Unfixed reference, monocular occlusion, and developmental dyslexia—a critique

D V M Bishop

From the Department of Psychology, University of Manchester

SUMMARY Stein and Fowler have proposed that poor binocular control of vergence eye movements is responsible for reading problems in a subset of dyslexic children, and that this subgroup is characterised by unstable performance on Dunlop’s reference eye test. Four predictions from this hypothesis are evaluated in the light of published evidence. First, it is shown that a substantial minority of good readers have unfixed reference. Second, the evidence for a raised prevalence of unfixed reference in dyslexics is reviewed and contradictory findings are discussed. Third, it is argued that there is little support for the view that dyslexics with unfixed reference make different types of reading errors from those with fixed reference: indeed many dyslexics with unfixed reference have non-visual, phonological difficulties. Finally, it is argued that studies which claim that monocular occlusion is a successful treatment for ‘visual dyslexia’ are methodologically flawed and do not provide convincing evidence for this view.

It has long been recognised that low intelligence or lack of opportunity cannot explain all cases of difficulty in learning to read. Although there have been a number of cogent objections to the use of this label,¹ the term ‘developmental dyslexia’ remains popular, and will be used here to refer to severe reading problems not attributable to sensory or intellectual impairment or lack of opportunity. Many early writers thought developmental dyslexia resulted from an impairment in the visual perception of written words,² but research over the past 20 or 30 years has not supported this idea. Cognitive deficits are reliably found in dyslexic children, but these are verbal rather than perceptual, suggesting that the basis of the reading problem is difficulty in reliably classifying, remembering, and/or segmenting speech sounds.³,⁴

However, it may be mistaken to look for a single explanation for dyslexia. Reading involves a complex range of cognitive processes, and there may be several types of dyslexia, each arising for a different reason and for which different types of remedy would be appropriate. Attempts have been made to subclassify developmental dyslexia,⁵ but there has been no consensus on a classification and no agreed criteria for distinguishing subtypes. Among those arguing for a distinction between dyslexic subtypes are Stein and Fowler,⁶,⁷ who proposed that a subset of children with ‘visual dyslexia’ have abnormalities of visuomotor integration detectable by a ‘reference eye’ test developed by Dunlop et al.⁸ They maintained that monocular occlusion helps stabilise performance on the Dunlop test and can lead to dramatic gains in reading ability of visual dyslexics. If these results are repeatable, then they are important in indicating an effective treatment for at least a subset of dyslexic children. However, a number of logical and methodological problems have been raised concerning this work.

Stein and Fowler: the theory

The claim made by Stein and Fowler is that unstable control of vergence eye movements with small targets is a sufficient, though not necessary, cause of reading difficulty. They regard ‘unfixed reference’ on the Dunlop test as an indicator of unstable vergence control. Four predictions follow from this:

(1) Children with unfixed reference will be poor readers.

(2) Not all poor readers will have unfixed reference, but there will be a raised prevalence of unfixed reference in dyslexic populations.

(3) Dyslexic children with unfixed reference
will make reading errors characteristic of 'visual dyslexia', unlike other dyslexics who will be more likely to have linguistic deficits.

(4) Procedures which stabilise performance on the reference eye test will be effective treatment for visual dyslexia.

Origins of Dunlop's reference eye test

Dunlop et al. described a new test which they regarded as a neurologically meaningful measure of eye dominance. The subject views two slides (Clement Clarke fusion slides F69 and F70) simultaneously, one with each eye, in a synoptophore. Each slide shows a house: one has a small tree on one side of the door, and the other has a large tree on the other side. When fusion is achieved, the subject sees a single house with a small tree on one side of the door and a large tree on the other, in the central 1° to 2° of the binocular field. If angle of the slides is increased or decreased smoothly and slowly as the subject continues to fixate on the door, most subjects report that one of the trees moves before fusion breaks. The stable image specifies the 'reference eye'.

