colorimeter. The subject was asked to describe any perceptual distortion of text or feelings of visual discomfort when, at each of 12 colours (12 hue angles spaced about 30° apart), the saturation of light falling upon the page was increased and then decreased. The saturation was increased over 5 s to CIE 1976 $s_u = 7$ and then decreased to $s_u = 0$ over a similar period. Hues at which distortion or discomfort was exacerbated were subsequently avoided. Those hues at which subjects reported that distortion abated or comfort improved were next explored by increasing the saturation and allowing adaptation to occur. At the optimal saturation, hue was then varied. Once the optimal hue had been obtained, saturation was again optimized. The above observations were obtained at a constant luminance of 23 cd m⁻². They were checked at luminances of 12 and 6 cd m⁻² in order to find the setting that gave maximum comfort.

Usually the optimal hue was selected by a process that involved the examiner providing two successive settings for comparison, but occasionally subjects adjusted the settings for themselves. Once the best setting had been obtained, the hue control was changed progressively by the examiner, first in one direction until distortions occurred and then in the other direction until distortions again occurred. In this way two chromaticities were obtained at which distortion was apparent, each having saturation similar to that of the optimal setting. Of these two sub-optimal chromaticities, that closest to the optimal was selected for the control lenses, provided that this chromaticity had not earlier been associated with aversion, in which case the other chromaticity was used. The optimal chromaticity was used for the experimental lenses.

The experimental and control lenses were dyed using the following procedure. A surface with even spectral reflectance (> 95% between 400 and 700 nm) lit by a 'white' (F3) fluorescent lamp was visible through a circular aperture (20 mm diameter) in the side of the colorimeter. A neighbouring aperture with dimensions similar to the first revealed the subject's viewing surface within the colorimeter, set to the appropriate hue and saturation. In a darkened room various combinations of tinted trial lenses were placed in front of the first aperture until the two surfaces appeared the same colour to an independent observer (A.W.) who had normal colour vision. The luminance of the colorimeter viewing surface was adjusted so that the matching was performed when the surfaces had similar luminance. As a result, the trial lenses had a chromaticity matched to that of the colorimeter setting, independently of the luminance. The trial lenses were dyed using two organic dyes from a selection of seven, according to the methods described elsewhere¹. The spectral transmission of the combination was even and, given the dyes used, was uniquely specified by the chromaticity. An independent laboratory (Cerium Visual Technologies Ltd) used a duplicate set of trial lenses as a standard against which to match the transmission of the spectacle lenses, when dyeing them using identical dyes. Two of the subjects who had uncorrected borderline myopia were given control lenses whose peak transmission was at a longer wavelength than the experimental lenses.
This was so that any preference for the experimental lenses could not be attributed to the myopic refractive error (greater refraction of shorter wavelengths placing longer wavelengths nearer the retina).

In summary, one pair of spectacle lenses (the experimental lenses) provided the 'optimal' colour. The other pair of spectacle lenses (the control lenses) had a colour that allowed perceptual distortion but was not aversive. Note that subjects did not see the trial lenses at any time. The spectacle lenses were selected on the basis of the colorimeter setting. Subjects were told simply that two pairs of glasses were being compared and they were not informed of any connection between the lenses and the colorimetry, which was described as a colour vision test.

The first pair of tinted spectacles (experimental or control, selected at random by the tinting laboratory) were available for a period of at least 1 month, during which subjects kept a daily symptom diary and noted when they wore the spectacles. At some time during the third week reading was assessed by the referring teacher, using the Neale Analysis of Reading Form 1. At the end of the first month the spectacle frames were reglazed with the second pair of lenses (control or experimental) and, after an interval without any coloured filters of at least 2 weeks, a further 1 month period of observation ensued with the second pair of tinted spectacles. During the third week of this observation period the Neale Analysis of Reading Form 2 was administered. The interval between receipt of the first and second lenses varied so that the periods when the spectacles were worn did not include school holidays. After the second pair of glasses had been worn for 1 month the child received a second optometric examination that included tests previously administered, together with a colorimetry examination. After the examination, the child attempted to match each pair of lenses with the colorimeter by setting the chromaticity to resemble the appearance of text when wearing glasses. All subjects then underwent a routine clinical assessment to select an optimal tint and were provided with spectacles free of charge.

