Irlen Colored Filters in the Classroom: A 1-Year Follow-Up

Stuart J. Ritchie1, Sergio Della Sala1, and Robert D. McIntosh1

ABSTRACT—Colored filters are used to treat Irlen syndrome (IS), a controversial disorder posited to be the cause of a substantial proportion of reading difficulties. Previously, we found that Irlen colored filters do not produce any short-term alleviation of reading difficulties in schoolchildren aged 7–12. Here, we tested whether colored filters show benefits with longer-term use, in a subset of the original sample. We measured reading rate with and without filters in 18 children diagnosed with IS, who had been using the filters for 1 year, and compared the progression of their reading ability across the year against 10 poor-reader control children. The Irlen-treatment group did not read any faster when using their colored filter, and showed no disproportionate gain in reading progress across the year compared to controls. We conclude that Irlen filters do not benefit reading, even after 1 year of use.

Irlen syndrome (IS), also known as Meares–Irlen syndrome, scotopic sensitivity syndrome, or visual stress, is a controversial diagnostic entity that purportedly causes visual distortions and illusions when an affected person views text or other high-contrast patterns (Irlen, 1991; Wilkins, 2003). The Irlen Institute posits that these symptoms are often the cause of reading difficulties in up to 46% of individuals with “reading problems, dyslexia, and learning difficulties” (Perceptual Development Corporation, 1998), but also that they can be alleviated by the use of individually prescribed colored filters (Irlen, 2010; Wilkins, 1994). The colored filter treatment is not marketed as a cure for reading difficulties, but is believed to remove a barrier to reading development (Irlen, 2010). These filters, either in the form of tinted lenses or colored plastic overlays, are used worldwide (Hyatt, Stephenson, & Carter, 2009, p. 321), often in school classrooms (Wilkins, 1994), and regularly receive mass media coverage.

Irlen syndrome (IS), also known as Meares–Irlen syndrome, scotopic sensitivity syndrome, or visual stress, is a controversial diagnostic entity that purportedly causes visual distortions and illusions when an affected person views text or other high-contrast patterns (Irlen, 1991; Wilkins, 2003). The Irlen Institute posits that these symptoms are often the cause of reading difficulties in up to 46% of individuals with “reading problems, dyslexia, and learning difficulties” (Perceptual Development Corporation, 1998), but also that they can be alleviated by the use of individually prescribed colored filters (Irlen, 2010; Wilkins, 1994). The colored filter treatment is not marketed as a cure for reading difficulties, but is believed to remove a barrier to reading development (Irlen, 2010). These filters, either in the form of tinted lenses or colored plastic overlays, are used worldwide (Hyatt, Stephenson, & Carter, 2009, p. 321), often in school classrooms (Wilkins, 1994), and regularly receive mass media coverage.

1Human Cognitive Neuroscience, University of Edinburgh

Address correspondence to Stuart J. Ritchie, Department of Psychology, University of Edinburgh, 7 George Square, Edinburgh EH8 9JZ, UK; e-mail: stuartjritchiel@gmail.com.

Considerable controversy surrounds the efficacy of this treatment; three recent reviews (American Academy of Pediatrics, 2009; Hyatt et al., 2009; Royal College of Ophthalmologists, 2009) and one systematic review (Albon, Adi, & Hyde, 2008) have concluded that it should not be recommended for individuals with reading difficulties until more rigorous research shows positive effects.

Our recent study (Ritchie, Della Sala, & McIntosh, 2011) drew similar conclusions. We administered reading tests with and without colored overlays to 61 primary school children aged 7–12 years, 77% of whom had been diagnosed with IS by an Irlen Institute diagnostician. Importantly, the children diagnosed with IS were not informed of the color of their prescribed overlay before testing. Under these masked conditions, the overlays failed to produce any significant increase in the reading rate, as measured by the Wilkins Rate of Reading Test (WRRT), or global reading ability, as measured on the Gray Oral Reading Test (GORT). We concluded that Irlen colored filters do not alleviate reading difficulties.

