Journal of Research in Reading, ISSN 0141-0423
Volume 23, Issue 3, 2000, pp 314–324

Typography in children's reading schemes may be suboptimal: Evidence from measures of reading rate

Laura E. Hughes and Arnold J. Wilkins

Visual Perception Unit, Department of Psychology, University of Essex

ABSTRACT

We investigated whether the layout of type in two popular children's reading schemes was suitable for the intended reading age. 120 children read four passages of text that adopted the typography of four reading stages in each of the two schemes. The size and spacing of the texts decreased with successive stages as the intended reading age increased. The reading speed of children aged 5 to 7 years decreased as the text size decreased: in particular, these children read fastest the text designed for 5 and 6 year olds. Older children aged 8 to 11 years were neither assisted nor disadvantaged by text size. Children of all ages, particularly those susceptible to visual stress, were found to make more errors on the smaller than on the larger text. We conclude that the reading development of some children might benefit from a larger text size and spacing than is currently the norm, and that no children would be disadvantaged by such a change.

INTRODUCTION

Typographic characteristics such as the font, the size of the type, the spacing between lines (interlinear spacing), the spacing between words and letters, and the contrast of the print and paper are all thought to influence legibility (e.g. Tinker 1963, Watts and Nisbett 1974, Walker 1992, Wilkins 1995). The contribution of these characteristics to the development of reading skills is important but has received limited study.

Early studies of the typography in children's books suggested conditions for optimal legibility but few were based on empirical research, and opinions were varied and rarely concurred. Shaw (1902) was one of the first to study standards of legibility and he based his recommendations for type size in children's books on German research. He suggested an x-height of 2.6mm for Grade 1 children reducing to 1.6mm for Grade 5. Tinker (1959) suggested a small font size (18–14pts for Grade 1 and

12pts for Grade 5) and leading (4–6mm for Grade 1 and 2mm for Grade 5). However, Burt (1959) advocated much larger font sizes (24pts for up to age 7 decreasing to 11pts for over age 12) and a word spacing equal to the letter *m* at minimum, especially for the younger children. Buckingham (1931) used one typeface (Monotype 8) and printed three different stories using a different size of typeface, line lengths and leading. These stories were given to 2,010 seven-year-old children to read, from whom he took comprehension and speed measures. He found that a 12 point type with a 14½ pica line length and a three point leading was read more rapidly.

Typically children's reading books have a tendency to reduce in x-height and line spacing as the reading age increases. Small print sizes are thought to make reading increasingly difficult and are more stressful to the visual system (Wilkins and Nimmo-Smith, 1987). The primary source of information for reading is visual (Massaro and Cohen, 1994) and the rapid and accurate decoding of text in turn promotes phonological awareness and competent reading (e.g. Carr, 1981; Jorm and Share, 1983). Manis et al (1990) suggest that poor translation from orthography to phonology is indicative of poor reading. Rumelhart (1977) and Stanovich (1980) suggest that if perceptual analysis is inadequate, reading will be adversely affected. Reducing perceptual difficulty in decoding text should therefore help improve general reading ability. Erdman and Neal (1968) and Blommaert and Timmers (1987) both demonstrated that adults found it more difficult to recognise words and letters when the print size was reduced. Cornelissen, Bradley, Fowler and Stein (1991) report an increase in visuo-perceptual problems, in proportion to a decrease in spatial dimensions of text. A control group of children was compared to a group of children with impaired visual processing who were expected to perform more poorly when their visual system was stressed. The children, of mixed reading abilities, were given three lists of single words, matched for linguistic complexity, each a different font size: *large* (24pts Helvetica, quadruple line spacing and three character spaces between each word); *medium* (12pts Helvetica single line spacing and one character space between each word); and *small* (9pts Helvetica half line spacing and one character space between words). The results showed that more errors were made by all the children as the print size reduced, however those with visual deficits made significantly more errors on the medium than on the large sized text.

