WHAT CHILDREN SEE AFFECTS HOW THEY READ

Piers Cornelissen
Lynette Bradley
Sue Fowler
John Stein

There is good evidence that children learning to read make rapid progress if they score highly on tests of phonemic awareness (Liberman et al. 1974, Lundberg et al. 1980, Wagner and Torgeson 1987). Conversely, many children who have difficulty learning to read are markedly poor at such tests (Bradley and Bryant 1978, Snowling 1981). Because of the powerful relationship between children’s phonological skills and their success in reading, this is one of the first areas of investigation when a child is referred with reading problems.

However, children with reading problems sometimes describe their difficulties in terms of what they see happening to letters and words on the page. For example, one 10-year-old boy reported to the first author that print often appeared ‘blurry’ or ‘zig-zaggy’. When asked to explain what he meant, he said that letters from the same line of text, or from the lines above or below, sometimes ‘moved over each other’. Often he would see a string of letters that were very different from those actually printed. He said that this ‘blurring’ might happen two or three times on any line of text, and was prevented by covering one eye with his hand. Routine ophthalmological tests, such as Snellen visual acuity, are usually normal in children like this boy, yet such reports are so suggestive of visual problems that they warrant further investigation.

Two kinds of study have demonstrated correlations between poor reading and poor performance on a variety of visual tasks, especially in six- to eight-year-old children (Willows 1990). In the first kind, attention was paid to the way in which retinal images are processed or analysed (Breitmeyer 1980, Lovegrove et al. 1986, Geiger and Lettvin 1987, Ruddock 1990). In the second, correlations between reading performance and eye-movement control were found (Zangwill and Blakemore 1972, Pavlidis 1981, Stein and Fowler 1982, Bigelow and McKenzie 1985). Unfortunately, as Hulme (1988) and Bishop (1989) have pointed out, these studies do not tell us whether poor visual processing affects reading performance, or whether cause and effect are the other way around—poor reading may impair performance on the visual tasks.

One way to find out whether impaired visual processing affects the way children read is to create a situation in which the visual system is stressed when they are reading. If we can show that children with impaired visual processing read differently when their vision is stressed, but that the same stress has no effect on children with no apparent visual impairment, we can suggest a direct link between the efficiency of visual processing and reading.
Two studies of adults have shown that the size of print can be important for reading. Erdmann and Neal (1968) and Blommaert and Timmers (1987) demonstrated that the threshold for recognizing words and letters depends on print size. Their adults found it more difficult to recognize words and letters when the print size was reduced. In another study, Parker (1972) found that school-children (nine to 13 years old) held the books they were reading closer if the print was smaller. These findings suggest that reduced print size taxes the visual system, so we used this to stress children’s vision in the present study.

One question we had to address was the way in which we might expect visual perceptual problems to affect children’s reading. Visual impairment could cause the print to be perceived as a scrambled or nonsensical version of what is actually on the page. Frith (1985) suggests that during the earliest stages of learning to read, children adopt a logographic strategy, in which ‘letter order is largely ignored and phonological factors are entirely secondary’; they may recognize words from salient graphic features alone. So when a child with visual impairment sees a scrambled version of a real word, s/he might refuse to attempt it or make a wild guess, which could be nonsense. Later on in their development, Frith says that children move into an alphabetic and then an orthographic stage. Children in the alphabetic stage try to work words out using a grapheme to phoneme conversion strategy; they probably look separately at each component in a word they do not know. At this stage, visually impaired children could be at a particular disadvantage: they might well be expected to translate a sequence of letters they perceive, incorrectly, to be a non-word.

However, Gough and Walsh (1991) and Baldwin (1990) have also described how children who have begun to use a grapheme-phoneme strategy for reading can make non-word errors for phonological reasons alone. This can happen when a child in Frith’s alphabetic stage of reading development misapplies grapheme-phoneme conversion rules. For example, if asked to read the word ‘sphere’, s/he may accurately identify two graphemic units, ‘sph-’ and ‘ere’, but might incorrectly translate the first unit as ‘spuh’ instead of ‘sf-’, and thus might utter the non-word error ‘spuh here’. For any given level of reading ability, effects of this kind would alter, depending on the linguistic complexity of words. Therefore in our experiment we took care to control for linguistic factors: we kept the mean word frequency, mean number of letters and syllables per word, mean word concreteness and spelling regularity constant in lists of words we gave the children to read.