Dunlop et al. reported that dyslexia was associated with 'crossed reference', that is, discrepancy between reference eye and handedness. However, later workers noted that many children do not perform consistently on this test and so cannot be said to have a reference eye. It is this lack of stable reference, rather than crossed reference, that Stein and Fowler regard as associated with dyslexia. They reinterpret the neurophysiological significance of this test, regarding 'unfixed reference' as reflecting poor control of vergence eye movements to small targets.

Is there a link between unfixed reference and reading ability in normal populations?

Bishop et al. used the Dunlop test in a study of 147 8½-year-old children registered at a rural general practice. No relationship was found between crossed reference and dyslexia, but there was a significant association between reading attainment and inconsistent performance on the Dunlop test. The 44 (32.1%) children who did not give consistent responses to the test or who did not report movement of either tree (unfixed reference) had a mean reading age 4.16 months below age level, whereas the 93 children with fixed reference (crossed or uncrossed) had a mean reading age 1.24 months above age level. However, those with unfixed reference had a lower mean IQ than those with fixed reference, and the difference in reading scores was no longer significant when IQ was used as a covariate.

It was concluded that the association between reading attainment and unfixed reference arose because less intelligent children had difficulty in maintaining fixation and discriminating between small indicators.

Newman et al. studied 323 children with a mean age of 8½ years, all of whom had WISC-R verbal and performance IQs of 90 or over. The orthoptist in this study was trained by Stein and Fowler to administer the Dunlop test, which was used to divide children into those with fixed reference (dominant eye on same side on eight or more out of 10 trials) or unfixed reference. Those with fixed reference (n=158) and unfixed reference (n=165) did not differ in age, mental age, or reading or spelling attainments.

Stein et al. gave the Dunlop test to 753 5- to 11-year-old children attending normal primary school. For 451 of these children reading test results were available. All had IQs above 90 on class-administered tests. There was a strong relationship between age and performance on the Dunlop test, with the proportion of children with stable reference increasing steadily from 52% in 5-year-olds to nearly 90% in 10-year-olds. Like Bishop et al. they found that around two-thirds of 8-year-olds had stable reference. Those 6- to 11-year-old children who had achieved stable reference were reading on average 6.3 months better than those who had not. Data on IQ were not reported.

Stein et al. argued that methodological differences were responsible for the failure of other studies to find an association between reading and Dunlop test performance, noting that considerable skill is necessary to obtain reliable results with the Dunlop test in children. However, there is an alternative explanation. Newman et al. analysed reading ages in relation to mental age, thus correcting for any IQ difference between children with fixed and unfixed reference: Stein et al. did not. Bishop et al. found that when no correction was made for IQ there was a significant difference between those with fixed and those with unfixed reference. However, when correction was made for IQ this difference disappeared. There is good evidence that both chronological and mental age affect performance on the Dunlop test, and it seems that results from all three studies can be explained in terms of these factors.

One point made by all three studies is that substantial numbers of normal readers have unfixed reference: for example, in the study by Stein et al. 24% of children whose reading was ahead of age level had unfixed reference on the Dunlop test. These findings pose a logical problem for any theory that maintains that lack of fixed reference on the Dunlop test is a sufficient cause of reading difficulties.
Unfixed reference and dyslexia

Is there a link between Dunlop test performance and dyslexia?

Stein and Fowler\textsuperscript{11} compared 80 dyslexics with 80 above average readers. Their mean age was just over 10 years. On the Dunlop test 63\% of the dyslexics but only one normal child had unstable reference, a highly significant difference.

Stein and Fowler’s 1985 study of 148 dyslexic children\textsuperscript{2} did not include a normal control group, but their data may be compared with those of their normative study.\textsuperscript{16} Stein (personal communication) confirms that the mean age of the dyslexic group was just over 9 years. In the normative study less than 20\% of 9-year-olds had unstable reference, whereas this was found in nearly 70\% of the dyslexics.