Individuals susceptible to visual stress may be expected to perform more poorly when reading closely spaced texts. Wilkins and Nimmo-Smith (1987) suggest that because text often resembles a pattern of stripes it can have aversive properties similar to patterns that induce illusions. Patterns of this kind are most aversive when they have a spatial frequency of about 1–3 cycles per degree and a duty cycle (ratio of bar width to separation) of 50%. As the separation of the lines of text increases, so the spatial frequency and the duty cycle of the 'stripes' both decrease, moving away from the stressful parameter range (Wilkins, 1995). The contrast of the text, and therefore the average contrast of the 'stripes' of text, is also an important factor: the stressful properties of stripes increase linearly with the logarithm of Michelson contrast. The contrast of the 'stripes' in text is reduced by the breaks between words and letters, but can still be high enough to be within the range of aversive patterns. In a study by Evans et al (1996), susceptibility to visual stress was tested using an appropriately stressful pattern of stripes, and to control for suggestibility a control pattern of wider stripes was also used. The stressful and control patterns were of similar size, space-averaged luminance and contrast but had very different spatial frequencies. Both patterns were presented to a group of experimental and control participants. The experimental group were children who were failing in reading and reported asthenopia or perceptual distortions. The experimental group reported significantly more illusions and anomalous visual effects with the stressful pattern than the placebo pattern and also reported significantly more illusions than the control group (children whose reading ability was at least average for their age). The study suggests that susceptibility to visual illusions in the aversive pattern may indicate a problem with reading that is related to perceptual factors.

A relationship between visual stress and reading is also supported by evidence from Conlon et al (1998). They gave participants a reading-like visual search task and found those who suffered symptoms of visual stress took longer to complete the search, particularly when the search was in a pattern that resembled stripes.

Wilkins et al (1996) designed a specific test to measure the visual aspects of reading independently of confounding factors, such as comprehension or linguistic complexity: known as the *Rate of Reading Test*. This is a simple test that minimises the linguistic and semantic aspects of reading and maximises visual difficulties. It comprises a list of 15 different words randomly ordered in a set of 10 lines, 15 words to a line. The text appears as a paragraph and lacks any meaning. The words used were selected from 110 most frequent words in children's reading books and the test can therefore be administered to both children and adults. For effective reading, accuracy and speed are both important contributors (Adams, 1990, Carr et al, 1990), but in the *Rate of Reading Test*, reading speed is the best single measure of performance because rate and accuracy are positively correlated: there is no speed/accuracy trade-off. The rate can be assessed by the number of words read correctly in one minute. The *Rate of Reading Test* differs from conventional reading tests that are designed to assess scholastic attainment in reading and compare one individual's performance with that of another. It is designed to compare each participant's performance under different visual conditions: reading speeds vary considerably from person to person. The spatial parameters of this test can be manipulated and the subsequent levels of visual stress can be measured under different visual conditions.

Wilkins et al (1996) analysed the reliability and the validity of the *Rate of Reading Test*. A significant effect of practice was reported, with the second presentation being performed 2.9% faster than the first, but the time taken to read three subsequent tests did not differ significantly from the second. The reliability on immediate retest was high, with correlations in excess of 0.9, reduced only slightly when the interval between tests was increased to 3 months.

The present study used versions of the *Rate of Reading Test* created in a typography that resembled that in children's reading schemes. The effects of the typography on speed and accuracy of reading were compared, with the aim of identifying whether the text was appropriately designed for the children who would be required to read it.

In order to choose two reading schemes that use typical typography, books from reading schemes dating from the 1940's to 1995 were collected from primary schools in South Yorkshire and Cambridge. At least two examples of different schemes, each from a different publisher were selected for each decade. The intended reading age of each book was listed in the NASEN *A–Z Graded list of Reading Books*, which