In the present study, children in the early stages of reading development were given three lists of single real words to read, matched for linguistic complexity. With each new list, print-size was progressively reduced. We expected this stress to evoke more non-word errors as print shrank by children with visual perceptual problems, but not to alter the proportion of non-word errors made by children without visual problems.

**Method**

The study group was selected from children referred for orthoptic assessment because of suspected reading difficulty, so our sample was biased in favour of children with visual problems. We used the Dunlop Test (DT) to identify a group of 45 (37 male, eight female) visually disabled children. We compared them with another group of 45 children (32 male, 13 female), who were matched for reading age, but who did not have a visual problem. Details of the two groups are given in Table I.

**Orthoptic and psychological assessment**

Every child was examined to exclude orthoptic and gross ophthalmological pathology before performing the DT. This is a subjective test of how reliably retinal information from both eyes is associated with oculomotor signals about eye position in the head. To generate our ‘visual’ versus ‘non-visual’ experimental groups, we treated the DT as a pass/fail visual task (see Stein and Fowler (1982) for details of administration). Children pass if they give the same response in eight or more trials out of 10, i.e. they have a stable or ‘fixed’ response to the test.
Otherwise they fail, and are said to have an unstable or 'unfixed' response.

We used a shortened form of the British Ability Scales (BAS) to assess performance IQ. This was calculated from the mean of the matrices and similarities subtest t-scores. A version of Bradley and Bryant's four-word rhyme detection task (1983) was used as a phonological test. We counted the total number of correct responses out of 16 items and assumed that the larger a child's score, the better the phonological skill.

**Experimental word-list design and administration**

At the beginning of the assessment each child was given the BAS single-word reading test. Next the child was given three experimental word lists of equal difficulty, selected on the basis of the BAS reading age. The choice of which triplet of lists was restricted to one of five possibilities, equivalent to five levels of difficulty, each corresponding to a bandwidth of the BAS reading-age score. The three lists at each level were matched for frequency (Thorndike and Lorge 1944). The midpoint for each bandwidth was adjusted so that children were likely to make 50 per cent errors. This ensured sufficiently large numbers of reading errors for later analysis.

The detailed design features of the word lists were as follows:

1. Each list comprised 15 regularly spelt and 15 irregularly spelt words. The maximum number of letters and syllables in each word was 10 and three, respectively. We only chose words with a concreteness value ±1 SD from the mean of 438 (range of 158 to 670) (Coltheart 1981).
2. Every list was printed three times, once in each of three conditions: large print—24-point lower-case Helvetica font with quadruple-line spacing and a minimum of three character spaces between each word; medium print—12-point lower-case Helvetica font with single-line spacing and one character space between each word; small print—9-point lower-case Helvetica font with half-line spacing and one character space between each word. An example of the appearance of part of three lists is shown in Figure 1.

The sequence in which the reading lists were administered was determined by two factors, randomized according to a double latin square: (1) print size (three levels) and (2) order in which the lists were given (three levels). This ensured that the 45 children who failed the DT read exactly the same 4050 words as the 45 children who passed, in exactly the same conditions.

**Statistical analysis**

The experimental word lists were scored as follows: (1) the total number of errors was counted, an error being defined as a real-word attempt at the target or a non-word error; (2) the total number of non-word errors was counted, these being defined as word errors which could not be found in a standard English dictionary; and (3) the total number of refusals to attempt a word (non-attempts) was counted.

