



ELSEVIER

Contents lists available at ScienceDirect

Vision Research

journal homepage: [www.elsevier.com/locate/visres](http://www.elsevier.com/locate/visres)

# A comparison of the effects of the colour and size of coloured overlays on young children's reading

Judit Veszeli, Alex J. Shepherd\*

Department of Psychological Sciences, Birkbeck, University of London, London, UK

## ARTICLE INFO

No. of reviewers: 1

### Keywords:

Visual stress  
Reading  
Visual clarity  
Coloured overlays  
Coloured reading rulers  
Primary school children

## ABSTRACT

This study compared the effects of the colour and size of overlays on reading time, reading errors and on the clarity of text with young primary school children. The sample comprised a non-clinical, typical, sample from an East London primary school. One hundred and six children aged between four and seven years were asked to read 11 short passages of text (60 words) either with full page overlays or smaller reading rulers (53 in each group). This sample included younger children than has often been tested before. The 11 short passages allowed an assessment of baseline reading performance (no reading aid) and performance while reading with each of a set of ten coloured reading aids. Two different, yet beneficial, colours were determined: the most effective and the clearest/most comfortable. Both of these measures are not usually recorded. All but four children had reduced reading times with one of the reading aids and all but one reported their aid improved the perceived visual clarity of the text: the size of the reading aid did not affect reading time or visual clarity significantly. The numbers of skipped words and errors/mis-read words also decreased when reading with the most effective and most comfortable reading aid. Near visual acuity was assessed with and without each child's most effective coloured aid. The most effective aid improved acuity in over a third of the children. Acuity has not been assessed in previous studies. As reported previously, different colours helped different children. In conclusion, coloured reading overlays reduced reading times on the reading test employed here and the size of the reading aid was not crucial to facilitate performance. The largest reductions occurred for the youngest readers, suggesting these aids may be particularly effective for early readers.

## 1. Introduction

Reading is one of the most important skills for young children to acquire. It has an impact on many other basic skills in the classroom, in the general learning process and in other daily activities. Reading ability can affect counting, writing, comprehending and phrasing. Reading difficulties are, however, common amongst primary school children. Those who struggle from the beginning often do not catch up with the curriculum later. When pupils leave primary school they are expected, as a minimum, to achieve a level 4 in reading (see *Participants* section, below, for a description of this level) at the SAT's testing in Year 6 (ages 10–11). Data from the *National Curriculum Assessments in England (2011/ (2011/ (2011/12, revised)*, however, has shown that 13% of pupils performed under this level. It is important, therefore, to examine claims of easily applicable methods that can help young children to improve their reading.

Since the 1980's, there have been many reports that tinted plastic sheets lain over text ("overlays"), smaller "reading rulers", which only

partially cover a page of text, or tinted lenses worn in spectacles, can improve the perceived visual clarity of printed text for some individuals and help with their reading (for reviews, see *Solan & Richman, 1990; Wilkins, 2002, 2003; Smith & Wilkins, 2007*). One study reported that 5% of children in mainstream schools, aged between 6 years 10 months and 8 years 6 months, read at least 25% more quickly with their preferred coloured overlays (*Wilkins, Lewis, Smith, & Rowland, 2001*). This group has also reported that up to 20% of children are likely to experience some benefit from overlays (reviewed in *Wilkins, 2003*). The efficacy of coloured overlays or rulers is controversial, however, with some fully endorsing them, particularly for children with reading difficulties, dyslexia or visually induced discomfort (reviewed in *Wilkins, 1995, 2003*), and others questioning whether beneficial effects are specific to colour or represent simply a placebo effect (see *College, 2009; Ritchie, Della Sala, & McIntosh, 2011*).

In this study, the reading performance of primary school children, aged between 4 and 7 years, was compared when they read with and without coloured overlays or the smaller reading rulers (*Fig. 1*). The

\* Corresponding author at: Department of Psychological Sciences, Birkbeck, University of London, Malet Street, London, WC1E 7HX, UK.  
E-mail address: [a.shepherd@bbk.ac.uk](mailto:a.shepherd@bbk.ac.uk) (A.J. Shepherd).

<https://doi.org/10.1016/j.visres.2019.01.006>

Received 6 June 2018; Received in revised form 24 December 2018; Accepted 8 January 2019

Available online 04 February 2019

0042-6989/ © 2019 Elsevier Ltd. All rights reserved.