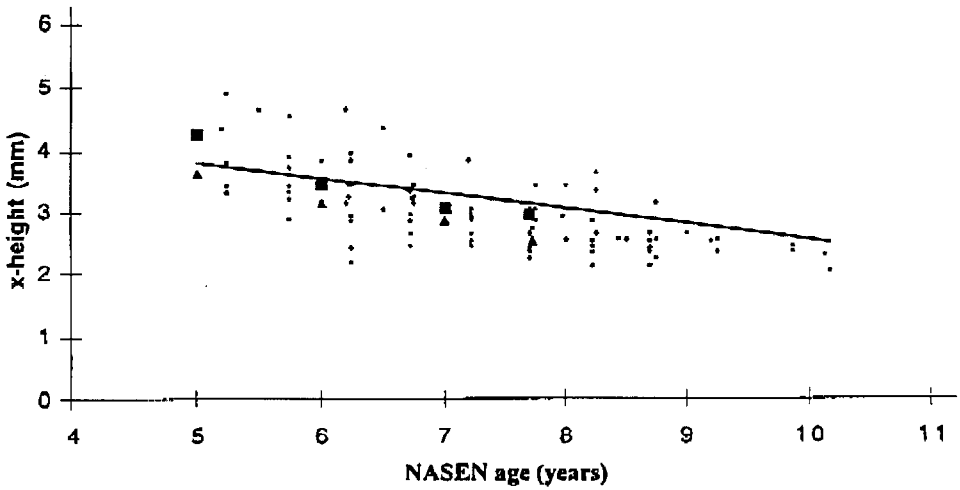

Figure 1. x-height (mm) as a function of NASEN age (years) for a sample of childrens' readers published between 1950 and 1995 (small points). The points for the *All Aboard* and *Oxford Reading Tree* schemes are shown by squares and triangles respectively.

provides a measure of readability calculated using a version of the Spache Readability Formula (Harrison, 1980) and the Fry Readability Graph (Fry, 1977). The x-heights of text from each reading book were measured, and the results are shown in Figure 1. From this figure it is evident that the x-height decreases as the intended reading age increases. The graph also shows, as segregated points, the values of x-height used in the two reading schemes selected for study: *Oxford Reading Tree* (Oxford University Press, 1986–1988) and *All Aboard* (Ginn and Co., 1994–1996). As can readily be seen, the selected reading schemes use x-heights that are typical and close to the mean which is shown by the solid line.

METHOD

Design

Rate of Reading Tests were used to measure reading fluency and were based on two popular reading schemes, *All Aboard* and *Oxford Reading Tree*. A representative sample of four books, covering reading ages 5 to 8 were selected from each reading scheme and used to create the reading tests. Children were randomly assigned to either of two groups. Both groups were first given an example *Rate of Reading Test*, used to familiarise the children with the task and reduce the effect of practice, and then were tested with four reading tests (one for each of four reading stages) with typography based on either of the two reading schemes. The reading tests were presented in a typical classroom setting for the child to read as they felt it comfortable to do so. Reading distance was determined by the child. In each group half the children were given four reading tests in the order of largest type to smallest and half in the reverse order. The children were tested for a period of 45 seconds rather than a minute and this was intended to remove severe tiredness effects anticipated in the younger children. A measure of susceptibility to visual illusions was also taken, using a visually stressful pattern viewed at a distance of approximately 40cm. A non-stressful pattern was also used to control for suggestibility.

Selection of participants

120 children were selected from years 1 to 5 in a mainstream primary school in Hove, UK. One child with epilepsy was excluded, together with three children for whom parental permission was not forthcoming.

Materials

From each reading scheme the following stages were included: *All Aboard* stages 2, 4, 6, 8 and *Oxford Reading Tree's* Sparrows, Owls, Robins and Jackdaws. From each stage the x-height, line spacing, word spacing and line length was measured. These measurements were used to create versions of the Rate of Reading Test with similar spatial parameters. Figure 2 a-d shows excerpts of each Reading Test.

A visually stressful pattern and control pattern were also created. The stressful pattern included some of the maximally stressful characteristics as set out by Wilkins

(a)

cat for the yo
dog see to is

(b)

cat for the you n
is come play loc

(c)

cat for the you no
play look and see

(d)

cat for the you not
play look and see y
mv the un and com

Figure 2. Excerpts of the *Rate of Reading Test* created using the letter size and spacing from the *All Aboard* reading scheme, as follows: (a) Test 1, (b) Test 2, (c) Test 3, (d) Test 4.

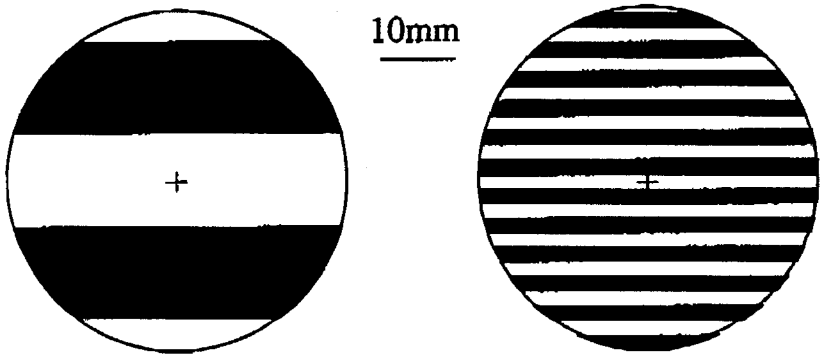

Figure 3. Visually stressful test patterns used in the study, showing a scale.