We used a multiple logistic regression model to look for DT × print size and DT × print size × rhyme task interactions to explain changes in the proportions of non-word errors that children made as print shrank. The number of non-word errors as a proportion of total errors made was entered as the dependent variable. The other variables controlled for in the model were IQ, age, reading age and phonological

<table>
<thead>
<tr>
<th>Performance IQ (BAS)</th>
<th>Dunlop Test* Passed (N = 45)</th>
<th>Dunlop Test Failed (N = 45)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean (SD)</td>
<td>107 (13·8)</td>
<td>109·6 (11·7)</td>
</tr>
<tr>
<td>Range</td>
<td>87–133</td>
<td>89–131</td>
</tr>
<tr>
<td>Reading age (BAS)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(yrs:mths)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean (SD)</td>
<td>7·9 (1·6)</td>
<td>7·8 (1·5)</td>
</tr>
<tr>
<td>Range</td>
<td>5·6–14·0</td>
<td>5·7–12·9</td>
</tr>
<tr>
<td>Rhyme task score**</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean (SD)</td>
<td>10·4 (3·4)</td>
<td>10·7 (3·1)</td>
</tr>
<tr>
<td>Range</td>
<td>3–16</td>
<td>4–15</td>
</tr>
<tr>
<td>Chronological age</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(yrs:mths)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean (SD)</td>
<td>9·3 (1·5)</td>
<td>9·5 (1·1)</td>
</tr>
<tr>
<td>Range</td>
<td>6·7–11·10</td>
<td>7·3–11·10</td>
</tr>
</tbody>
</table>

*No significant differences were found between children who passed the DT and children who failed it for any of these measures.

**Number of correct responses to 16 items.
rough smell beginning calm quarrel costume slept debt victim prayer stumble wreck fifteen threw object patient sergeant payment column religion weigh navy crumb banker agreement angel bump whisper amount bright view telegraph copper tourist slide reflection footstep shriek finish midnight swift profession short buy hold farm dog thunder scent autumn scarlet blush circus haul rescue triumph charity outfit chew bud product hang brief attack foreigner dwelling thief punish assault scholar beauty hermit chasm magnet nymph art ice stirred fantasy cruel rhapsody seduction parade phase receipt contents trend plug yawn sandy delirium wrote million meeting square idle freight alter council exposure isle antique offend rumble grizzly vacuum faint contract scheme garden

rough smell beginning calm quarrel costume slept debt victim prayer stumble wreck sergeant payment column religion weigh navy crumb banker agreement angel bump telegraph copper tourist slide reflection footstep shriek finish midnight swift profession thunder scent autumn scarlet blush circus haul rescue triumph charity outfit chew bud foreigner dwelling thief punish assault scholar beauty hermit chasm magnet nymph art rhapsody seduction parade phase receipt contents trend plug yawn sandy delirium idle freight alter council exposure isle antique offend rumble grizzly vacuum faint

Fig. 1. Appearance of reading lists (reduced scale), showing relative change in print size and line spacing. (These are not the original lists.)

TABLE II
Mean total error scores and mean non-word error scores of children who passed and children who failed Dunlop Test

<table>
<thead>
<tr>
<th></th>
<th>Dunlop Test</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Passed (N = 45)</td>
<td>Mean (SD)</td>
<td>Failed (N = 45)</td>
</tr>
<tr>
<td>Total error score</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Print size</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Large</td>
<td>11.1 (0.6)</td>
<td>9.3 (0.8)</td>
<td></td>
</tr>
<tr>
<td>Medium</td>
<td>11.3 (0.7)</td>
<td>10.3 (0.7)</td>
<td></td>
</tr>
<tr>
<td>Small</td>
<td>11.9 (0.6)</td>
<td>10.7 (0.7)</td>
<td></td>
</tr>
<tr>
<td>Non-word error score*</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Print size</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Large</td>
<td>31.3 (4.0)</td>
<td>25.2 (3.3)</td>
<td></td>
</tr>
<tr>
<td>Medium</td>
<td>30.7 (3.2)</td>
<td>35.4 (3.8)</td>
<td></td>
</tr>
<tr>
<td>Small</td>
<td>28.5 (3.8)</td>
<td>35.1 (3.6)</td>
<td></td>
</tr>
</tbody>
</table>