Our study, like most previous work in this area, focused on the immediate effects of colored filters (though see Noble, Orton, Irlen, & Robinson, 2004; Robinson & Foreman, 1999a, 1999b). We compared reading with and without filters at a single time point in poor-reader children who had been prescribed the filters very recently. It is clear that the colored filter theory predicts an immediate benefit at this stage, but a further key prediction concerns the longer-term benefits. As the filters are intended as a long-term aid, to be used on a continuing basis, they should facilitate reading even after extended periods of use. More crucially, because the filters purportedly remove a prior barrier to reading development, we should expect the children using them to begin to catch up with their peers (those not diagnosed with IS) in terms of reading development; they should therefore show disproportionate gains in reading fluency and comprehension within the first year of treatment.

This study aimed to assess reading outcomes 1 year after a school-level intervention by the Irlen Institute, and to provide practical data on what educators might expect after such an intervention. Only two previous experiments have examined Irlen filters over the longer term. Noble et al. (2004) reported large gains in reading ability over 3 months for two groups...
of 31 and 40 children using colored filters (see Statistical Power section for more details). More mixed results (improved accuracy and comprehension, but not rate of reading or reading strategy) were reported for a sample of 38 children using colored filters over 20 months (Robinson & Foreman, 1999a, 1999b). However, this study has been criticized for various methodological problems such as failure to include an eye examination as part of the study and inappropriate statistical analyses (Albon et al., 2008; Hyatt et al., 2009). In addition, the 95% confidence interval for the reading accuracy results reported by Robinson and Foreman (1999b) crosses zero, potentially indicating a lack of effect (Albon et al., 2008). There is a clear need for more evidence on this issue.

In the present study, we followed up as many children as possible from our previous sample (Ritchie et al., 2011) 1 year after their IS diagnosis and colored filter prescription. Little evidence exists on the continuation rate of Irlen filters (as opposed to intuitive filters, a separate system that has been evaluated; Wilkins, Lewis, Smith, & Rowland, 2001). We first ascertained the number of children who were still using their prescribed overlay after 1 year. Next, we evaluated the reading rate with and without colored filters, and global reading ability, comparing poor-reader children who had been diagnosed with IS and were still using their colored filter with poor readers who had not been diagnosed with IS and had never used colored filters.

If colored filters benefit reading in children with a diagnosis of IS, then these children should read more fluently when using their filter than when using a filter of a different color, or when using no filter. Finally, we compared, for the first time, the progression of reading abilities of children diagnosed with IS, those who had discontinued treatment did not differ significantly from those who had continued in terms of mean age (9.60 vs. 9.68 years; \( t[45] = .79, p = .48 \)), mean Mini Mental State Examination (MMSE) score (23.64 vs. 24.27; \( t[43] = .54, p = .59 \)), WRRT performance with a colorless overlay (81.28 vs. 77.83; \( t[38.97] = .43, p = .67 \)), GORT oral reading quotient (ORQ) (78.50 vs. 76.41; \( t[44] = .52, p = .61 \)), percentage of orthoptic tests failed at examination in the original study (13.71% vs. 17.43%; \( t[45] = .91, p = .37 \)), or overlay benefit at the original study (1.00 vs. .90 extra words with the prescribed overlay than the colorless overlay; \( t[45] = .03, p = .98 \)). That is, in terms of the baseline measures available, there was no obvious pattern to distinguish the IS children who continued treatment from those who did not.

During the study period, we were able to ask 13 of the children in the latter group about why they had discontinued treatment. Three reported that they had found that the overlay made reading more difficult; two had visited an optician and had their vision corrected, which they saw as a replacement for the overlays; one had lost the overlays and not sought replacements; one could not remember having received an overlay; and six did not give a specific reason. Owing to the variation in responses, it is difficult to know whether these children should be regarded as children for whom the filter did not work, children for whom the filter worked only initially, children who lacked motivation to persist with the treatment, or (most probably) some mixture of these and other descriptors. Given the uncertain status, and likely heterogeneity of this group, we decided not to include “treatment-discontinued” children in the present study’s behavioral analyses.