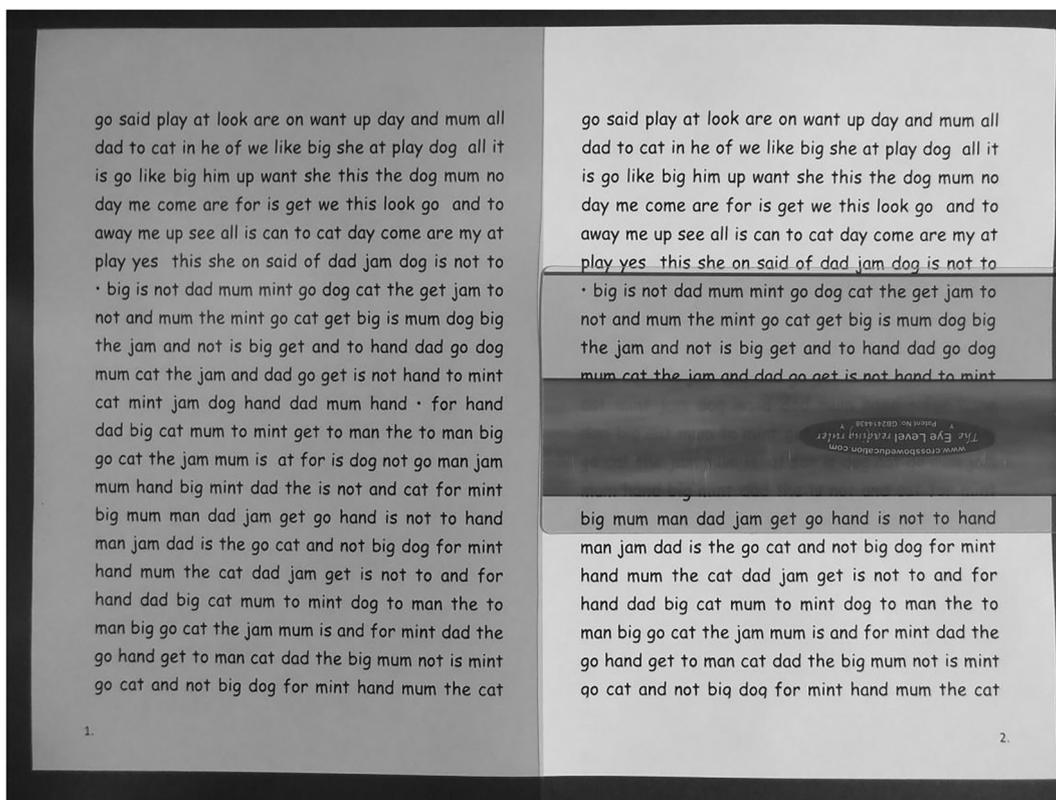


Fig. 1. A greyscale photograph of the overlays (left hand page) and reading rulers (right hand page) supplied by Crossbow Education Ltd. The reading aids were not grey, they had the chromaticities depicted in Fig. 2.

children in this study had not been identified as having reading difficulties or dyslexia, so the literature on the use of coloured reading aids for those conditions, or on the possible causes of such reading difficulties, is not reviewed here (see Albon, Adi, & Hyde, 2008 for a review). The children tested in the present study were simply a sample from a mainstream primary school in East London. The focus of the study was on possible differential effects of the size of the reading aid, and which colours were selected. Younger children than are commonly assessed were included to assess whether these reading aids can improve the acquisition of reading in those who have just begun to learn to read short, high frequency, words fluently.

Some authors have assessed reading performance for single words and words in sentences, in which case the context in which a word appears, or the meaning of the sentence or of the whole text passage, can influence single word identification, reading speed and comprehension (e.g. Tyrell, Holland, Dennis, & Wilkins, 1995; Wilkins, 2003; Henderson, Tsogka, & Snowling, 2012). For this reason, it can be argued that these tests confound comprehension of context with visual clarity of text and that this may obscure the effects of interventions such as overlays (Wilkins, Jeanes, Pumfrey, & Laskier, 1996; Wilkins, 2003). The Rate of Reading Test (RRT) is an alternative assessment that eliminates context and focuses on reading speed and errors. It presents short, high frequency, words in a random order, so the passages cannot make sense. Each word thereby has to be read, it cannot be guessed from the context in which it is set, and this should enable the effects of different reading aids to be more readily apparent (1996, 2003; Wilkins et al., 1996; Wilkins, 1994). This test has been used extensively in prior research (overviewed in Wilkins, 2003).

This study used a modified version of the RRT to compare reading performance with coloured overlays and matching coloured reading rulers. While performance on this test may not generalize to normal reading, the perceptual ease the reading aids have been reported to afford might generalize to enhance normal reading. It is important to

remove context when comparing reading across different ages, however, and the RRT has been used widely for similar reasons: to look at perceptual clarity without the bias of contextual flow, particularly when testing early readers with different reading abilities and less habituated reading styles. For these reasons, the literature on word frequency, word identification and eye movements during natural sentence reading, is not reviewed here (see Rayner, 1998; 2009; Hyatt, Stephenson, & Carter, 2009 for reviews).

Once context is removed, there are still several factors that can affect individual word recognition. Familiarity, frequency, repetition, spelling-sound regularity, and the age of acquisition of the words, all can help the reader's speed and accuracy of word recognition (Ellis, 1993). Typographical features such as the font, text justification, character-, word- and line-spacing can also have a profound effect on reading performance (Hughes & Wilkins, 2000; Wilkins, 2003). These factors, with the exception of age of acquisition, were considered in the construction of stimuli used in this study.

Printed text is an aggregation of letters in horizontal lines, and the typically high contrast letters on the page, and repeated patterns of letter formations, may create anomalous visual effects, such as illusory colours, shimmering and vertical motions throughout the body of the text. These illusions can be accompanied by somatic reactions, such as nausea, eyestrain, headache and tiredness (2003; Conlon, Lovegrove, Chekaluk, & Pattison, 1999; Harle, Shepherd, & Evans, 2006; Shepherd, Beaumont, & Hine, 2013; Wilkins, 1995). Readers who are affected by these discomforting symptoms may complain about the “blinding whiteness” of the page and the experience of “moving”, “tilting” and “blurring” letters and words. These unpleasant experiences do not indicate problems such as refractive or binocular impairment, but may indicate visual stress, also referred to as visual discomfort or pattern glare (Evans, Patel, & Wilkins, 2002; Shepherd et al., 2013; Wilkins et al., 1984; Wilkins, 2003). Indeed, visual stress can occur with normal acuity and normal scores on other standard optometric tests (Evans

et al., 2002). It can also occur in individuals whose reading level is not impaired, however, so reading performance and visual stress can be dissociated (Evans, Allen, & Wilkins, 2009; Wilkins, 2003). Visual stress can, nevertheless, result in restless behaviour patterns, lack of attention, aversion to reading, and consequential underachievement.