This result is in striking contrast to those of Bishop \textit{et al.} and Newman \textit{et al.} One might wonder whether the discrepancy arises because of differences in study design: when the base rate of a disorder in a normal population is low, true associations may be difficult to detect because the number of disordered children is small. For instance, if one estimates that dyslexia affects around 5\% of children, then only about seven children in the study by Bishop \textit{et al.} would be affected, and an increased prevalence of unfixed reference might be difficult to detect in such a small sample. However, both Bishop \textit{et al.} and Newman \textit{et al.} did supplementary analyses of their data in which they selected from their samples those children with poor reading scores relative to intelligence. In neither case was there any suggestion of a higher rate of unfixed reference in this group. Nor was unfixed reference remarked upon by Dunlop \textit{et al.}\textsuperscript{13} in their original study of 15 dyslexics. Apart from work by Stein and his colleagues the only other study to find this association was by Bigelow and McKenzie,\textsuperscript{15} who reported that nine out of 14 (64\%) dyslexics had unstable performance on the Dunlop test, compared with only three members (21\%) of a control group, composed of younger children matched on reading age.

If for the moment we ignore the negative findings and accept that a true association may exist between unfixed reference and dyslexia, we then have to consider what this might mean. One explanation is that visual problems cause dyslexia, but this is not the only possibility. There could be a causal link in the opposite direction, so that experience of reading affects performance on the Dunlop test. But this interpretation is difficult to reconcile with the data from Bigelow and McKenzie, and with a study by Stein \textit{et al.},\textsuperscript{16} which showed that dyslexics were impaired on vergence eye movement control even when compared with younger normal children with whom they were matched on reading ability.

Another possibility is that unfixed reference may be a correlate of dyslexia without playing any causal part in reading problems. There have been previous demonstrations of poor sensorimotor skills in children with specific developmental disorders, including dyslexia.\textsuperscript{18-24} One interpretation is that motor immaturity and delayed language development are indicators of a single underlying cause—delay of neurological maturation.\textsuperscript{29} An alternative, though not incompatible, interpretation is that attentional disorders associated with dyslexia disrupt performance on a range of tasks.\textsuperscript{26} Unfixed reference could be regarded as a further example of this type of association, reflecting neurological immaturity and/or poor attention in dyslexic children.

Differences between dyslexics with fixed and unfixed reference

As well as finding a dramatic increase in the frequency of unfixed reference in their dyslexic group Stein and Fowler\textsuperscript{11} reported that all the dyslexics with unfixed reference were ‘visual dyslexics’. However, they gave no criteria for ‘visual dyslexia’; it is not clear who made this diagnosis (the two reports of this study differ on this point), and they did not state how many ‘visual dyslexics’ had fixed reference. Altogether the published account is too unclear for us to evaluate the claim that performance on the Dunlop test relates to type of dyslexia.

Rather more detail is provided in the study by Stein and Fowler,\textsuperscript{15} in which 148 dyslexic children were treated with monocular occlusion. Forty-five of these children had fixed reference when first assessed. All children were seen by a reading specialist who ‘arbitrarily categorised each child according to the predominant type of error he/she made when performing simple reading, writing, sequencing, rhyming, and alliteration tasks’. Visual errors were defined as losing place on the page, having to point with a finger to keep the place, missequencing, and reversing and rotating letters and words when reading or writing. Phonemic errors involved failure to find rhymes or alliterations for 10 common words, and sequencing errors were failure to name the days of week or months of year. (Note that the behaviours regarded as evidence of visually based reading problems are open to other interpretations.\textsuperscript{28} It is unfortunate that no task of non-verbal visual discrimination or sequencing was included.)

The relationship between the classification of children as ‘visual’ or ‘non-visual’ dyslexics and performance on the Dunlop test was not reported. Apparently this was because dyslexic children did not neatly subdivide into visual and non-visual types. Many children made both ‘visual’ and ‘auditory'
errors. It is impossible to discover from the published account how error type related to unfixed reference, since these ‘mixed dyslexics’ were first excluded and only then were the remaining children compared on the basis of Dunlop test results.

If unfixed reference is an indication of a visual basis for dyslexia, then one would expect those with unfixed reference to have less evidence of early speech delay. However, dyslexics with fixed and unfixed reference did not differ in experience of speech therapy.