**METHOD**

**Participants**

*Original Sample*

The recruitment sample was the group of 61 children who had taken part in our previous study (Ritchie et al., 2011). All had been identified by their teachers as having below-average reading ability. Children with a diagnosis of autism were excluded from the original sample, and thus were not included in the follow-up study. At the time of the previous study, all of these children were assessed by an Irlen diagnostician, and 47 of them received a diagnosis of IS while 24 did not.

**Treatment Continuation**

For the children diagnosed with IS, we obtained information from the school and from parents/guardians about whether the child was still using a colored filter. Twenty-two children with IS were still using a form of colored filter treatment after 1 year. In 14 cases, this was the colored overlay prescribed originally, while 8 children had progressed to colored-lens spectacles. An analysis of the records from the original study showed that of the 47 children who had been diagnosed with IS, those who had discontinued treatment did not differ significantly from those who had continued in terms of mean age (9.60 vs. 9.68 years; \( t[45] = .79, p = .48 \)), mean Mini Mental State Examination (MMSE) score (23.64 vs. 24.27; \( t[43] = .54, p = .59 \)), WRRT performance with a colorless overlay (81.28 vs. 77.83; \( t[38.97] = .43, p = .67 \)), GORT oral reading quotient (ORQ) (78.50 vs. 76.41; \( t[44] = .52, p = .61 \)), percentage of orthoptic tests failed at examination in the original study (13.71% vs. 17.43%; \( t[45] = .91, p = .37 \)), or overlay benefit at the original study (1.00 vs. .90 extra words with the prescribed overlay than the colorless overlay; \( t[45] = .03, p = .98 \)). That is, in terms of the baseline measures available, there was no obvious pattern to distinguish the IS children who continued treatment from those who did not.

**Follow-Up Groups**

Of the 22 IS children who had continued treatment, 18 were available for and completed follow-up testing (5 using lenses, 13 using overlays). Of the 13 children not diagnosed with IS, 10 were available for and completed testing. The follow-up testing was therefore completed by two groups, an Irlen-treatment group (\( n = 18 \)) and a non-Irlen group (\( n = 10 \)). An analysis of the records from the original study showed that the Irlen-treatment group was older than the non-Irlen group (8.70 vs. 9.78 years; \( t[26] = 2.28, p = .03 \)), but the groups did not differ on MMSE score (24.20 vs. 23.61; \( t[26] = 59, p = .70 \)), or colorless WRRT performance (87.90 vs. 79.39; \( t[26] = .24, p = .24 \)). The Irlen-treatment group did have a lower GORT ORQ than the non-Irlen group (74.83 vs. 91.00; \( t[26] = -3.63, p = .001 \)).
Statistical Power
As far as we are aware, the most comparable long-term study of the effects of Irlen filters is that of Noble et al. (2004), who reported that in their two groups of children the filters were associated with an increase in reading grade equivalents of an average of 1 year 5 months during the 3-month treatment period. After this period, the authors suggested that “any further rate of development may [have been] at normal grade expectations rather than at accelerated levels” (p. 21). While insufficient information was included by Noble et al. (2004) to calculate a precise effect size, on the basis of their results we would expect the effect of the overlays on reading to be moderate to large. Power analysis using G*Power 3.0 (Faul, Erdfelder, Lang, & Buchner, 2007) indicated that with our 18 Irlen-diagnosed participants, we had 80% power to detect any cross-year effect above approximately $d = .61$.

Reading Measures
The WRRT (Wilkins, Jeanes, Pumfrey, & Laskier, 1996) consists of four 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 min, and deviations are recorded. As in our previous study, the larger font size form of the WRRT was used. Two extra word lists were created to allow for the procedure described below. A “practice” sheet, read for 30 s before first administration of the task, was also created.

The GORT (Wiederholt & Bryant, 2001) is a global reading measure in which participants read stories aloud while being assessed for accuracy and fluency, then answer comprehension questions. An overall age-standardized ORQ is then calculated. Form B of the GORT was used in our original experiment, so the parallel Form A was used for this follow-up. During GORT testing, Irlen-treatment children used their usual colored filter (overlay or lenses), while non-Irlen children used no filter.

Procedure
The children were tested individually in a quiet room at their school (Newark Primary School, Port Glasgow, or Port Glasgow High School).

