Irlen Colored Overlays Do not Alleviate Reading Difficulties
Stuart J. Ritchie, Sergio Della Sala and Robert D. McIntosh
Pediatrics; originally published online September 19, 2011;
DOI: 10.1542/peds.2011-0314

The online version of this article, along with updated information and services, is located on the World Wide Web at:
http://pediatrics.aappublications.org/content/early/2011/09/15/peds.2011-0314
Irlen Colored Overlays Do not Alleviate Reading Difficulties

WHAT’S KNOWN ON THIS SUBJECT: Colored overlays are recommended by Irlen Institute specialists to alleviate Irlen syndrome (also known as visual stress), a disorder posited to be a cause of reading difficulty in many children. Both the intervention and the disorder, however, are controversial.

WHAT THIS STUDY ADDS: This study, in which methodologic problems that have been noted in previous studies were avoided, revealed that Irlen colored overlays have no immediate effect on reading ability in children with reading difficulties.

abstract

OBJECTIVES: To test the efficacy of Irlen colored overlays for alleviating reading difficulties ostensibly caused by Irlen syndrome, a proposed perceptual disorder with controversial diagnostic status.

PARTICIPANTS AND METHODS: Sixty-one schoolchildren (aged 7–12 years) with reading difficulties were assessed by an Irlen diagnostician. We used a within-subject study design to examine differences in reading rate across 3 conditions: using an overlay of a prescribed color; using an overlay of a nonprescribed color; and using no overlay. In a subset of 44 children, all of whom had a diagnosis of Irlen syndrome, we also used a between-group design to test the effects of Irlen colored overlays on a global reading measure.

RESULTS: The Irlen diagnostician diagnosed Irlen syndrome in 77% of our poor readers. We found no evidence for any immediate benefit of Irlen colored overlays as measured by the reading-rate test or the global reading measure.

CONCLUSIONS: Our data suggest that Irlen colored overlays do not have any demonstrable immediate effect on reading in children with reading difficulties. Pediatrics 2011;128:e932–e938

AUTHORS: Stuart J. Ritchie, MSc, Sergio Della Sala, MD, PhD, and Robert D. McIntosh, PhD

aHuman Cognitive Neuroscience, University of Edinburgh, Edinburgh, United Kingdom

KEY WORDS
education, reading, dyslexia, learning disorders, color, ocular physiological processes

ABBREVIATIONS
IS—Irlen syndrome
MMSE—Mini-Mental State Examination
WRRT—Wilkins Rate of Reading Test
GORT—Gray Oral Reading Test
ORQ—oral reading quotient

All 3 authors designed the study. Mr Ritchie and Dr McIntosh collected and analyzed the data. All 3 authors contributed to the writing of the article.

www.pediatrics.org/cgi/doi/10.1542/peds.2011-0314
doi:10.1542/peds.2011-0314
Accepted for publication Jun 8, 2011
Address correspondence to Stuart J. Ritchie, MSc, Psychology Department, University of Edinburgh, 7 George Square, Edinburgh EH8 9JZ, United Kingdom. E-mail: stuartjritchie1@gmail.com

PEDIATRICS (ISSN Numbers: Print, 0031-4005; Online, 1098-4275).
Copyright © 2011 by the American Academy of Pediatrics

FINANCIAL DISCLOSURE: The authors have indicated they have no financial relationships relevant to this article to disclose.
The use of colored filters to alleviate reading difficulties is recommended by Irlen Institute specialists internationally. These filters have received regular mass media exposure, and their use is increasingly accepted in schools. Proponents of this method have suggested that ∼12% of the population, and up to 46% of individuals with dyslexia and learning disabilities, suffer from Irlen syndrome (IS) (also known as visual stress, Meares-Irlen syndrome, and scotopic sensitivity syndrome), a proposed perceptual disorder that causes an individual to experience visual distortions and illusions when viewing text, and interferes with reading ability. Individually prescribed colored filters, either tinted spectacle lenses or colored sheets of plastic (overlays), are posited to alleviate these visual distortions, removing an “obstacle” to reading instruction. This proposed reduction of visual distortions would hypothetically improve text discrimination and reading rate in the short-term, and facilitate the development of reading skills in the longer term.