In order to accept that the Dunlop test distinguishes subtypes of dyslexic children, one would need to show that a blind, binary classification of dyslexics as ‘visual’ or ‘non-visual’ on the basis of reading errors predicted Dunlop test performance, and/or that children with unfixed reference made significantly more visual errors and significantly fewer non-visual errors than those with stable reference (without first excluding children who did not fit the hypothesis). Such analyses have not been reported. The data presented by Stein and Fowler pose problems for their theory in demonstrating that many dyslexic children with unfixed reference have verbal difficulties that extend beyond reading.

**Treatment of reading problems by monocular occlusion**

The most important claim for a causal link between unfixed reference and reading problems is that dyslexic children with unfixed reference can be helped by monocular occlusion. This claim is made in the two studies of dyslexic children by Stein and Fowler described above.

In the first of these Stein and Fowler studied 30 ‘visual dyslexics’ who were asked to wear spectacles with the left lens occluded for all reading and close work. The attainments of these children were compared with those of 30 normal readers and 15 untreated visual dyslexics. After six months the reading age of the treated children had increased by an average of 13.5 months, compared with an increase of only 3.94 months in the untreated group: a statistically significant difference.

Unfortunately, the account of the study lacks details that would make a proper evaluation possible. We are not told the ages of the different groups of children, though it is evident from other studies that age is a crucial variable in determining Dunlop test performance. No details are given of the criteria for matching controls and dyslexics. The variance of reading scores in the treatment phase is not stated.

A further problem is that, as Stein and Fowler recognise, their result could reflect a placebo effect. Children and their parents and teachers were aware whether they were in a treatment or control group, and this could have affected outcome.

In 1985 Stein and Fowler reported a study designed to overcome the methodological problems of their earlier work. 148 dyslexic children were randomly assigned to treatment and placebo groups. Those in the treatment group were given spectacles with one lens occluded, the remainder were given plain spectacles. After six months the children’s reading was assessed, and the spectacles were removed from children whose reference remained or became fixed. All who still had unfixed reference continued to wear or were then given occluding spectacles to wear. After a further six months this procedure was repeated.

Monocular occlusion was significantly associated with development of fixed reference: this occurred in 51% of those treated with monocular occlusion and 24% of the placebo group. Furthermore, there was a significant difference in reading scores between children who established a reference eye during the course of the study and those who did not.

These data look impressive, but, as with the 1981–2 study, there are some methodological problems, and evaluation is hampered by omission of a number of important details, such as age, initial reading level, and IQ of the different groups. Data from one group of children (unfixed-occluded-unfixed-occluded-fixed) were not given, and it is often unclear how the data were analysed.

Further, Wilsher pointed out that, although assignment of children to treatment was described as random, those with unfixed reference were much more likely to have been treated with occlusion. This suggests that the person assigning children to treatment groups believed in the efficacy of occlusion and was reluctant to deny children with unfixed reference the opportunity of treatment. Stein et al. answered by stating that the imbalance arose because 53 children were lost from the study after the spectacles had been handed out. These children constitute a very high proportion of the study population, and one might ask why they were not mentioned in the original paper. This seems especially surprising, since the authors described another 53 children who did not participate in the study, 24 because they were not dyslexic, 9 with ophthalmological conditions, and 20 who failed to return for follow-up. This sort of omission not only weakens one’s confidence in the report of this study, it also introduces a worrying possibility of bias. If there was a marked difference between treated and placebo groups in the frequency of drop-outs, then this suggests that parents and children were not blind to treatment status. Stein et al. suggest that it was the lack of improvement associated with placebo rather than awareness of
Unfixed reference and dyslexia