*Total non-word errors/total errors × 100.

with three and two levels, respectively. The regression model is:

\[
\text{Logit (non-word error proportion)} = \text{IQ} + \text{Age} + \text{Reading age} + \text{Rhyme score} + (\text{Print size} \times \text{DT}) + (\text{Print size} \times \text{DT} \times \text{Rhyme score})
\]

We used a package called GLIM (General Linear Interactive Modelling) to run the regression analysis. GLIM provides an estimate of the regression coefficient, expressed as a log odds ratio, with its standard error. Thus the significance and magnitude of the risk of each independent variable in predicting non-word errors could be calculated. Risk values were expressed as odds ratios: values greater than one represent increased risk; values less than one represent decreased risk.

Results
There was a general trend for all children, whether they passed or failed the DT, to make more errors as print size was reduced (Table II). The largest increase in the total number of errors occurred in children who failed the DT when print size was changed.
from large to medium. However, children who failed the DT showed a significant increase in the proportion of non-word errors when print size was reduced from large to medium (10.2 per cent, p < 0.05) (Table II). There was negligible change between medium and small print. In contrast, there was no significant change as print shrank in the proportion of non-word errors made by children who failed the DT. Children who failed the DT made fewer non-word errors than children who passed the DT when they read large print, but they made more non-word errors with medium and small print.

Table III shows the results of the regression analysis. Most importantly, there was a significant interaction between DT and print size (p < 0.01). There was also a significant three-way interaction between DT, print size and rhyme score (p < 0.0001).

Discussion
It has been shown that many children with reading difficulty perform less well than normally reading controls on a number of different visual tasks, but there is little direct evidence for any one of a number of possible explanations for these correlations. For example, superior performance on the visual tasks may be a result of the wider reading experience of better readers; there may be no relationship between inefficient low-level visual processing and reading; or visual problems may disturb the reading process. In this paper we have attempted to show that a visual impairment (as measured by the DT) can affect reading, at least when visual stress is applied.

Our finding that there was a trend for all children to make more errors as print size was reduced is consistent with the findings of Erdmann et al. (1968) and Blommaert et al. (1987), who showed that reducing print size makes words more difficult to read. But of greatest importance was the finding that only the children who failed the DT read differently as print size was reduced; they made more non-word errors. Therefore, since the only factor changed was print size, this result must have a visual basis. Notably, the increase in non-word errors was mostly restricted to the change between large- and medium-size print. This suggests that the visual mechanism responsible may have become saturated over this range of print size, so that no additional effect was observed when print size changed from medium to small. Furthermore, this change in error pattern was associated with a modest increase in the total number of errors. These results suggest that children who fail the DT may make reading errors for visual reasons alone.

What visual physiological mechanism could explain the non-word errors made by children who fail the DT? Reading demands frequent small changes in binocular fixation. Children who fail the DT are said to be unable to maintain binocular fixation as steadily as children who pass it. This has been referred to as unstable binocular control (Stein et al. 1987). Therefore, during fixation these children’s eyes may aim inaccurately at points in front of or beyond the plane of the page. Consequently they are likely to suffer intermittent physiological diplopia,
TABLE IV
Relationship between non-word error score, print size, phonological ability and Dunlop Test

<table>
<thead>
<tr>
<th>Print size</th>
<th>Phonological ability**</th>
<th>Mean non-word error score*</th>
<th>Dunlop test</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Pass</td>
</tr>
<tr>
<td>Large</td>
<td>Good</td>
<td>35.3</td>
<td>34.7</td>
</tr>
<tr>
<td></td>
<td>Poor</td>
<td>26.4</td>
<td>16.1</td>
</tr>
<tr>
<td>Medium</td>
<td>Good</td>
<td>35.6</td>
<td>41.4</td>
</tr>
<tr>
<td></td>
<td>Poor</td>
<td>24.8</td>
<td>29.7</td>
</tr>
<tr>
<td>Small</td>
<td>Good</td>
<td>34.7</td>
<td>44.7</td>
</tr>
<tr>
<td></td>
<td>Poor</td>
<td>21.4</td>
<td>25.9</td>
</tr>
</tbody>
</table>