Both the existence of IS and the efficacy of the colored-filter treatment are controversial in scientific and medical circles. Several recent reviews, including a review from the American Academy of Pediatrics, have noted that the studies performed to investigate the condition are often of poor quality.Methodologic issues include preselection of participants who have already reported a benefit from the filters, inadequate blinding/masking, small sample sizes, and inappropriate statistical analysis. Studies for which results were claimed to demonstrate improved reading with filters contrast with studies in which no such effects were observed. Given the widespread use of the filters, the lack of reliable evidence for their beneficial effects is cause for concern. Thus, the investigation of colored filters under more rigorous conditions is important.

In the present study we tested the short-term efficacy of Irlen colored overlays in a sample of schoolchildren with reading difficulties who were assessed by an Irlen diagnostician. To our knowledge this study was the first investigation of Irlen colored overlays in which the participants had no explicit knowledge of their prescribed color before being tested.

**METHOD AND PARTICIPANTS**

**Participants**

Participants were 75 children who attended Newark Primary School, Port Glasgow, Scotland. All participants were selected by their classroom teachers for assessment by an Irlen diagnostician, an intervention that was ongoing at the school, because these children were judged to be below-average readers. Parent/guardian consent was received for all participants, and the study was approved by the ethics panel of the University of Edinburgh Psychology Department.

**Outcome Measures**

Irlen colored overlays are available in 10 colors with both matte and shiny sides. In this study the overlays were used in conjunction with the reading tests as detailed below. The overlays were provided by the Irlen diagnostician, but the Irlen name and other details printed along the top and bottom edges were removed. A colorless overlay (a clear cellulose acetate sheet, both sides of which were shiny) was also used.

The Mini-Mental State Examination (MMSE), often used to assess dementia in older adults, has been shown to have validity in children as an assessment of general cognitive ability, and was chosen for use in the present study because of its brief administration time.

The Wilkins Rate of Reading Test (WRRT) consists of lists of 15 familiar words arranged in 10 lines, each with a different random word order, in closely spaced type and a small font size. Each 150-word test is read for 1 minute, and deviations are recorded. As advised (Arnold Wilkins, PhD, written, January 25, 2010), we used the form of the WRRT with a larger font size to suit the age and reading ability of the participants. The test has 4 variations; 2 extra variations were created to allow for the procedure described below. A practice sheet, to be read for 30 seconds before first administration of the task, was also created.

In addition, we assessed participants by using the Gray Oral Reading Test, fourth edition (GORT), a measure of global reading ability that gives 3 results: reading fluency; reading comprehension; and an oral reading quotient (ORQ).

**Procedure and Randomization**

Diagnosis of IS involves a 1-to-1 meeting with a diagnostician who has expertise with this condition. Participants perform tasks such as counting symbols within pictures and looking at text through various overlaid colors, while answering questions about visual discomfort and distortions (e.g., blurring, movement, flickering, glowing). A process of elimination is then used to find an optimal color from the 10 available. The session lasts ~30 minutes. Ordinarily, if IS is diagnosed, a participant is informed of the diagnosis and their prescribed filter color at the end of the session. To maintain treatment masking in this study, however, we arranged with the Irlen diagnostician to simply inform all of the children that additional testing would take place in future. Aside from this change, the diagnostic session pro-
ceeded exactly as it would have if our study had not been taking place. We accepted the diagnoses from the sessions because we aimed to test the efficacy of the treatment as it is typically administered by Irlen practitioners. The cost of the diagnostic session was covered by the local council, no parents or guardians paid fees toward the cost of the Irlen assessment.

For children in whom IS was diagnosed (the Irlen group), the Irlen diagnostician informed Dr McIntosh of each child’s diagnosis and prescribed color. Children who did not have an IS diagnosis (the non-Irlen group) were nevertheless assigned a mock prescribed color by Dr McIntosh, to match approximately the frequency of each color prescribed to children in the Irlen group. Each prescribed color was paired with a nonprescribed color, selected (if possible) from the complementary part of the spectrum. Colorless overlays were used as an additional comparison condition for each child.