Coloured overlays and lenses have been reported to reduce perceptual distortions and visual stress experienced from text, alleviate fatigue, improve the clarity and comfort of text and so increase the speed of reading (Evans et al., 2009; Wilkins, 2003). In this study, in addition to recording reading times, each child was asked about the perceptual clarity and comfort of a text passage when viewed using one of the two differently sized coloured reading overlays.

A few previous studies have compared the effects of overlay size on reading performance in children, but not with the materials, task and age range used here. Waldie and Wilkins (2003), for example, tested children aged between 8 and 12 years with A4 sized *Intuitive Overlays* (Wilkins, 1994), and trimmed A4 *Intuitive Overlays* that covered just a text passage and not the margins of the paper (80 mm × 92 mm). They used only one coloured aid per child, which was determined as the one giving the text the greatest clarity and comfort during an assessment in the previous school year. They found that those children who had continued to use their overlay read more quickly with either their A4 overlay or the smaller trimmed version, compared to the no reading aid condition. The children who had stopped using their overlays did not, however, read more quickly with either reading aid. They concluded that, when aids are beneficial, the size of the reading aid may not be crucial and smaller aids may be useful with smaller book sizes. This was not a direct comparison of overlays and the smaller reading rulers as used here, however, and they recommended further research was needed.

Smith and Wilkins (2007) compared the performance of 7–8 year olds when they were asked to read with (i) A5 *Intuitive Overlays*, available in nine shades (ii) a smaller cut down *Intuitive Overlay* in the shape of a ruler, or (iii) reading rulers from *Crossbow Education Ltd*. They reported reading speed improvements with the A5 *Intuitive Overlays*, and with the *Intuitive Overlays* cut down to the reading ruler size, but little improvement with the *Crossbow Education* reading rulers. They attributed the difference to the many more colours that are available with the *Intuitive Overlays* (27 in total, if the nine coloured overlays are used in pairs), so the optimum colour for an individual is more likely to be found, compared to the five colours then available for the *Crossbow Education* reading rulers.

Now ten colours are available as overlays, and as reading rulers, from *Crossbow Education Ltd*. This study, therefore, aimed to make an up to date comparison between the two sizes of reading aid, to see whether the colour and/or size of the reading aids from a single manufacturer affect the children's reading performance. As recommended in the procedure for the RRT (Wilkins, 2003), the children were asked about the clarity and comfort of the text with each reading aid (either ten overlays or ten reading rulers) and reading performance was also measured. The procedure differed, however, in one important respect: reading performance was assessed with all of the available colours, and not just the one judged by each child to give the text the greatest perceptual clarity and comfort. It would, thereby, be possible to assess whether the coloured aid that gave the text the greatest perceptual clarity, or the one that made it most comfortable, yielded the best reading performance. This has not been assessed before. Binocular visual acuity was also assessed with and without each child's optimum coloured aid. Finally, younger children were included (four year olds): those who have just begun to read words rather than letters, compared to the mostly older children involved in the studies cited above. Younger children were included to determine whether the coloured overlays may help early readers in a similar way to the effect reported for older readers. Older children who struggle with reading may have had difficulties earlier, when they transition from reading single objects (letters) to reading words (multiple letters). Visual blur or glare, rather

than cognitive or learning issues, may impact this transition, hence, the younger group were included to test this issue. Only those children, who were tested at the age of four and were able to read simple words correctly were included. Children aged less than four were not included to avoid floor effects. The government in the UK is currently pushing improving the reading of children aged five, as those who struggle at age five are reported to be six times less likely to reach expected standards at age 11 than others (see <https://www.gov.uk/government/news/boost-for-school-standards-with-primary-literacy-drive>, 2018).

## 2. Method

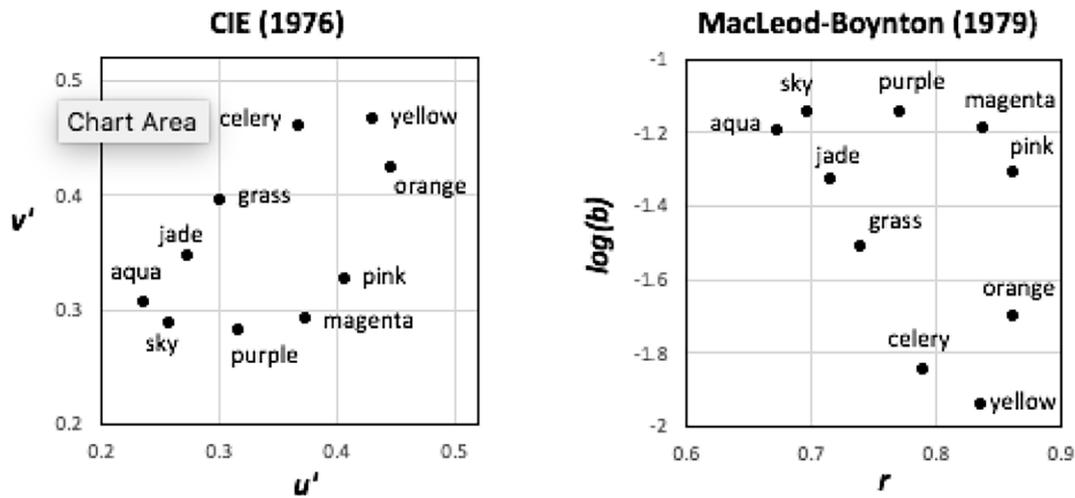
### 2.1. Stimuli

The ten coloured reading aids from *Crossbow Education Ltd* consist of A4 plastic sheets, or smaller plastic strips in the shape of a ruler, that are placed over text (Fig. 1). The overlays measure 210 × 297 mm and the reading rulers measure 206 × 74 mm. The A4 overlays were cut down into A5 sheets, measuring 148 × 210 mm, to fit the text passages prepared for testing. Both overlays and rulers are glossy on one side, and matte on the other, and were positioned with the matte side uppermost to avoid direct specular reflections from the overhead lighting. Their chromaticities are plotted in Fig. 2 in the CIE (1976)  $u',v'$  chromaticity diagram, for consistency with reported chromaticities in many earlier studies (for example, Wilkins, 1994; 1995, 1996, 2003, Wilkins et al., 1994, 1996, 2001), and, for comparison, in the MacLeod and Boynton (1979)  $r, \log_{10}(b)$  chromaticity diagrams<sup>1</sup>. The overlays and rulers are differentiated by the following names: purple, magenta, pink, orange, yellow, celery, grass, jade, aqua, sky-blue (hereafter "sky").