*Non-word error score = (non-word errors/total errors) × 100.
**Phonological ability is ‘good’ if child made ≤5 errors, and ‘poor’ if a child made >5 errors out of 16 items in rhyme task.

in which case the letters seen by the two eyes may appear incorrectly superimposed. This situation may be further complicated by binocular retinal rivalry, when components from the non-corresponding images from the two eyes may be variably suppressed (Davson 1980). Even in adults, it is difficult to determine the eye of origin of each retinal image (Blake and Cormack 1979). It is clear that binocular instability of this sort could lead to children experiencing visual confusion and perceiving incorrect letter sequences, which they could then translate as non-words.

All our children were more likely to make non-word errors if they had good phonological skills and higher reading ages. This corroborates the findings of Gough and colleagues (1983), who used a pseudoword reading task to measure children’s phonological skill. They found that children’s ability to read pseudowords correlated positively with the proportion of non-word errors they made in an oral reading test. The strong relationship between phonemic awareness and non-word errors in our data suggests that many of our children were able to apply a phonological reading strategy. Hence they could be at Frith’s alphabetic stage of reading development.

On the one hand we have accounted for children’s non-word errors in terms of unstable binocular control. On the other hand we have appealed to children’s phonological skills to explain their non-word errors. So have we created a dilemma? We propose an explanation which reconciles these two findings in a complementary way.

Table IV shows that for each print size children always made more non-word errors if they had good rather than poor phonological skills. This discrepancy was significantly more marked for children who failed the DT. Children with good phonological skills are more likely to work out words they do not recognize at first sight by using a grapheme-phoneme strategy. To do this successfully they must look at, or fixate, each component in a word separately. As Gough et al. (1983) have shown, they may sometimes misapply grapheme-phoneme conversion and make a non-word error; the children see what is on the page, but decode it wrongly.

However, children who fail the DT and who use a phonological reading strategy may be further disadvantaged. Like other children, they may misapply phonological rules and make non-word errors, but because they have to fixate each grapheme separately, they are also at risk of suffering visual confusion, perceiving a scrambled version of what is actually on the page, which they also decode as a non-word. In short, children who fail the DT may have both a phonological and a visual/phonological route for making non-word errors, and further multiple regression analysis supported this explanation (see Table III).

What are the implications of our findings? Previous studies have reported how well the DT discriminates good from poor readers. It has been claimed that up to 65 per cent of genuine poor readers referred for orthoptic assessment fail the DT. This means that 35 per cent of the poor readers in these studies passed the DT (Bigelow and McKenzie 1985, Stein and Fowler 1985), and in our study some of the better readers failed it. These facts show that if the DT is used alone to diagnose reading problems it would produce a significant number of false positive diagnoses, just as some good readers do poorly in phonological tests (Bryant and Bradley 1985).

Nevertheless, those who fail the DT are more likely to make non-word errors when reading single words. Unstable binocular control therefore potentially interferes
with reading, even of otherwise good readers, particularly in the early stages of reading, and especially when children attempt to read new words using an alphabetic strategy. This could have important remedial implications, as some children who appear to experience difficulty using an alphabetic strategy may in fact have competent phonological skills. It is also possible that some children who fail the DT, but who nevertheless make reasonable progress in reading, have learned to compensate, for example by adopting an orthographic strategy more quickly.

In this paper we have only considered the non-word errors that children made when reading because we wanted to make testable predictions about how stressing the vision of visually impaired children would affect the way they read single real words. One clear prediction was that they would make more non-word errors as the print size of the words was reduced. However, it is also possible that children of different ages, who are reading contextual material and who have been taught different reading strategies, might deal in other ways with visually confused words. We know that dyslexic people can misread real target words as other real words, nonsense target words as real words and vice versa (Snowling 1980, 1981; Kochner et al. 1983), so we do not suggest that under normal conditions a child with unstable binocular control will only make non-word reading errors. If the child perceives an incorrect grapheme string, it is feasible that s/he could generate any of the error types currently described in the literature.