The WRRT was administered in ABCBAC form: each participant read in each condition (prescribed color, nonprescribed color, and colorless overlay) twice. The trial order list was prepared such that within each sequential series of 12 participants, the 6 possible trial orders for the WRRT were counterbalanced against the 2 possible overlay assignments for the GORT (prescribed color, colorless overlay). For each participant, Mr Ritchie was notified of the trial orders for the WRRT and the GORT (but not of the child’s diagnostic status) by 2 separate e-mails from Dr McIntosh before participant testing began.

On a different day from their Irlen diagnostic session, participants sat in a quiet room in their school that had fluorescent lights switched on, as well as natural light from a window. Mr Ritchie began by administering the MMSE, then opened the first e-mail to ascertain the trial order for the WRRT. The WRRT was then completed, and the second e-mail was opened before the GORT was administered. This session took ~30 to 40 minutes.

Throughout the procedure no reference was made to the claimed effects of the overlays, and any questions asked by the children on this subject were redirected. The overlays were simply placed over the text before reading began; no explanation was given or attention drawn to them. The overlays were thus imbued with as little significance as possible.

Finally, all children were assessed on a third day by a qualified orthoptist, from whom the child’s diagnostic status was also masked. The following measures were taken: unocular near and distance visual acuity, ocular motility, convergence, accommodation, fusional reserves, and saccades. A cover test for strabismus was also performed. This session took no more than 30 minutes. If orthoptic problems were discovered, parents or guardians were sent a letter that referred their child to their local eye clinic for treatment.

RESULTS

Participant Characteristics

Thirteen children were deemed unsuitable for diagnosis by the Irlen diagnostician because of their absence from school or test failure. Of the 58 children who attended both sessions, 39 (67%) were diagnosed IS (mean age: 10.6 years; SD: 1.19 years) and 19 (33%) were excluded from further analysis. Thirty-four of the 39 children from the Irlen diagnostic group completed the orthoptic assessment. The Irlen diagnostician diagnosed IS in 47 of the 61 children (77%). Participant characteristics are summarized in Table 1. All children completed the MMSE, and 59 completed the orthoptic assessment. As expected, scores on the MMSE were correlated positively with chronological age (\(r_{59} = .36; P = .002\)), with baseline WRRT scores (those that were obtained when the colorless overlay was used) for all participants (\(r_{58} = .40; P = .002\), and with GORT ORQ averaged across the 2 experimental conditions (\(r_{58} = .41; P = .001\)). There was no significant difference between the Irlen and non-Irlen groups for the mean MMSE score (\(t_{59} = -.73; P = .47\)) or the mean baseline WRRT score (\(t_{59} = .98; P = .33\)), which indicated that the 2 groups were similar in general cognitive ability as well as reading proficiency.

Results of comparison of Irlen and non-Irlen by Mann-Whitney U tests showed no significant differences in the percentage of individuals with orthoptic problems, ie, the percentage of the group for which any orthoptic tests were failed (\(U = 251.50; z = -1.00; P = .32\), or in the mean percentage of orthoptic tests failed (\(U = 262.50; z = -.71; P = .46\)). The mean percentage of orthoptic tests failed did, however, correlate negatively with baseline WRRT score (\(r_{59} = -.27; P = .04\)). Thus, more orthoptic problems were associated with slower reading rates, but children in the Irlen group did not suffer significantly more orthoptic problems than children in the non-Irlen group.

Table 2 shows the frequency of each overlay color prescribed to the Irlen

<table>
<thead>
<tr>
<th>TABLE 1 Characteristics of Irlen and Non-Irlen Groups</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group</td>
</tr>
<tr>
<td>Age, mean (SD), y</td>
</tr>
<tr>
<td>Gender-frequency, male/female, n</td>
</tr>
<tr>
<td>MMSE score, mean (SD)</td>
</tr>
<tr>
<td>Children with any failed orthoptic tests, n (%)</td>
</tr>
<tr>
<td>Orthoptic tests failed, n (mean %)</td>
</tr>
</tbody>
</table>