The code for the mask was divided between A.W. and Cerium Visual Technologies, neither having knowledge of the identity of the lenses.

The study was approved by the Cambridge Health Authority Local Research Ethics Committee and the Institute of Optometry Research and Ethical Committee.

Results

Characteristics of sample

The average age at first appointment was 12 years 2 months (SD 1 year 9 months, range: 9 years 9 months–15 years 5 months).

Attrition

Of the 68 children (42 boys and 26 girls) who entered the study, 15 did not finish. A further 16 failed to complete the symptom diary, one of whom refused to wear the control lenses. Optometric data were available for 53 children, Neale scores for 45.

Optometry

The optometric characteristics of the subjects will be described in more detail elsewhere. No ocular pathology was detected and none of the subjects had a strabismus or any manifest incomitancy. The only optometric reasons for excluding subjects or postponing entry to the study were: decompensated heterophoria (according to conventional clinical criteria); or blurred vision due to uncorrected refractive or accommodative anomaly. Nine subjects were excluded for these reasons.

Visual acuity and orthoptic tests were carried out with any refractive correction that was usually worn at the distance used. The mean binocular near visual acuity was 0.05 LogMAR units. Most optometric findings were within clinical norms, with the exception of the mean amplitude of accommodation which was 12.3 D. This is less than the value of 15.2 D which would be predicted for a sample of this age by Hofstetter's formula and concurs with other research that has found a reduced amplitude of accommodation to be a correlate of dyslexia.

The prevalence of colour vision abnormalities was similar to that expected in a sample of this size and gender ratio on both the Ishihara and City University tests. Although there has been a suggestion that the prevalence of such abnormalities is greater in children with learning difficulties, this runs counter to the body of evidence. The Ishihara test (available for 33 subjects) revealed 9% with more than four errors, the accepted criterion for failure. We would have expected three such failures amongst the boys (two were obtained) and 1.4 among the girls (one obtained). None of those with more than four errors were amongst the 36 subjects who completed diaries and contributed to the analysis of symptoms below. The City University test detected errors in 17% of the 40 subjects tested, slightly less than the 24% expected. One fifth of the subjects had a family history of colour vision defects. The Farnsworth–Munsell 100-Hue test was administered to 32 subjects. The mean error score was 122 (SD 77). In most cases the scores were 'anarchic' and it was difficult to discern a profile.

Most of the children, 72%, had previously seen an optometrist, but only 18% wore glasses; 27% gave a history of orthoptic anomalies ('turning eye', eye exercises or patching); 42% gave a history of headaches severe enough for the child to be kept off school; and 57% gave a family history of migraine.

When questioned about reading, 76% of the parents reported that their child tired easily, and 90% said the child omitted lines or words and reread them.

Maintenance of the masked control

At the second optometric examination 49 of 52 subjects said that one or both pairs of glasses had helped: 31 subjects preferred the first pair of glasses, 17 the second and 4 expressed no preference. Overall, 22 subjects preferred the experimental spectacles and 26 preferred the control; 23 subjects 'did not know' which pair matched the original preferred colorimeter setting; 10 thought the experimental glasses were the ones that matched and 11 were of the opposite opinion. Subjects were not able reliably to replicate the appearance of text through either pair of glasses when viewing text in the colorimeter and adjusting hue (given the appropriate setting for saturation).
Chromaticities of the lenses

The chromaticities of the experimental lenses are shown by the points in Figure 1. A line joins the chromaticity of the experimental lens (open symbol) with that of the control. The average separation of experimental and control lenses is shown by the horizontal dotted line in the inset of Figure 1. The separation was small (CIE 1976 \( h_{ab} \)) between the experimental and control lenses was small (37.3°), but because the direction of the difference varied randomly, the hue averaged across subjects was closely similar for experimental and control lenses (\( h_{ab} = 243.2° \) and 243.7°, respectively). The average saturation (CIE 1976 \( s_{ab} \)) of the experimental and control lenses did not differ (\( s_{ab} = 1.26 \) and 1.16 respectively; \( P = 0.22 \)).