For the WRRT, an ABCBBA design identical to that of our previous study (Ritchie et al., 2011) was used, with each participant reading twice in each of three conditions: prescribed filter, nonprescribed filter, and colorless filter. For the non-Irlen group we used the mock “prescribed” color from the original study (these had been chosen to match approximately the frequency of colors prescribed to children in the Irlen group). The “nonprescribed” filters were determined by the same fixed pairings of filter colors as used in the original study—chosen, where possible, from the complementary end of the spectrum. The Irlen-treatment children who had progressed to colored lenses used these as their prescribed filter; all other filters used were plastic Irlen overlays. The order of the three conditions within the ABCBBA design was cycled from participant to participant within each group.

For the GORT, each participant read the stories in their favored filter condition. That is, all Irlen-treatment children used their usual filter (overlay or lenses), while non-Irlen children used no filter.

After completing both reading tests, the children were debriefed—the Irlen-treatment group were asked what difference they felt their filter made to their reading, while the non-Irlen group were simply thanked for their participation. After the study was completed, class teachers were surveyed about their pupils’ colored filter use; for each child, they were asked whether, in the period after their initial diagnosis, the child used the filter “always,” “regularly,” “sometimes,” or “not at all.”

RESULTS

Usage
Information from teachers on overlay/filter usage was available for 17 of the 18 IS children. This indicated that, after first being diagnosed with IS and prescribed a filter, four of the children “always” used their filter in class (three of these were the children who used colored lenses). Eight used it “regularly” (these included the remaining two colored lens users), while the remaining five used their filter “sometimes.” This indicates a good general level of compliance with the treatment among the treatment-continued children.

WRRT Results
Table 1 shows the scores for the Irlen-treatment and non-Irlen groups in each filter condition (prescribed filter, nonprescribed overlay, colorless overlay) in this follow-up study. Given that the children in the Irlen-treatment group had been using the filters for a full year, suggesting that they felt that the treatment improved their reading, we predicted that they would read significantly faster in the prescribed filter condition than in the other conditions. A $3 \times 2$ analysis of variance (ANOVA) tested for differences between the three WRRT conditions for the two groups (Irlen-treatment and non-Irlen). There was no main effect of condition ($F[2, 52] = .34, p = .71$), or of group ($F[1, 26] = 1.49, p = .23$), and no significant condition $\times$ group interaction ($F[2, 52] = 2.08, p = .14$). The filters made no measurable difference to the reading rate in either group.

This overall lack of filter benefit on reading rate replicates, in this subsample, the findings made in our original study at the time of filter prescription. Figure 1 shows the filter benefit (the number of words per minute more that were read with the prescribed filter than with the colorless filter) for each child both for the original study and at follow-up. The filter
benefit scores are scattered evenly around zero for the original study, and also for the follow-up, with the exception of two Irlen-treatment children who showed a large positive filter benefit at follow-up. These two children are responsible for driving the numerical trend toward higher prescribed filter scores apparent in Table 1. However, the overall group effect at follow-up remains nonsignificant, indicating that the clinical significance of the filters is low. Moreover, there is no significant correlation between the filter effect in 2010 and in 2011 ($\rho[26] = .16, p = .40$), indicating a very low test–retest reliability of any filter effect.

To compare the progression of reading rate between the two groups, across the year, we analyzed WRRT scores from the Irlen-treatment group reading in the prescribed filter condition and the non-Irlen groups reading in the colorless overlay condition (we would expect the groups to be most comfortable in these conditions). Using a $2 \times 2$ ANOVA to test for effects of group (Irlen-treatment, non-Irlen) and time (first study, follow-up), we found only a significant main effect of year ($F[1, 26] = 12.68, p = .001, \eta^2_p = .33$), confirming that the reading rate increased between study and follow-up. There was no significant main effect of group ($F[1, 26] = 1.45, p = .24$), and no significant year × group interaction ($F[1, 26] = .12, p = .74$), indicating that the improvement in reading rate across the year was similar between the two groups. This is not consistent with a prior obstacle to reading progress having been removed for the Irlen-treatment group. These results are illustrated in Figure 2a.