The Irlen diagnostician diagnosed IS in 47 of the 61 children (77%). Participant characteristics are summarized in Table 1. All children completed the MMSE, and 59 completed the orthoptic assessment. As expected, scores on the MMSE were correlated positively with chronological age (\(r_{59} = .36; P = .002\)), with baseline WRRT scores (those that were obtained when the colorless overlay was used) for all participants (\(r_{58} = .40; P = .002\), and with GORT ORQ averaged across the 2 experimental conditions (\(r_{58} = .41; P = .001\)). There was no significant difference between the Irlen and non-Irlen groups for the mean MMSE score (\(t_{59} = -.73; P = .47\)) or the mean baseline WRRT score (\(t_{59} = .98; P = .33\)), which indicated that the 2 groups were similar in general cognitive ability as well as reading proficiency.

Results of comparison of Irlen and non-Irlen by Mann-Whitney U tests showed no significant differences in the percentage of individuals with orthoptic problems, ie, the percentage of the group for which any orthoptic tests were failed (\(U = 251.50; z = -1.00; P = .32\), or in the mean percentage of orthoptic tests failed (\(U = 262.50; z = -.71; P = .46\)). The mean percentage of orthoptic tests failed did, however, correlate negatively with baseline WRRT score (\(r_{59} = -.27; P = .04\)). Thus, more orthoptic problems were associated with slower reading rates, but children in the Irlen group did not suffer significantly more orthoptic problems than children in the non-Irlen group.

Table 2 shows the frequency of each overlay color prescribed to the Irlen
group, along with the colors allocated in the matching process the non-Irlen group.

**Effect of Overlays on Rate of Reading (Within-Subject Design)**

Sixty participants successfully completed the WRRT. During testing, 3 of the children in the Irlen group stated that they knew their filter color, and we determined that 2 of these children had in fact been using their filter for some days before the experiment. This situation compromised treatment masking for both participant and experimenter, so data from these children were separated from the main analysis.

Figure 1 illustrates the mean WRRT scores for the Irlen group (n = 43) and the non-Irlen group (n = 14) for each overlay condition. The means, as well as individual scores, for the 3 participants with a nonmasked treatment (n = 3) are also included, for illustrative comparison.

A mixed-design analysis of variance was performed on WRRT scores, with the between-subject factor of group (Irlen, non-Irlen) and the within-subject factor of overlay condition (colorless overlay, nonprescribed overlay, prescribed overlay). The main effect of the group was not significant (F1,55 = .24; P = .63), nor was there a significant effect of overlay condition (F2,110 = .16; P = .85) or an interaction (F2,110 = .91, P = .41). Thus, colored overlays, whether of the prescribed color or not, did not facilitate reading rate compared with a colorless overlay.

Figure 2 illustrates for all children the percentage change in reading rate, relative to the colorless overlay condition, for both the prescribed and the nonprescribed overlay. Because a common criterion for the identification of a positive effect of a colored filter is an increase in reading rate of 5% over baseline, dotted lines have been added to denote a 5% increase and a 5% decrease in reading rate with the prescribed overlay. Almost as many of the participants showed a 5% or greater decrease in reading rate (18 children) as showed a 5% increase (19 children). The most dramatic benefits were shown by the 2

### TABLE 2 Prescribed and Allocated Colors According to Group

<table>
<thead>
<tr>
<th>Group</th>
<th>Frequency of Color Prescribed/Allocated</th>
</tr>
</thead>
<tbody>
<tr>
<td>Irlen (n = 47)</td>
<td>Aqua: 15</td>
</tr>
<tr>
<td>Non-Irlen (n = 14)</td>
<td>Aqua: 7</td>
</tr>
</tbody>
</table>

All colored overlays were used matte-side up.
children whose data are displayed in the top right of Fig 2, for both of whom the treatment was not masked (they were already aware of their prescribed overlay color, which meant that we could not disentangle any true treatment effect from a possible placebo effect in these cases).

**Effect of Overlays on Oral Reading Test Outcome (Between-Group Design)**

Sixty of the 61 participants completed the GORT. The poor general reading ability of the children in this sample was confirmed by an overall mean GORT ORQ of 80.00 (SD: 13.64), which was 1.3 SDs below the test’s standardized mean of 100.