There are still many questions to be answered. We have identified an effect on reading which must have a visual basis. Now we need to find a way to identify more precisely the visual mechanism underlying the effect we have found, as the DT is difficult to administer and its results are inconsistent (Newman et al. 1985). Only then can we discover how prevalent the phenomenon is among young children, and its relationship to their progress in school.

In summary, some children with reading difficulty are able to give strikingly clear reports of what they see happening to letters and words on the page. The results of this study show that a group of mixed-ability children who failed the Dunlop Test made more non-word errors when their visual systems were stressed. This suggests that, even if they represent a minority of children with reading problems, what some children see may affect how they read.

Accepted for publication 5th March 1991.

Authors’ Appointments
*Piers Cornelissen, Physiology Department;
Lynette Bradley, Experimental Psychology Department;
University of Oxford, Parks Road, Oxford OX1 3PT.
Sue Fowler, Orthoptic Department, Royal Berkshire Hospital, Reading.
John Stein, Physiology Department, University of Oxford.

*Correspondence to first author.

SUMMARY
Children of mixed abilities were given three single-word lists to read, matched for linguistic complexity. The visual component of the task was made harder by reducing the print size with each new list. The reading errors made by children who did and who did not have a visual impairment were compared. The visually impaired children’s pattern of reading errors changed as their vision was stressed by the reduction in print size; their errors became non-words (neologisms). This finding suggests a link between the efficiency of visual processing and the accuracy of reading of these children.

RÉSUMÉ
Ce que les enfants voient modifie leur lecture
Des enfants de capacités variées reçurent des listes de trois mots isolés à lire, appariés pour la complexité linguistique. La composante visuelle de la tâche fut rendue plus difficile en réduisant la taille d’imprimerie avec chaque nouvelle liste. Les erreurs de lecture faites par les enfants que avaient ou n’avaient pas de troubles visuels ont été comparées. La distribution de erreurs de lecture chez les enfants à trouble de vision étaient modifiée lorsque leur vision était altérée par la réduction de la taille d’impression; leurs erreurs devinrent des non-mots (néologismes). Ces données suggèrent un lien entre l’efficacité de la fonction visuelle et la précision de la lecture chez ces enfants.

ZUSAMMENFASSUNG
Was Kinder sehen, bestimmt wie sie lesen
Kinder mit unterschiedlichen Fähigkeiten bekamen drei Listen mit einzelnen Wörtern zu lesen, geordnet nach sprachlichem Zusammenhang. Die visuelle Komponente der Aufgabe wurde dadurch erschwert, daß die Buchstaben mit jeder neuen Liste kleiner wurden. Die Lesefehler wurden bei den Kindern mit und ohne Sehbehinderung verglichen. Bei den visuell behinderten Kindern änderte sich mit
zunehmender Beanspruchung der Augen durch kleinere Buchstaben das Muster der Lesefehler; ihre Fehler waren Worte neubildungen. Diese Befunde zeigen, daß es eine Verbindung zwischen der visuellen Verarbeitung und der Lesegenauigkeit bei diesen Kindern gibt.

RESUMEN
Lo que el niño ve afecta a como lee
A niños con habilidades mixtas se les dió tres listas de monosílabos para ser leídas, según una complejidad linguistica. El componente visual de la tarea se aumentó en complejidad reduciendo el tamaño de la letra impresa en cada nueva lista. Se compararon los errores de lectura hechos por los niños que tenían o no alteración visual. El patrón de errores de lectura de los niños con alteración visual cambiaba al aumentar la dificultad visual con la reducción del tamaño de las letras; sus errores se convirtieron en neologismos (no palabras). Este hallazgo sugiere un enlace entre la eficiencia del procesado visual y la precisión de la lectura en estos niños.

References
Pavlidis, G. T. (1981) 'Do eye movements hold the key to dyslexia?' Neuropsychologia, 19, 57-64.