(1995). It had a square-wave luminance profile, a contrast of 80%, a duty cycle of 50% and spatial frequency of 2.5 cycles per cm (3 cycles per degree at a viewing distance of 50cm). It had a very small radius (2.5cm) in order to avoid adverse somatic effects. The control pattern differed in spatial frequency, having a spatial frequency of 0.28 cycles per cm. Figure 3 shows examples of the patterns used.

Procedure

Children were taken individually into a quiet room and presented first with an example *Rate of Reading Test* which they were asked to read as quickly as possible. They were then presented sequentially the four *Rate of Reading Tests* based on either the *All Aboard* stages or the *Oxford Reading Tree* stages. They were asked to read these four short passages as quickly as they could; the number of words read in 45 seconds was noted and any errors made. Errors included omissions and commissions of words or lines. Following the *Rate of Reading Tests*, the children were presented with two patterns, the control pattern first and then the experimental pattern; each was presented for about 5 seconds. The children were asked whether the lines appeared to move, bend or blur and which colours they saw, if any.

Before participating in the study, all the children had undergone a test of reading attainment as part of the school curriculum. The tests used were the *Schonell Reading Test*, in the case of the children in years 1 to3, and the *Salford Reading Test* in the case of the children in years 4 and 5.

RESULTS

Presentation effects

Each child received the four *Rate of Reading Tests*. One group received the test in ascending order of text size and a second group in descending order of size. The effect of test order was assessed by an analysis of variance. No significant effect was obtained, suggesting that the observed differences in performance could be attributed to the text size rather than the order in which the tests were presented.

Reading speed versus accuracy

There was a negative correlation between reading speed and errors. (The errors were calculated as a percentage of the number of words read). Pearson's correlation coefficient $r = -0.167$, *All Aboard*; $r = -0.223$, *Oxford Reading Tree*. As reading speed increased the number of errors decreased indicating that reading speed was a suitable single measure of reading performance.

Analyses of reading speed

Figures 4a and 4b display the reading speeds over the four Reading Tests for each age group.

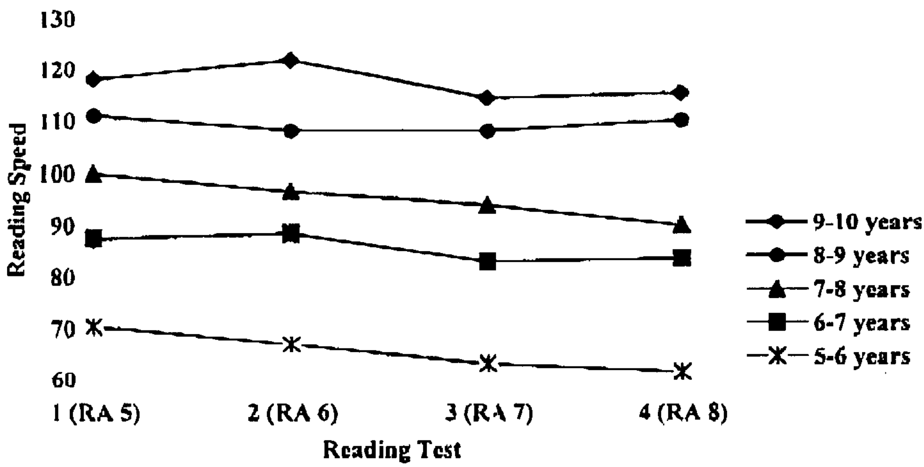

Figure 4a. Mean words per minute on each version of the *Rate of Reading Test* (1–4), reproduced using the typography of the *All Aboard* stages, as appropriate for the reading age (RA) shown.