Because of the low numbers of non-Irlen children (n = 7 per treatment group), this analysis focused exclusively on the Irlen group. The Irlen group (n = 44) was split into a prescribed overlay group and a colorless overlay group; characteristics of the groups are shown in Table 3.

Results of independent samples t-tests showed there were no significant differences in age (t(42) = −.38; P = .71), mean MMSE score (t(42) = .07; P = .94) or baseline WRRT score (t(54.70) = .33; P = .75). For the GORT measures, there were no significant between-group differences in fluency (t(42) = .48; P = .65), comprehension (t(42) = −.43; P = .67), or ORQ (t(42) = .13; P = .90). The prescribed filters, then, did not show evidence of significant facilitation of reading on any of the GORT measures compared with the colorless overlays.

**DISCUSSION**

In this study we found that Irlen colored overlays have no clinically or statistically significant immediate effect on reading ability in poor readers. In a sample of children who, to the best of our knowledge and ability, had their diagnostic status and the color of their prescribed overlay masked, neither prescribed nor nonprescribed colored overlays had any effect on reading ability as measured with the WRRT or the GORT. The lack of attention drawn to the filters, in addition to the double-masked design, is a possible explanation for the absence even of any placebo effects, despite the inherently salient nature of the colored filter intervention.

Drawing any conclusions from a subsample as small as those for whom the treatment was not successfully masked (n = 3) requires extreme caution, but the difference in the pattern of WRRT scores between this subsample and the 2 main experimental groups is nonetheless striking. Two of the children with a nonmasked treatment experienced a dramatic overlay effect (see top right corner of Fig 2) that would be consistent with the claims made by some proponents of colored filters. Of course, the possibility cannot be excluded that masking was compromised for these children precisely because their therapeutic response was so dramatic that their prescribed overlay color could not be effectively hidden from them. The observations are also consistent with the more prosaic possibility that colored filters may exert a strong placebo effect among children who sincerely believe that the use of these filters will assist their reading. This possibility further underscores the need for careful experimental masking in any future studies of colored filters.

Our results did not test the nature of IS, or its existence as a meaningful diagnostic entity. Not only were there no differences in treatment effect between Irlen and non-Irlen groups, no overall differences were found between the 2 groups in MMSE score, baseline reading ability, or orthoptic problems. Proponents of colored filters would predict this equivalence, because they emphasize that individuals with IS/visual stress are not necessarily poorer readers, and that the condition is not caused by low-level visual disorders. In support of these theories, it has been noted that no reliable orthoptic correlates of the condition have been found. However, the few studies in which adequate control for all potential orthoptic problems has been included have yielded less positive results of assessments of the efficacy of the filters. In the present study, the Irlen group did have a numerically higher rate of orthoptic problems than the non-Irlen group (54.3% vs 38.5%); the failure of this difference to reach statistical significance may have been attributable to the unequal group sizes (46 in the Irlen group and 13 in the non-Irlen group) and, especially, to the low number of non-Irlen children, which resulted in low statistical power to detect between-group differences in orthoptic problems.

**TABLE 3 Age and Test Scores for Randomly Assigned Groups**

<table>
<thead>
<tr>
<th>Group</th>
<th>Age, Mean (SD), y</th>
<th>MMSE Score, Mean (SD)</th>
<th>WRRT Baseline Score, Mean (SD)</th>
<th>GORT Fluency, Mean (SD)</th>
<th>GORT Comprehension, Mean (SD)</th>
<th>GORT ORQ, Mean (SD)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prescribed overlay</td>
<td>9.73 (1.20)</td>
<td>23.86 (4.09)</td>
<td>80.25 (33.80)</td>
<td>4.68 (3.28)</td>
<td>7.82 (2.10)</td>
<td>77.50 (13.97)</td>
</tr>
<tr>
<td>(n = 22)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Colorless overlay</td>
<td>9.59 (1.18)</td>
<td>23.95 (4.10)</td>
<td>83.01 (20.59)</td>
<td>5.09 (2.33)</td>
<td>7.50 (2.79)</td>
<td>78.05 (13.65)</td>
</tr>
<tr>
<td>(n = 22)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Our sample size was more than sufficient to test the claim that the Irlen method benefits 46% of individuals with “reading problems, dyslexia and learning difficulties.” Although we cannot state the proportion of our sample in which dyslexia or learning disabilities would be formally diagnosed, it would likely have included a substantial number of children with such a diagnosis. Indeed, the very low mean GORT ORQ score in our sample population gives reason to characterize our sample as reading disabled. Even more tellingly, IS was diagnosed in 77% of the children in our sample, and even in these children no differences in reading with and without overlays were found.