Table 1  Outcome data after the first six months

<table>
<thead>
<tr>
<th></th>
<th>UOU 29</th>
<th>UNF 11</th>
<th>UNU 28</th>
<th>UOF 31</th>
<th>FNF 29</th>
</tr>
</thead>
<tbody>
<tr>
<td>n</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>111.0</td>
<td>118.2</td>
<td>110.0</td>
<td>113.6</td>
<td>111.3</td>
</tr>
<tr>
<td>SD</td>
<td>12.14</td>
<td>15.54</td>
<td>15.63</td>
<td>17.59</td>
<td>12.83</td>
</tr>
<tr>
<td>Initial reading age</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>79.9</td>
<td>83.5</td>
<td>80.8</td>
<td>88.5</td>
<td>83.0</td>
</tr>
<tr>
<td>SD</td>
<td>10.13</td>
<td>9.49</td>
<td>9.99</td>
<td>10.65</td>
<td>8.16</td>
</tr>
<tr>
<td>Reading T score† (initial)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>34.1</td>
<td>34.0</td>
<td>34.7</td>
<td>38.5</td>
<td>35.3</td>
</tr>
<tr>
<td>SD</td>
<td>4.61</td>
<td>3.52</td>
<td>4.17</td>
<td>5.80</td>
<td>4.25</td>
</tr>
<tr>
<td>Matrices T score</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>46.1</td>
<td>44.4</td>
<td>47.8</td>
<td>47.9</td>
<td>48.5</td>
</tr>
<tr>
<td>SD</td>
<td>8.32</td>
<td>6.59</td>
<td>6.21</td>
<td>10.15</td>
<td>9.62</td>
</tr>
<tr>
<td>Similarities T score</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>50.8</td>
<td>47.9</td>
<td>50.4</td>
<td>53.9</td>
<td>49.5</td>
</tr>
<tr>
<td>SD</td>
<td>7.57</td>
<td>9.21</td>
<td>8.36</td>
<td>7.79</td>
<td>7.14</td>
</tr>
<tr>
<td>Reading age gain (months)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>5.6</td>
<td>7.1</td>
<td>5.7</td>
<td>11.6</td>
<td>8.8</td>
</tr>
<tr>
<td>SD</td>
<td>4.71</td>
<td>7.70</td>
<td>6.81</td>
<td>9.94</td>
<td>8.18</td>
</tr>
</tbody>
</table>

* Dunlop test coding: F = fixed reference (stable vergence); U = unfixed reference (unstable vergence).
† Treatment coding: O = occluded; N = placebo or untreated.

being in a placebo group that caused children to drop out, but if that were the explanation, then more children with fixed reference should have dropped out, as they too were making poor progress.

In response to requests for further details about this study, Dr. Stein kindly agreed to provide the raw data for reanalysis to investigate how far variables such as intelligence, age, and initial reading level may have affected gains in reading age. The data from the first six months of the study are summarised in Table 1. There are some discrepancies between the data given here and those reported in the *Lancet*. In particular, the standard errors in the published account are much smaller than those derived from standard deviations shown here. Dr. Stein has confirmed that these arose through computational error and the figures given here are the correct ones.

The prediction from Stein and Fowler's theory is not that occlusion will be effective for everyone, but rather that its efficacy will depend on initial status. Thus if children are divided on the basis of initial status (fixed (F) or unfixed (U) reference) and treatment (occluded (O) or placebo (N)) the prediction is that there will be an interaction, so that treatment has a significant effect only on the initially unfixed group. When data for gain in reading age are analysed in a two-way analysis of variance neither main effect is significant (F ratios less than 1), and nor is the interaction (F(1, 140) = 2.92). A supplementary analysis just on children who did not initially have stable reference also gives a non-significant difference between treated and untreated groups (F(1, 96) = 2.9). Thus we cannot reject the null hypothesis that treatment has no effect on reading.

Nevertheless, on the basis of a different type of analysis, Stein and Fowler argue that their results do support the theory. It seems reasonable to suppose that occlusion would improve reading scores only if it brings about fixed reference. Stein and Fowler therefore subdivided both treated and untreated groups of children according to whether stable reference was achieved after six months. When this is done, significant differences in reading age gains can be shown between groups. If the six groups shown in Table 1 are compared on reading age gain in a one-way analysis of variance, a significant main effect is found (F(5, 138) = 2.71; p<0.05). If groups UOU and UNU are combined and compared with groups UOF and UNF a significant difference in reading age gain is found (t(96) = 3.06; p<0.01), though there is no significant difference between UOF+UNF versus FOF+FNF (t(86) = 1.29, NS). It is on this basis of the result that Stein and Fowler maintain that attainment of stable vergence eye movements facilitates reading development.