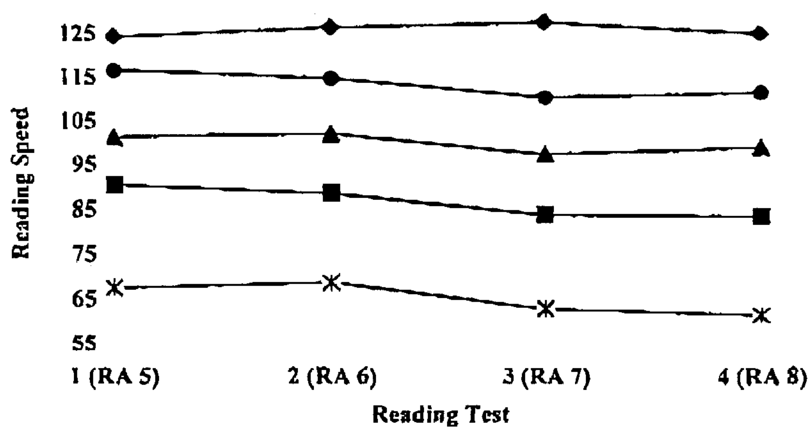

Figure 4b. Mean words per minute on each version of the *Rate of Reading Test* (1–4), reproduced using the typography of the *Oxford Reading Tree* stages, as appropriate for the reading age (RA) shown (legend as Figure 4a).

The reading speeds of the children aged 5 to 8 decreased over the four reading tests as size decreased. Contrast analysis was used to identify a linear trend (the size of the main effect attributable to linearity is represented by Page's L). In order to increase the power of the statistical test the children were grouped within each scheme into the age groups: 5–7, and 8–11. A significant trend was found in both schemes for the younger age groups (5–7) [*All Aboard*: L = −598, F (1,81) = 6.87, $p<0.05$; *Oxford Reading Tree*: L = −460, t (81) = 1.85, $p<0.05$] indicating that as text size decreased so reading speed also decreased. No such trend was observed for either of the older age groups.

There were no significant differences for individuals with above or below average reading age between reading rates on the four tests.

Analyses of errors

In the analysis of the errors the children were grouped by ages as in the analysis of reading speed (ages 5–7 and 8–11). Table 1 shows the percentage of errors made on each of the reading tests for each age group (the percentage was calculated as the number of errors made per number of words read).

The errors made by the children on each reading test were also analysed. Significant differences were observed between the number of errors made on the small text (Test 4), and on the large text (Test 1) in all groups [*All Aboard* age 5–7, $p<0.0004$, age 8–11 $p<0.01$, *Oxford Reading Tree* age 5–7, $p<0.0001$, 8–11 $p<0.016$]. A significant linear trend was also identified for each group [*All Aboard* age 5–7: L = 333.47, F(1,81) = 41.7; age 8–11: L = 133.2, F(1,93) = 10.4; *Oxford Reading Tree* age 5–7: L = 307.97, F(1,81) = 30.66; age 8-11 L = 115.4, F(1,93) = 10.88]. This indicates that all children made more errors on the smaller texts (Tests 3 and 4).

Analyses of susceptibility to visual stress

Further analyses were carried out to identify whether certain groups of children were differentially affected by the smaller text, particularly those susceptible to visual stress, as measured by the response to the visual stress pattern. The errors demonstrated the greatest difference between the children reporting illusions in the visual stress pattern and those reporting no illusions. In the age group 5–7 both the illusion and no-illusion groups made significantly more errors on Tests 3 and 4 than on

Table 1. Mean number of errors made by children in the two age groups on each reading test.

	N	Test 1	Test 2	Test 3	Test 4
All Aboard					
Ages 5–7	28	1.99	2.85	3.31	5.8
Ages 8-11	32	2.63	3.25	3.65	3.88
Oxford Reading Tree					
Ages 5-7	28	2.03	3.03	3.83	5.43
Ages 8-11	32	2.95	2.55	2.71	4.1

Tests 1 and 2 [*All Aboard*: illusion $t(11) = 2.502$, $p = 0.015$ (1 tail); no illusion $t(14) = 2.171$, $p = 0.023$ (1 tail) *Oxford Reading Tree*: illusion $t(14) = 3.76$, $p = 0.001$ (1 tail); no illusion $t(12) = 1.919$, $p = 0.001$ (1 tail)].