Use of the lenses

The experimental glasses were worn on 72.0% of days for which they were available, for an average of 2.1 h per day. The control glasses were worn on 71.6% of days for an average of 2.1 h. The differences did not approach significance.

Relief of symptoms

Days on which the spectacles were worn were divided into those on which headache and eye-strain occurred, and days that were free of these symptoms. The mean number of days on which the experimental spectacles were worn was 18.4 of which 13.9 (71%) were symptom-free. The control spectacles were worn for 17.9 days of which 11.9 (66%) were free of symptoms. Figure 2 shows the proportion of days with symptoms when experimental glasses were worn and when control glasses were worn. There are 19 points above the diagonal and 11 below, with six at the origin (i.e. showing no symptoms with either pair of glasses). For each subject the number of days with symptoms and the number without were calculated separately for days on which the experimental spectacles were worn and those on which the control spectacles were worn. A 2 \( \times \) 2 table expressed for each subject the contingency between the presence or absence of symptoms and the type of spectacles worn. Fisher’s exact probability test showed this association to be significant in seven children. All seven showed a significant reduction in symptoms with the experimental spectacles, indicated by solid points in Figure 2.

From each contingency table the log–odds ratio was calculated using the formula:

\[
\ln \left( \frac{\text{number of days with symptoms when 'control' glasses worn}}{\text{number of days without symptoms when 'control' glasses worn}} + \frac{1}{2} \right) - \ln \left( \frac{\text{number of days with symptoms when 'experimental' glasses worn}}{\text{number of days without symptoms when 'experimental' glasses worn}} + \frac{1}{2} \right)
\]

The value of the log–odds for each subject provided a measure of the relative benefit of the experimental glasses. When the individual probabilities of the log–odds were analysed using the methods described by Fleiss (pp. 165–168), the combined log–odds indicated significantly fewer symptoms with the experimental lenses (\( \chi^2(1) = 13.3, P = 0.002 \)). The distribution of log–odds across subjects was heterogeneous (\( \chi^2(22) = 48.0, P = 0.001 \)) indicating that the difference in the effect of experimental and control lenses was greater for some subjects than for others. The average value of the log–odds was significantly greater than zero (\( t(29) = 2.33, P = 0.026 \)), indicating that even when the individual probabilities of the log–odds were ignored, the experimental glasses were associated with greater clinical benefit in the group as a whole.

There was no significant correlation between the value of
the log–odds and the transmission, hue angle or saturation of the experimental or control lenses, or the difference between the two in transmission or chromaticity. There was a non-significant tendency for the log-odds to be greater when the placebo was presented first ($t(29) = 1.79, P = 0.083$).

**Neale analysis of reading**

The Neale Analysis of Reading gave reading ages for speed, accuracy and comprehension and these were 9.37, 9.57 and 10.28 years respectively when the experimental lenses were worn, and 9.22, 9.42 and 10.12 years respectively when the control lenses were worn. The differences were not statistically significant, even when raw scores for equivalent passages were compared. The Diagnostic Tests were administered during the baseline period before the lenses were worn. These tests do not have age-norms, but their raw scores are comparable with those of the other test forms which were used when the glasses were worn. It is therefore of interest that accuracy scores obtained using the Diagnostic Tests (without glasses) were much poorer than those obtained with the other test forms (when the glasses were worn; $P < 0.006$, t-tests).