**GORT Results**

The mean GORT fluency, comprehension, and overall ORQ scores are shown for each group at each time in Table 2. The overall ORQ scores are also plotted in Figure 2b. A $2 \times 2$ ANOVA showed a significant main effect of group ($F[1, 26] = 14.93, p = .001, \eta^2_p = .37$), and of time ($F[1, 26] = 7.23, p = .01, \eta^2_p = .22$), indicating that the Irlen-treatment group ORQ was below that of the non-Irlen group, and that ORQ decreased in both groups during the year. Crucially, no significant group by
time interaction was found ($F[1, 26] = .14, p = .71$), indicating that the groups’ reading performance changed comparably across the year. This is not consistent with a prior obstacle to reading progress having been removed for the Irlen-treatment group. Finally, a correlation between the ORQ scores in the first study and at follow-up was strong and highly significant, confirming the test–retest reliability of this reading measure ($r[26] = .87, p < .001$).

One potential complication in comparing the GORT scores across time is that in the first study half of the children were administered the GORT with a colored overlay and half without, as dictated by the between-groups design of that study. For the present Irlen-treatment group, 8 had used their prescribed overlay at the first assessment, and 10 had not. It should be noted that any bias caused by this factor would have been to the advantage of the Irlen-treatment group at follow-up, as all were using their prescribed filter at follow-up, while fewer than half had used their prescribed filter in the original study. Given that the Irlen-treatment group did not perform differentially better at follow-up, this potential source of bias seems to have had no relevant influence. Nonetheless, in order to double check, we ran an additional repeated-measures ANOVA, confined to the Irlen-treatment group, to compare the effects of time across the subgroups that had and had not used their prescribed filter at the first GORT assessment. This confirmed the main effect of time ($F[1, 16] = 5.30, p < .05$), reflecting the reduction in GORT ORQ across the year, but found no main effect of subgroup ($F[1, 16] = .31, p = .59$), nor time by subgroup interaction ($F[1, 16] = 3.64, p = .08$). This provides further reassurance that the variation in the original GORT testing condition did not distort the present results.

**Debrief**

At debrief, all 18 of the Irlen-treatment group children reported positive subjective effects of the overlays. Fourteen children reported a reduction of specific purported IS symptoms—for example, “the words stopped moving around on the page,” “the page is less blurry,” “the overlay stops the white page from hurting my eyes,” and “the page feels zoomed in”—while four reported general statements of preference—for example, “I prefer reading with this overlay,” and “this overlay makes my reading much better.”

**DISCUSSION**

This follow-up study examined the effects of Irlen colored filters after 1 year of use in the same sample tested in our original study (Ritchie et al., 2011). Using the same reading tests—the WRRT and the GORT—we reassessed the children who were still using their colored overlay or colored lenses, and compared them with those who had not been diagnosed with IS. To our knowledge, this is the first 1-year follow-up study on the Irlen treatment. On the WRRT, no significant differences were found between reading with colored filters and colorless overlays, and no between-group differences in reading progress across the year were found. On the GORT, both groups significantly decreased in skill across the year, and again, no differences in the trajectory between the groups were found.

From a practical perspective, it is useful to note that around 53% of children diagnosed with IS and prescribed a colored filter will discontinue the treatment within 1 year. Continuation rates for Irlen overlays have not been previously reported, but research using Intuitive overlays—a comparable treatment—in a sample with normally distributed reading ability has found very similar continuation rates (Wilkins et al., 2001).

The most immediate and parsimonious explanation of the WRRT results is that the Irlen overlays have made no difference; both groups improved their reading rate to a similar degree. This was despite all of the Irlen-treatment