The reading stimuli were based on the RRT developed by Wilkins and colleagues (Wilkins et al., 1996; Wilkins, 2003). Fifteen high frequency words were selected and repeated four times: big, is, not, dad, mum, mint, go, dog, cat, the, get, jam, to, and, hand. The resulting 60 words comprised the test passages and were presented over five lines in a pseudorandom order. The constraints on word order were (i) that the same word was not immediately repeated; (ii) that each of the 11 reading passages had the same number of words (13) and approximately the same number of characters (37–38) in each complete line (iii) that the words in each line on each of the 11 test passages were the same and were just presented in a different order. With these constraints, each line and text passage presented the same level of difficulty (Fig. 1). Eleven different versions of the test passages were created, so that the children read a different sequence with each colour of overlay or ruler to avoid learning effects. It is important to emphasize that without a contextual flow, the children were not able to determine any pattern or predict the occurrence of the next word, therefore, the children had to focus on word to word reading, without making predictions or guesses.

The font, words and task were not identical to those used in the RRT: they were adjusted to better represent what the children were practiced and familiar with in their primary school. Thus, the words were printed in a font the children were used to (Comic Sans MS, font size 14). The line spacing was set at 1.15 lines to ensure that the youngest children would also be able to read them. Some of the words in the RRT, such as "see", "look", "play" were omitted as they contain digraphs that were unfamiliar to the poorest readers, considering the reading skills typical for their primary school and the age range of between four and seven years. Consequently, the words selected had

<sup>1</sup> For comparison with the CIE (1976)  $u',v'$  uniform colour space, the ordinate of the MacLeod and Boynton (1979) diagram has been  $\log_{10}$  transformed following Shepherd (1997, 1999): the scaling of the ordinate in the original MacLeod and Boynton (1979) diagram is reported as arbitrary. Shepherd (1997, 1999) reported colours around neutral are approximately equally perceptually spaced in this diagram following this transformation.



**Fig. 2.** The chromaticities of each of the reading colours plotted in the CIE (1976) and MacLeod and Boynton (1979) chromaticity diagrams. The CIE (1976) diagram is an approximately uniform colour space where equal distances between pairs of points should appear approximately equally different. For comparison, the MacLeod-Boynton  $r, \log_{10}(b)$  diagram is also an approximately uniform colour space for colours around neutral (after Shepherd 1997, 1999). These chromaticities should be taken as a guide only as there are slight differences in tint with different printed batches of overlay or ruler. The lightness of the overlays varied. The rank order from lightest to darkest was: yellow, celery, orange, grass, pink. These were recorded by measuring the luminance of each coloured aid lain on a white matte tile and lit with a white, equal energy, light source. Values for the remaining coloured aids were comparable to pink. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

**Table 1**

Participant details for each group and each reading level. N = number, M = male, F = female, RT = reading time in the baseline (no reading aid) condition. Age is given as the average in years (yrs)  $\pm$  1 standard deviation (SD). Numbers in parentheses indicate the age range for each reading level. Errors were mis-read words.

Reading level	N	Age $\pm$ 1 SD (yrs)	M	F	Average RT $\pm$ 1 SD (secs)	Number of children who skipped N words			Number of children who made N errors						
						N =	1	2	3	N =	1	2	3	4	5
<b>OVERLAY GROUP</b>															
1/2*	14	5.3 $\pm$ 0.7 (4.1–7.0)	10	4	51.7 $\pm$ 8.9	1	–	–	–	5	3	1	–	–	–
3	12	5.3 $\pm$ 0.8 (4.1–6.4)	4	8	46.8 $\pm$ 8.3	2	3	–	–	2	1	–	1	–	–
4	9	5.9 $\pm$ 0.6 (5.3–6.3)	8	1	49.9 $\pm$ 10.8	2	–	1	–	3	–	–	–	–	–
5	11	6.5 $\pm$ 0.6 (5.9–7.3)	4	7	41.1 $\pm$ 11.8	4	–	–	–	3	–	–	–	–	–
6/7**	7	6.6 $\pm$ 0.3 (6.1–7.0)	4	3	36.0 $\pm$ 5.5	1	2	–	–	4	–	–	–	–	–
<b>READING RULER GROUP</b>															
1/2*	15	5.7 $\pm$ 0.6 (4.7–6.7)	7	8	50.4 $\pm$ 8.6	–	–	–	–	6	2	–	–	–	–
3	13	5.8 $\pm$ 0.4 (5.3–6.7)	7	6	48.8 $\pm$ 15.2	2	–	–	–	3	1	–	–	–	–
4	10	5.7 $\pm$ 0.5 (5.1–6.9)	2	8	44.3 $\pm$ 5.4	–	–	–	–	–	1	–	–	–	–
5	11	6.5 $\pm$ 0.3 (6.0–7.1)	4	7	45.3 $\pm$ 10.3	1	1	–	–	3	2	–	–	–	–
6	4	6.4 $\pm$ 0.3 (6.1–6.7)	3	1	33.7 $\pm$ 9.0	1	–	–	–	1	–	–	–	–	–

easy phonics and were known by all of the pupils: primary school children are assessed with high frequency word lists from Reception (age three years), several times a year. By selecting high frequency words, familiarity or practice were unlikely to affect each child’s performance over the test session.