As Wilsher noted, there are problems with this analysis. In dividing subjects by reference eye status at follow-up, one has moved from classification in terms of an independent variable over which there is experimental control to classification in terms of a dependent variable. This makes interpretation of results ambiguous: performance on the Dunlop test at follow-up may relate to reading outcome, but we cannot be certain of the direction of causality, or indeed that any causal relationship is implicated. Other factors, discussed above, such as influence of reading experience on Dunlop test performance, could be responsible for this relationship. Alternatively, some third variable, such as rate of matura- tion, might affect both stability of reference eye and reading ability without the two being directly causally related.

Scrutiny of the data presented in Table 1 suggests that this latter interpretation may be the correct one, and that initial reading status might be a crucial variable determining outcome. When *initial* reading ages of all six groups are compared in a one way analysis of variance, a significant main effect is found (F(5, 138) = 3.31; p<0.01), and a specific comparison between groups UOU + UNU vs UOF + UNF with a t test shows a significant difference in favour of those with fixed reference at follow-up (t(96) = 3.3; p<0.01). Thus those groups that show most gain in reading ability over the first six months also start off with higher reading ages. An analysis of covariance was carried out to compare reading age gain scores
for the six groups shown in Table 1 after scores had been adjusted for initial reading age. The main effect did not reach significance (F (5, 137)=1.54).

Reading age is a function both of chronological age and of severity of dyslexia. By transforming reading age scores into T scores one can examine the effect of severity of reading problem independently of age. The correlation between initial reading T score and reading age gain over six months is statistically significant (r (142)=0.242; p<0.01), as is the correlation with total reading age gain over the whole 12 months of study (r (142)=0.439; p<0.01). When the analysis of covariance was repeated using initial T score as covariate when comparing reading age gains of the six groups in Table 1, the main effect again did not reach significance (F (5, 137)=1.68). This confirms that the degree of initial reading impairment can account for differences in outcome between groups. Regardless of age, the more severe a child’s initial problem, the less likely he is to make good progress.

The study continued for a further six months, with a partial replication of the first phase. All children who were found to have unstable reference were treated with occlusion, whereas occlusion was discontinued for those with stable reference. Results for this second phase of the study are shown in Table 2.

As all ‘unfixed’ children were treated at this stage, it is not possible to compare treated and untreated groups, but we can subdivide children in terms of whether stable reference is achieved after occlusion, by comparing the two of the initially untreated groups, namely, UNUOF and UNUOU. The amount of reading age gain made in the second phase of the study is not significantly different for these two groups, regardless of whether the T score at the start of the second six months is entered as a covariate (Anova: F (1, 26)=1.3, NS; Ancova: F (1, 25)=2.4 NS). Thus the difference between those who developed fixed reference and those who did not which was observed in the first phase of the study was not replicated in the second phase.

Finally, when the eight groups shown in Table 2 are compared in an analysis of variance on total reading age gain made over 12 months, no significant difference is found (F (7, 136)=1.77: the F ratio is even smaller if the initial reading T score is entered as a covariate).

In conclusion there is no evidence that monocular occlusion in children with unfixed reference results in improved reading scores. When direct comparisons are made between treated and untreated children, differences in amount of reading progress are not statistically significant. There is a significant relationship between development of stable reference and improvement in reading, but this is explicable in terms of differences in initial reading ability. Those children who developed stable reference in the first phase of the study tended to be those whose initial reading problems were less severe.

This work is based on a paper presented at a meeting of the Academia Rodimensis Pro Remediacione at the Royal Society on 16 October 1987. I would like to thank Dr John Stein for inviting me to that meeting in order to put forward a critical view of his work, and for generously providing his raw data for analysis.

References
Unfixed reference and dyslexia


Accepted for publication 25 February 1988.