**Table 2**

Mean Gray Oral Reading Test Scores for Each Group for Both Years

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Irlen-treatment</td>
<td>3.94 (2.41)</td>
<td>3.11 (2.08)</td>
<td>7.56 (2.43)</td>
<td>7.44 (1.72)</td>
<td>74.83 (11.87)</td>
<td>71.67 (10.09)</td>
</tr>
<tr>
<td>Non-Irlen</td>
<td>7.70 (2.71)</td>
<td>7.20 (2.15)</td>
<td>9.30 (1.57)</td>
<td>8.40 (1.90)</td>
<td>91.00 (10.10)</td>
<td>86.80 (10.88)</td>
</tr>
<tr>
<td>Total</td>
<td>5.29 (3.08)</td>
<td>4.57 (2.87)</td>
<td>8.18 (2.29)</td>
<td>7.79 (1.81)</td>
<td>80.61 (13.60)</td>
<td>77.07 (12.56)</td>
</tr>
</tbody>
</table>

Note: Standard deviations are given in parentheses. Comp. = comprehension; flu. = fluency; ORQ = oral reading quotient.
group children indicating they felt a subjective benefit when reading with overlays. The placebo effects that might be expected to accompany such beliefs were evidently not strong enough to improve reading scores on either of the tests while using the preferred overlay.

Only two children had large “overlay difference” scores at follow-up, as shown in Figure 1; the outlying individual with the largest score stopped reading in the colorless overlay condition after an average of 26 s, stating that her eyes were tired. This was not the case in the original study, when she was able to complete reading for a full minute in both iterations of the colorless overlay condition. In addition, as can be seen in Figure 1, this individual read substantially faster in the colorless overlay condition in the original study (i.e., had a negative overlay effect score), suggesting her performance was highly variable.

As the GORT ORQ results are age-standardized, the ORQ score should remain constant across time, provided that the children are improving in line with the standardized norms. In fact, while the original and follow-up ORQ scores were highly correlated, they significantly decreased in both groups across the year. This may reflect a genuine loss of global reading ability. As noted above, given that the Irlen treatment is intended to remove a barrier to reading ability (Irlen, 2010), one would expect that the individuals with IS who had been using the Irlen treatment for a full year would no longer be disrupted by the visual symptoms of the disorder, and would make concomitant gains in their scores disproportionate to those of the non-Irlen group. That this has not occurred, and the two groups changed parallel to one another (Figure 2b), suggests that the overlays have made no measurable difference to the global reading ability of the IS-diagnosed children in our sample.

Irlen (2010) has claimed that colored filters are often life-changing, resulting in greatly improved reading ability and thereby improved educational attainment. Such claims predict large effect sizes for the treatment. While the present study had a small sample size, it had adequate power to detect medium-to-large effects of the order required for practical or educational significance; these effects should have been particularly large in the within-subject filter comparison. Smaller effects might have remained undetected by our study, but we do not believe these would have had important practical significance (Ferguson, 2009).

Our study has a number of limitations beyond its small sample size. First, the control group of poor readers not diagnosed with IS is not the ideal comparison, which would be a group of children diagnosed with IS who did not receive Irlen treatment. Unfortunately, no such sample was available. Second, the small number of controls limits the generalizability of the between groups comparison. Nevertheless, the similarities in reading trajectories between the Irlen-treatment and control groups (Figure 2) are striking. Third, Irlen (2010) claims that colored filters can have effects other than improving reading ability, such as a reduction in headaches or visual discomfort. While the children in the present study were included on the basis of their poor reading ability, it is possible that the filters had other effects across the year that the study did not measure. Finally, it should be noted that conclusions drawn from our results are limited to the Irlen Institute’s colored filter method; other methods, such as the Intuitive system (Wilkins, 1994) exist, and utilize different diagnostic techniques. On the other hand, our study has real-world relevance and value, being a direct assessment of a school-level intervention for reading, as conducted by the Irlen Institute.

Data from this study suggest that the filters provided no benefit to reading ability, at least in the vast majority of cases, even after 1 year of continued use. These results, combined with those from our previous study (Ritchie et al., 2011), imply that the Irlen colored filter treatment does not have a statistically or clinically significant effect on reading, either in the short or longer term.

Acknowledgments—Our thanks go to the staff and pupils at Newark Primary School and Port Glasgow High School, Port Glasgow, Scotland. We are also grateful to two anonymous reviewers for their suggestions. Stuart J. Ritchie is supported by an ESRC +3 PhD Scholarship.

NOTE

1 Null results were also obtained using similar analyses between all filter conditions.

REFERENCES