In the RRT, the task is to read as many words as possible in one minute, starting at the first row on the page. Here, the 60 test words were embedded in a page of text and positioned to start on the seventh line and to continue for five lines, so that the test passage had text above and below it. The passage that comprised the words for the reading task was dot-marked at both ends (see Fig. 1). This modification to the RRT was introduced to standardise the task for children with differing reading abilities and avoid floor-effects. In addition, it:

- (i) was more economical on each child’s attention span, and their patience, compared to the RRT, as the majority completed the task in less than one minute and there were eleven reading tasks to complete;
- (ii) still met the criterion of being goal oriented as the child worked towards the end-dot (a trait often attributed to primary school children);

- (iii) presented lines of text above and below each standard passage, which should maximise the possibility of provoking visual stress in those susceptible to it and enable any beneficial effects of the overlays or rulers to be more clearly seen.

As recommended in the *Intuitive Overlays* assessment pack, each reading test consisted of two columns of text (120  $\times$  167 mm, surrounded by 20 mm margins) printed on landscape layout A4 white paper (125 gsm). These two text columns mimicked two pages of an opened book that the children would be familiar with in their educational setting. The test passage was embedded within the left hand column of text, the right hand column served as a comparison for simultaneous judgements of text clarity and comfort with, and without, each coloured reading aid when each aid was placed over the left hand column (Fig. 1). The instructions used followed those stipulated in the RRT (Wilkins et al., 1996; Wilkins, 2003) and the guidelines in the screening pack from Crossbow Education.

**2.2. Participants**

The participants were four to seven year old children recruited from

**Table 2**

A. Reported favourite colours prior to testing in each group (i.e. not the favourite choice of overlay or ruler). B. The number (N) of children who preferred each coloured reading aid, compared to the plain white page of text. C. The number of children who had the fastest reading time with each coloured reading aid, for overlays and rulers. D. The number of children who reported the clearest and most comfortable text with each colour, for overlays and rulers, compared to the baseline no reading aid condition. E. The number of children who had the slowest reading time with each coloured reading aid, for overlays and rulers.

A. Favourite colours			B. Preference over none			C. Fastest colour			D. Most comfortable colour			E. Slowest colour							
Overlay	N	Ruler	N	Overlay	N	Ruler	N	Overlay	N	Ruler	N	Overlay	N	Ruler	N	Overlay	N	Ruler	N
blue	14	blue	19	yellow	45	aqua	47	celery	10	sky	10	yellow	12	yellow	17	none	15	jade	15
red	8	red	11	grass	42	magenta	46	yellow	8	magenta	10	pink	8	magenta	7	jade	6	none	8
pink	8	pink	8	magenta	40	sky	45	aqua	7	pink	7	purple	7	celery	6	magenta	5	grass	8
purple	8	yellow	6	celery	40	orange	45	magenta	6	aqua	7	aqua	7	pink	6	yellow	5	orange	4
green	5	purple	5	sky	40	celery	43	pink	4	grass	5	magenta	6	aqua	5	aqua	4	celery	4
orange	3	green	2	aqua	39	yellow	42	purple	4	celery	4	celery	5	sky	5	celery	4	purple	4
yellow	3	orange	1	purple	38	jade	42	sky	3	purple	3	orange	3	orange	2	orange	4	magenta	3
black	2	white	1	orange	36	grass	42	orange	3	none	2	jade	3	purple	2	pink	3	aqua	3
white	1	black	0	pink	35	pink	40	jade	3	orange	2	sky	2	grass	2	purple	3	sky	2
grey	1	grey	0	jade	34	purple	38	grass	3	yellow	2	grass	0	jade	1	grass	3	yellow	1
								none	2	jade	1	none	0	none	0	sky	1	pink	1

Year 1 to Year 3 of a primary school in North-East London (Table 1). As the participants were children, consent was obtained by sending information letters to the parents and caregivers giving them the opportunity to opt-out if they did not wish their child to participate. The study received ethical approval from the Department of Psychological Sciences ethics committee, Birkbeck College, University of London and in accordance with the Code of Ethics of the World Medical Association (Declaration of Helsinki). One parent of the 118 who replied that they wished their child to be excluded and eleven of the youngest children were excluded as they could not read the test words sufficiently fluently.

Of the remaining 106 pupils, 17 were in Year 1, 67 were in Year 2 and 22 were in Year 3. Within each year group there were children reading at multiple levels using the classification levels used by their school, which was based on the then national curriculum in the UK. The classification levels are incremental and vary from 1 (lowest) to 7 (highest). All of the pupils classified at these levels can read high frequency short words. In addition, level 1 readers can distinguish words, letters and spaces, blend phonemes and use phonic knowledge to attempt unknown words. Level 2 readers can, in addition, read familiar words in simple contexts. Level 3 readers can blend and segment sounds in consonant clusters to use in reading. Level 4 readers can add long vowel phonemes to aid reading and use a range of strategies to decode words. Level 5 readers can, in addition, read polysyllabic words and can read aloud with intonation and expression. Level 6 readers can, further, recognise prefixes and suffixes to help to understand the meaning of words in context. Level 7 readers can also read independently and recognise consonant diagraphs (e.g. kn, wr, ph). As there were only three pupils who read at level 1, they were combined with those reading at level 2. Similarly, there was only one pupil who read at level 7, who was thereby included with the group reading at level 6 (see Table 1).