In the age group 8–11 only the illusion groups made significantly more errors on Tests 3 and 4 than on Tests 1 and 2. Thus it appears that younger children and those susceptible to visual stress are more likely to make errors when reading smaller text [*All Aboard*: illusion $t(22) = 1.799$, $p = 0.043$ (1 tail); *Oxford Reading Tree*: illusion $t(16) = 2.175$, $p = 0.022$ (1 tail)].

DISCUSSION

A significant decrease in reading speed was observed over the four typographic styles for the children aged 5 to 7, whereas children aged 8 to 11 showed no significant effect of text parameters on their reading speed. The analysis of the errors revealed that all children in both schemes and in both age groups made more errors on the smaller texts (Tests 3 and 4) than on the larger (Tests 1 and 2). A negative correlation was observed between reading speed and errors which suggests that there was no trade-off between speed and accuracy and that a slow reading speed indicates an increase in mistakes.

Further analyses of visual stress and reading demonstrated that text can be visually stressful and thereby affect reading ability. Children who were susceptible to visual stress performed significantly more poorly when asked to read the smaller texts (Tests 3 and 4), which were more visually stressful than the larger texts.

These results suggest that reading speed and accuracy could be increased by presenting children with a text having a larger more widely spaced typeface. It might be expected that children should read text that is designed for their age group faster and more accurately than other text. However, as Figures 4a and 4b show, this did not always occur. The children in the 5–7 age group performed better on the larger text, designed for those aged 5–6, than the smaller text designed for older children. The children in the age group 8–11 also performed better on the larger text than the text designed for them, although the difference was not significant.

Cornelissen et al (1991) reported an increase in the number of errors made by children as text size decreases. Contrary to expectation, they failed to find that children with impaired visual processing were more adversely affected by small text than those without apparent problems. Most of the children in the study by Cornelissen et al were affected, suggesting that rather than deficient visual processing, a decrease in the legibility of the reading material was the cause of an increase in errors.

In reading schemes, the child has to contend not only with an increase in the complexity of the linguistic and semantic content but also with an increase in visuo-perceptual complexity. It might be helpful if the child were able to concentrate on developing comprehension and phonological skills rather than perceptual deciphering skills. In this study there was no increase in comprehension complexity over the four tests, yet the younger children showed a significant decrement in performance. Taxing the child's reading skills further by adding an increase in semantic and linguistic complexity might be expected to further decrease performance.

In Reading Test 1 of both reading schemes the typography was designed for material with one or two lines at most on a page. When the typeface and style were used for a paragraph of text, as in the *Rate of Reading Test*, the interword spacing was large relative to the interline spacing. Nevertheless, the children read this text *more quickly* than the text which had been designed to be set as a paragraph.

In both schemes the spatial frequency of the successive lines of text increases towards the range at which illusions are most frequently seen (Wilkins 1995). Reading Test 4 for both schemes is within the stressful range. This may explain the greater number of errors made by the children who were susceptible to visual stress, and also suggests parameters which should be avoided when designing text layout.

The results indicate that the present x-height and spacing of text designed for children with reading ages of 7 and 8 years may be sub-optimal. The range of x-heights used in the present study was limited, and, as can be seen from Figure 1, further reduction in x-height is common in material used by older children. Although overall the 8–11 year olds did not show a significant decrease in reading speed with x-height, this is unlikely to hold in books designed for them, where x-height is further reduced. These books are likely to have a text layout which resembles a pattern of stripes with visually stressful properties. Those children susceptible to visual stress might start to struggle when their reading age develops.

Perhaps children would benefit from the postponement of the size reduction in reading schemes until a later reading age, keeping text large while reading skills develop. The greatest benefit would be for the younger children aged 5–7 who have difficulty with the visual processing of smaller text, as demonstrated by their slower reading speeds and lower level of accuracy on this text. The children who are susceptible to visual stress would also benefit; these children made more errors on the smaller than on the larger text, regardless of age. The results suggest that keeping text size and spacing large while children learn to read will be beneficial to some children and detrimental to none.

Acknowledgements

The authors wish to thank Philippa Greenfield and Andrea Cove-Smith for their permission to reproduce the data in Figure 1, which they collected.

REFERENCES

ADAMS M.J. (1990) *Beginning to read: Thinking and Learning about Print*. Boston, MA: Massachusetts Institute of Technology.