For reasons described in the introduction, the selection of subjects was deliberately heterogeneous; the subjects could be loosely categorized as suffering from headaches, asthenopia, and reading difficulty. This raises the question of which of these categories was most strongly associated with the benefit from coloured filters. The sample size within each of these categories is too small for any conclusive analyses. The correlation between log–odds and the history of headaches, symptoms of asthenopia when reading, and reading age and reading retardation (regardless of IQ) were all small and non-significant. It seems that children who benefit from colour cannot be simply described as a subset of those with reading difficulty, or of those with headache, or asthenopia. It remains possible that, as Irlen has claimed, ‘Irlen Syndrome’ is a unitary condition that is sometimes associated with each of these other conditions.

**Discussion**

The clinical improvements recorded in the headache and eye-strain diaries cannot be readily interpreted in terms of placebo effects. The experimental and control lenses were similar in colour, subjects were not told that one pair was designed to be less effective than the other, and there was always a period of several weeks between the colorimetry and the issue of the first pair of glasses, and between the first and second pair. The double-mask was confirmed by the demonstration that the subjects were unable to distinguish the experimental and control lenses and by their inability to match the colour of either pair of lenses using the colorimeter. The attrition rate was high, but, given the random allocation of subjects to treatment condition, and the maintenance of the double-mask it is difficult to see how the results can be interpreted as an artefact of attrition.

The finding that both pairs were, to varying degrees, reported to be of benefit may suggest that the control and experimental tints were so similar that the subjects were unaware of the slight difference in symptoms with each pair. A more powerful clinical effect might have been obtained by increasing the chromaticity difference between the two pairs, but this would have weakened the double-masked design. Nevertheless there was, on average, a greater reduction in symptoms when the experimental glasses were worn. These had a chromaticity similar to that which had previously reduced perceptual distortions and visual discomfort when text was observed in the colorimeter. The difference in the chromaticity of experimental and control lenses was small, indicating that the clinically effective tint is idiosyncratic and needs to be determined with precision. The specificity of the therapeutic tint explains the failure of previous attempts to demonstrate the effectiveness of coloured glasses despite anecdotal reports of success. Evidently the Intuitive Colorimeter provides one way of determining a therapeutic tint and has the advantage that the effects of the tint can be assessed rapidly whilst the eyes are colour adapted.

It seems unlikely that vision therapy would have afforded our subjects the same symptomatic relief as coloured glasses (as has been suggested elsewhere), for the following reasons: subjects used any refractive correction that was necessary to alleviate refractive blur at distance or near; the sample did not include cases of strabismus or decompensated heterophoria; the difference between the experimental and control lenses was not, on average, of a type that would have alleviated heterophoria by modifying accommodation; and no ocular or ocular motor mechanism has been suggested that could account for the degree of specificity of the beneficial colour.

The physiological basis of the efficacy remains obscure. Tints have reduced epileptic photosensitivity and different photosensitive patients show different effects of colour. In some patients the trigger involves focal cortical hyperexcitability, as evidenced by the selective response to gratings with different orientations. Minimal cortical hyperexcitability has been proposed as a neurological basis for visual discomfort and perceptual distortions in people without epilepsy, particularly those with migraine. Many of the children had migraine in the family, and a family history of migraine has previously been shown to be far more common amongst those who find tints beneficial. Individuals with migraine have colour preferences different from age- and sex-matched controls. The above observations are consistent with the notion that the tint changed the pattern of excitation in striate and pretectate cortical areas so as to reduce the excitation in locally hyperexcitable regions.

**Acknowledgements**

The goodwill of the children, their parents, their teachers and their schools is gratefully acknowledged. We are particularly grateful to the following teachers and head-teachers: R. Bather, I. Gliddon, P. Hawkey and C. Waugh. We are also indebted to K. C. Holland and E. Pimm for advice. We thank Clement Clarke International, CIBA Vision and Vistechn Consultants Inc. for donating equipment. The work was funded by the Medical Research Council, the Institute of Optometry, Cerium Optical Products, Dollond and Aitchison and the Paul Hamlyn Trust.

**Disclosure of interest**

None of the authors has a direct financial interest in precision tinting. The Medical Research Council owns the rights to the Intuitive Colorimeter® and associated tinting system which is marketed under licence by Cerium Visual Technologies, Tenterden, Kent, UK.

References