Binocular near visual acuity was assessed for each child with a Lighthouse near visual acuity chart, viewed at 40 cm (with glasses for the three who wore them). None was diagnosed with dyslexia.

The allocation of participants was equally distributed between the overlay and reading ruler groups. Fifty-three children were tested with A5 overlays, and 53 with reading rulers. Each group had approximately the same number of children from each reading level (Table 1).

2.3. Procedure

The 30–40 min long assessment session was carried out during school time, in a quiet, small room that was similar to a classroom and lit by natural and artificial light. After recording their age and gender, each child was asked several short questions about their reading experiences/history and about their favourite colours. To familiarize them

with the task, they were then shown an example of the test pages, introduced to the reading task, and informed about its non-contextual nature. The instructions for the reading task were taken from Wilkins et al. (1996) and Wilkins (2003). Thus, prior to testing, each child had seen the reading passages and understood they were reading words in a nonsensical order rather than reading normal sentences. Separate text examples were read by the experimenter with each child until the child expressed a wish to read for themselves (which took less than a minute). The experimenter was known to the children and the children approached the task as a normal reading session. All included in the study, from 4 to 7 years, understood the task. They were offered breaks if they wished, but none did so.

For each reading task, the stimuli were positioned on a flat table in front of the children, with a viewing distance of approximately 40 cm. The children read aloud, and this was recorded with a Sony digital voice recorder (ICD-PX820) so that errors could be tabulated subsequently. The time taken to read the 60 words was recorded with a digital stopwatch. The experimenter, rather than the child, pointed to the start of each consecutive line, to ensure each child did not skip lines. The reading rulers exposed four lines of text (Fig. 1) and the experimenter moved the reading ruler down each page when needed as the child reached the fourth line visible through the reading ruler.

Both groups first read the 60-word text passage with no reading aid, to provide the baseline reading rate. Each child then read another ten 60-word text passage with each of the ten coloured reading aids, either overlay or reading ruler. Each colour was paired with a different page of text. On each trial, before the reading, each child was asked whether the coloured (left hand side) or the plain white side of the page (right hand side) was more comfortable to see and easier to read. The reading time (RT) was then recorded. At the end, each of the colours that the child had said made the text more comfortable to view than the plain white text was compared sequentially side by side. Comparisons of the most comfortable of each pair of colours continued until only one particular colour remained, namely the colour that gave the most comfortable visual experience (after Wilkins, 2003). The experimenter emphasised to the children that they needed to choose the ‘most comfortable colour’ based on its effect on the clarity and visibility of the text and not on their favourite colour. All children understood what comfortable, clear and visible (easy to see) meant. The children seldom chose their own favourite colour over the most comfortable colour (Table 2). For 78 children the colours were presented in the following order: magenta, celery, sky, pink, yellow, grass, aqua, purple, orange and jade, following the protocol described by Wilkins (2003) and Crossbow Education. This order is recommended by both so that similar, as well as opposite (“complementary”), colours are not presented in near proximity (Wilkins, 2003, p. 60). Complementary colours are

those that, when mixed together additively, produce a neutral, grey or white colour. (The aim to avoid complementary colours may be questioned, however, in the cases of the sequences magenta/celery and orange/jade.) To allay concerns about possible order effects such as practice, fatigue or the sequential use of near complementary colours, 28 children were presented with the coloured reading aids in random order. Presentation order was addressed in the statistical analyses.

When the reading tasks were completed, each child’s binocular near visual acuity was assessed with a Lighthouse LOGMAR near visual acuity chart. This chart has two sides, with different letter orderings on each side. Acuity was assessed from one side of the chart without any overlay and from the other side covered with the most effective coloured overlay (the one that produced the fastest RT), to determine whether the colour enhanced the children’s visual acuity. Finally, the children were rewarded with stickers and were offered their preferred reading materials to keep so that they could continue to use them as long as they wanted to.

### 3. Results

#### 3.1. Statistical analyses

The data collected were reading times (RT), the number of skipped words, and the number of errors/mis-read words, when each child read with no reading aid (baseline) and with each of the ten coloured reading aids. RTs were then extracted for three conditions:

- baseline;
- the fastest RT achieved by each child with one of the ten colours (“the most effective colour”, Table 2, Column C);
- RTs with the colour judged by each child to give the text the greatest clarity and to be the most comfortable to view (“the most comfortable colour”, Table 2, Column D).

The RTs with the most effective and most comfortable colours were necessarily individually determined for each child. Within each group (overlay or ruler), therefore, these RTs reflected performances with a selection of different colours for different children. Separate comparisons of the two groups’ performance with each of the ten colours were, therefore, not performed as each comparison would necessarily combine the fastest, slowest and in-between RTs for different children, which would eradicate the meaningfulness of such an analysis (Table 2). For example, overall, the difference in RT between the fastest and slowest RTs was, for both groups, approximately 10 s and these differences, and intermediate RTs, were spread across the different colours. It was for this reason that RTs for the above three conditions were, instead, extracted.

Baseline RTs were normally distributed for each group, however, the RT data for the most effective and the most comfortable coloured aids were not: they were positively skewed (Kolmogorov-Smirnov tests). This was rectified using a natural logarithm transformation so parametric analyses could be conducted on the most effective and most comfortable coloured aids. Reading level had a limited range, so associations between reading level and RTs were assessed with the non-parametric Spearman’s correlation coefficient. The data for the number of skipped words, and the number of errors/mis-read words, for the most effective and the most comfortable colours, compared to baseline, were also not suitable for parametric testing and so were assessed with Mann-Whitney U tests, and changes in each of these measures were assessed with Sign tests. Most of the statistical analyses were performed using PASW statistics version 23.0 (SPSS Inc., Chicago, IL, USA), apart from the Sign tests, which were calculated by hand.