BLOMMAERT, F.J. and TIMMERS, H. (1987) Letter recognition at low contrast levels: effect of letter size. *Perception*, 16, 421-432.

BUCKINGHAM, B.R. (1931) *New data on the Typography of Text Books*. 30th Year Book. Part II: The National Society For Study Of Education. Chicago: University of Chicago Press, pp. 93–125.

BURT, C. (1959) *A Psychological study of Typography*. London: Cambridge University Press.

CARR, T.H. (1981) Building theories of reading ability: On the relation between individual differences in cognitive skills and reading comprehension. *Cognition*, 9, 73–114.

CARR, T.H., BROWN, T.L., VAVRUS, L.G. and EVANS, M.A. (1990) Cognitive skill maps and cognitive skill profiles: Componential analysis of individual differences in children's reading efficiency. In T.H. Carr and B.A. Levy (Eds.) *Reading and its development: Component skills approaches*. New York: Academic Press.

CONLON, E., LOVEGROVE, W., HINE, T., CHEKALUK, E., PIATEK, K. and HAYES-WILLIAMS, K. (1998). The effects of visual discomfort and pattern structure on visual search. *Perception* 27, 21-33.

CORNELISSEN, P., BRADLEY, L., FOWLER, S. and STEIN, J. (1991) What children see affects how they read. *Developmental Medicine and Child Neurology*, 33(9), 755-762.

ERDMAN R.L. and NEAL A.S. (1968) Word legibility as a function of letter legibility, with word size, word familiarity and resolution as parameters. *Journal of Applied Psychology*, 52, 403-409.

EVANS, B.J.W., WILKINS, A.J., BROWN, J., BUSBY, A., WINGFIELD, A., JEANES, R., and BALD, J. (1996) A preliminary investigation into the aetiology of Meares-Irlen syndrome. *Opthamic Physiological Optics*, 16(4), 286-296.

FRY, E. (1977) *Elementary Reading Instruction*. New York: McGraw Hill.

HARRISON C. (1980) Readability in the Classroom. London: Cambridge University Press: Cambridge.

JORM, A. and SHARE, D. (1983) Phonological recoding and reading acquisition. *Applied Psycholinguistics*, 4, 103.

MANIS, F.R., SZESZULSKI, P.A., HOLT, L.K. and GRAVES, K. (1990) Group differences in phonological decoding. In T.H. Carr and B.A. Levy (Eds.) *Reading and its development: Component skills approaches*. New York: Academic Press.

MASARO, D. and COHEN, M. (1994) Visual, orthographic, phonological and lexical influences in reading. *Journal of Experimental Psychology: Human Perception and Performance*, 20, 1107-1129.

RUMBLHART, D.E. (1977) Toward and interactive model of reading. In S. Dornic and P. Rabbitt (Eds.) *Attention and Performance IV*. Hillsdale, NJ: Erlbaum.

SHAW, E.R. (1902) *School Hygiene*. London: Macmillan.

STANOVICH, K.E. (1980) Toward and interactive-compensatory model of individual differences in the development of reading fluency. *Reading Research Quarterly*, 16, 32-71.

TINKER, M.A. (1959) Print for children's textbooks, *Education* (USA), 80(1), 37-40.

TINKER, M.A (1963) *Legibility of Print*. Ames, IA: Iowa State University Press.

WALKER, S. (1992) *How it looks: A teachers guide to typography in children's reading books*. Reading:, Berks: The Reading and Language Information Centre of the University of Reading.

WATTS, L. and NISBETT, J. (1974) *Legibility in children's books: A review of Research*. Windsor, Berks: NFER Publishing Co.

WILKINS A.J. (1995) *Visual Stress*. London: Oxford University Press.

WILKINS, A.J., JEANES, R.J., PUMFREY, P.D. and LASKIER, M. (1996) Rate of reading test: Its reliability, and its validity in the assessment of the effects of coloured overlays. *Ophthalmic and Physiological Optics*, 16(6), 491-497.

WILKINS, A.J. and NIMMO-SMITH, M.I. (1987) The clarity and comfort of printed text. *Ergonomics*, 30(12), 1705-1720

Address for correspondence: PROFESSOR ARNOLD WILKINS, Visual Perception Unit, Department of Psychology, University of Essex, Colchester CO4 3SQ, UK. Email: arnold@essex.ac.uk