On the basis of our results the possibility cannot be discounted that, in a small subset of individuals, the use of colored filters may lead to clinically significant improvements in reading, and this subset is too small for trials of the precision of the tint is of great importance, and given the nature described here to detect. Why, then, have significant sample-wide effects been found in previous trials? As mentioned above, by selecting for inclusion only individuals who had previously been using colored lenses, researchers may have biased their results toward the positive. In addition, the statistical techniques used may have increased the likelihood of finding a positive result.

In our study we focused on diagnoses made by an Irlen practitioner. Therefore, caution is required in generalizing our results to other colored filter systems, such as Wilkins’ Intuitive system. Given the widespread usage of the Irlen system and the Intuitive system, it is surprising that the 2 systems have never been directly compared experimentally. Nevertheless, the available overlay colors are broadly similar in both systems, and a similar process of elimination is used in both to select the optimal color. One would expect a large overlap in diagnosis, especially given the very similar conceptions of the symptoms of IS and visual stress.

On the other hand, for colored lenses, Wilkins’ system uses the “Intuitive Colorimeter,” which allows a very precise tint to be selected. It may be that the precision of the tint is of great importance, and given the nature of our study, our results do not shed light on the use of such methods.

CONCLUSIONS

Our results suggest that Irlen colored overlays have no immediate effect on reading in children with reading difficulties, even among those in whom IS has been diagnosed. On the basis of these results, and the lack of convincing evidence from previous studies, we recommend that parents, schools, health care professionals, and government bodies carefully consider the totality of the evidence before expending time, resources, and hope on this controversial treatment.

ACKNOWLEDGMENTS

This study was supported by the Moray Endowment Fund of the University of Edinburgh.

We are very grateful to the teachers, support staff, and pupils at Newark Primary School. We thank Liz McKelvie (Irlen Scotland) for her assistance, and Nadia Northway and Jillian Campbell (Glasgow Caledonian University) for arranging and carrying out the orthoptic examinations.

REFERENCES

1. Irlen H. Reading by the Colors: Overcoming Dyslexia and Other Reading Disabilities Through the Irlen Method. New York, NY: Avery Publishing Group; 1991
3. Wilkins A J Reading Through Colour. Chichester, United Kingdom: John Wiley and Sons; 2003
15. Menacker SJ, Breton ME, Breton ML, Radcliffe J, Gole GA. Do tinted lenses improve the reading performance of dyslexic chil-


<table>
<thead>
<tr>
<th><strong>Irlen Colored Overlays Do not Alleviate Reading Difficulties</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>Stuart J. Ritchie, Sergio Della Sala and Robert D. McIntosh</td>
</tr>
<tr>
<td><em>Pediatrics</em>; originally published online September 19, 2011;</td>
</tr>
<tr>
<td>DOI: 10.1542/peds.2011-0314</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>Updated Information &amp; Services</strong></th>
<th>including high resolution figures, can be found at:</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th><strong>Permissions &amp; Licensing</strong></th>
<th>Information about reproducing this article in parts (figures, tables) or in its entirety can be found online at:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><a href="http://pediatrics.aappublications.org/site/misc/Permissions.xhtml">http://pediatrics.aappublications.org/site/misc/Permissions.xhtml</a></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>Reprints</strong></th>
<th>Information about ordering reprints can be found online:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td><a href="http://pediatrics.aappublications.org/site/misc/reprints.xhtml">http://pediatrics.aappublications.org/site/misc/reprints.xhtml</a></td>
</tr>
</tbody>
</table>