#### 3.2. Baseline reading performance

There were no significant differences between the ages of the

children in the overlay and ruler groups [ $t(1\ 0\ 4) = 0.66$ , NS] nor between their initial reading levels (coded as 1 to 7 [ $t(1\ 0\ 4) = 0.54$ , NS]). The baseline RTs decreased with increasing reading level, as expected, in both groups (Table 1). Overall, however, there were no significant differences between the two groups’ baseline RTs [ $t(1\ 0\ 4) = 0.7$ , NS]. There were also no significant group differences in the number of children who skipped words or made errors (mis-read words) at each reading level in the baseline condition (two Mann-Whitney U tests to compare groups for each measure, NS, Table 1). Thus, both groups had a comparable range of reading abilities prior to being tested with the reading aids. Finally, near binocular acuity did not correlate significantly with the baseline RT for either type of reading aid [overlay:  $r_s(53) = -0.12$ , NS; reading ruler:  $r_s(53) = 0.07$ , NS].

#### 3.3. Preference for each coloured reading aid

All but one child said they preferred the text covered with at least one of the coloured reading aids, compared to no reading aid (Table 2B), and all but four children had decreased RTs using at least one of the coloured reading aids (Table 2C). Overall, the reading rulers were preferred slightly more frequently than the overlays, compared to the no-reading aid condition (Table 2B). Given these frequencies, an unexpected result was that when the children were offered a reading aid to take home with them, 37/53 (70%) who were assessed with rulers selected an overlay rather than a ruler (Table 3). It is possible that because the overlays were larger they were seen to be more desirable to a child. There were no clear effects of age or reading level on the choice of overlay or reading ruler to take home, however, so it was not the case that it was the younger children who preferred the larger reading aids (Table 5): on the whole, children at each age group preferred to take away the larger aids. The most effective and the most comfortable colours were chosen to be taken home in approximately equal numbers.

#### 3.4. Changes in reading times (RT) with the coloured reading aids

The maximum decrease in RT with the most effective colour was 23.1 s (s) in the overlay group and 18.9 s in the reading ruler group, compared to baseline. Overall, the average decrease with the most effective colour was 8.3 s (22%) in the overlay group and 7.8 s (20%) in the reading ruler group (Table 4). For the most comfortable colour, the maximum decrease in RT was also 23.1 s in the overlay group and 17.7 s in the reading ruler group. The average decrease with the most comfortable colour was 6.1 s (16%) in the overlay group and 5.5 s (14%) in the reading ruler group (Table 5).

Two two-factor mixed ANOVAs on the  $\log_e$  transformed RT data were conducted with group as the between-subjects factor (type of reading aid: overlay vs ruler) and colour condition as the within-

**Table 3**  
The reading aids and colours that the children chose to take home with them for each group. N = number who selected each colour.

OVERLAY GROUP				READING RULER GROUP			
Colour	Total N	Overlay	Ruler	Colour	Total N	Overlay	Ruler
yellow	12	9	3	yellow	10	7	3
celery	7	3	4	magenta	9	7	2
magenta	7	5	2	celery	7	5	2
aqua	6	4	2	sky	7	2	5
purple	5	3	2	pink	6	4	2
sky	5	5	0	aqua	5	4	1
pink	5	4	1	purple	3	2	1
jade	3	2	1	grass	3	3	0
orange	3	1	2	jade	2	2	0
grass	0			orange	1	1	0
none	0			none	0		

**Table 4** Reading times (RT), skipped words, errors and median acuity for each reading level calculated using the reading aid that produced each child's fastest reading time (the most effective colour). The 'average RT change' values are relative to the no reading aid condition (Table 1). Negative values indicate that the children read more quickly with the reading aid. The number of skipped words and errors should also be compared to values in Table 1. The acuity measures are metric letter sizes. The Lighthouse near visual acuity chart presents metric letter sizes (M) and their Snellen equivalents: 0.3 M = 20/16, 0.4 M = 20/20, 0.5 M = 20/25, 0.6 M = 20/32, 0.8 M = 20/40, 1.25 M = 20/63. Only one child had the latter acuity and it was improved with the use of their most effective overlay. \*\* $p < 0.005$ , \* $p < 0.01$ . See text for further details.

Reading level	Average RT ± 1 SD (secs)	Average RT change ± 1 SD (%) change)	Number of children who skipped N words							Number of children who made N errors							Acuity no aid (range)	Acuity with aid (range)
			N = 1	2	3	4	5	6	N = 1	2	3	4	5	6				
<b>OVERLAY GROUP</b>																		
1/2	42.5 ± 8.9	-8.6 ± 4.6 (20)	1	-	1	-	-	-	-	-	-	-	-	-	-	0.45 (0.3–0.6)	0.4 (0.4–0.6)	
3	37.6 ± 8.3	-10.6 ± 5.3 (28)	1	1	-	-	-	-	-	1	-	-	-	-	-	0.5 (0.3–0.8)	0.45 (0.3–0.8)	
4	40.1 ± 10.5	-9.9 ± 3.9 (25)	2	-	-	-	-	-	-	2	-	-	-	-	-	0.5 (0.3–0.5)	0.4 (0.3–0.5)	
5	33.3 ± 8.8	-7.8 ± 4.5 (23)	1	-	1	-	-	-	-	4	1	-	-	-	-	0.5 (0.3–0.6)	0.4 (0.3–0.6)	
6/7	33.9 ± 5.3	-2.1 ± 2.5 (6)	-	1	1	-	-	-	-	-	-	-	-	-	-	0.4 (0.3–1.25)	0.3 (0.3–0.5)	
Overall	38.0 ± 9.1	-8.3 ± 5.0 (22)	-	-	-	-	-	-	-	-	-	-	-	-	-	0.5	0.4**	
<b>READING RULER GROUP</b>																		
1/2	41.6 ± 7.9	-8.7 ± 4.8 (21)	2	2	1	-	-	-	-	5	2	1	1	-	-	0.4 (0.3–0.6)	0.4 (0.3–0.8)	
3	41.0 ± 13.2	-7.7 ± 5.0 (19)	1	-	-	-	-	-	-	2	-	-	-	1	-	0.5 (0.3–0.6)	0.4 (0.3–0.6)	
4	36.6 ± 3.6	-7.7 ± 5.6 (21)	2	-	-	-	-	-	-	2	-	-	-	-	-	0.5 (0.3–0.6)	0.5 (0.3–0.6)	
5	38.2 ± 8.4	-7.1 ± 3.9 (19)	4	-	-	-	-	-	-	2	1	-	-	-	-	0.5 (0.3–0.5)	0.4 (0.3–0.5)	
6	27.3 ± 6.6	-6.4 ± 2.4 (23)	-	-	-	-	-	-	-	-	-	-	-	-	-	0.55 (0.4–0.6)	0.45 (0.4–0.5)	
Overall	38.7 ± 9.5	-7.8 ± 4.6 (20)	-	-	-	-	-	-	-	-	-	-	-	-	-	0.5	0.4*	

subjects factor (baseline vs the most effective colour; baseline vs the most comfortable colour). Each ANOVA revealed a significant effect of colour condition [most effective colour:  $F(1,104) = 332, p < 0.001$ ; most comfortable colour:  $F(1,104) = 116, p < 0.001$ ] but no significant difference between the groups (overlay or ruler) nor a significant interaction between group and colour condition [all  $F_s < 1, NS$ ]. Thus, the RTs with the most effective or the most comfortable coloured reading aids were significantly faster than the initial assessment with no reading aid (baseline) (Tables 4 and 5). Note that the baseline vs the most effective colour comparison is *not* simply comparing the fastest with the slowest times, as the slowest reading times were not always achieved in the baseline condition (Table 2).

The most effective and most comfortable colours differed for 62 of the children (29 in the overlay group, and 33 in the reading ruler group). To compare directly performance for the most effective and the most comfortable colours overall, RT difference scores from baseline were entered into another two-factor ANOVA as a within-subjects variable (type of colour: 'most effective' vs 'most comfortable'), together with type of reading aid as a between-subjects variable (overlay or ruler). These difference scores were normally distributed (Kolmogorov-Smirnov tests,  $p > 0.05$ ). There was a significant main effect of type of colour [ $F(1,104) = 59, p < 0.001$ ] but no other significant effects (both  $F_s < 1, NS$ ). So, when the most effective and the most comfortable colours differed, the RTs were faster with the most effective colour, compared to the most comfortable colour, but again, the effects of the two types of reading aid (overlay or ruler) on performance did not differ.

Binocular visual acuity was also tested with no reading aid and with the most effective coloured aid (Table 4). Overall, acuity improved statistically significantly with the coloured aid for both groups [Sign tests—overlay: total changes = 27, negative changes (improvement in acuity) = 22, positive changes (deterioration in acuity) = 5,  $p < 0.001$ ; ruler: total changes = 26, negative changes = 23, positive changes = 3,  $p < 0.001$ , see Table 4]. Thus, in each group, over 40% of the children experienced an improvement in acuity with their most effective coloured aid.

### 3.5. Correlations between reading level and changes in reading times

It was hypothesised that the coloured reading aids may have more of an impact on the younger children, and less of an impact on the older, more competent, readers. There were indeed significant positive correlations between the children's reading levels (coded as 1 to 7) and their improvement in RTs when they used their most effective coloured overlay [overlays:  $r_s(53) = 0.37, p < 0.005$ ; rulers:  $r_s(53) = 0.12, NS$ , one tailed tests]. The positive correlation indicates that the less experienced readers showed the largest improvements from the usage of their most effective coloured overlay, compared to the more experienced readers. Thus, the largest improvement was observed with children whose reading skill was low (Fig. 3A).

In the overlay group again, there was also a significant positive correlation between the children's reading levels and their improvement in RTs when they used their most comfortable colour [overlays:  $r_s(53) = 0.23, p < 0.05$ ; reading rulers: [ $r_s(53) = -0.01, NS$ ] (Fig. 3B). In the reading ruler group, the changes in RT with overlays or rulers were fairly uniform across the reading levels, resulting in the lack of significant association between RT and reading level.

### 3.6. Skipped words and errors

The number of skipped words and the number of errors/mis-read words were calculated for the baseline, the most effective colour and the most comfortable colour. Generally, each type of aid in each condition decreased both types of errors compared to the baseline rate, but only two of the decreases were statistically significant (Table 6).

**Table 5** Reading times (RT) and errors for each reading level calculated using the reading aid that each child reported made the text clearest and easiest to view (the most comfortable colour). The average RT change values are relative to the no reading aid condition (Table 1). Negative values indicate that the children read more quickly with the reading aid.

Reading level	Average RT ± 1 SD (s)	Average RT change ± 1 SD (% change)	Number of children who skipped N words						Number of children who made N errors				Proportion of children who chose overlays or (rulers) to take home	
			1	2	3	4	5	6	N=	1	2	3		4
<b>OVERLAY GROUP</b>														
1/2	45.8 ± 9.4 (13)	-5.3 ± 5.4	1	-	-	1	-	-	4	2	1	-	-	0.71 (0.29)
3	40.3 ± 10.6 (22)	-7.8 ± 6.2	1	2	-	-	1	-	1	-	2	1	-	0.67 (0.33)
4	40.4 ± 10.2 (25)	-9.6 ± 3.8	1	-	-	-	-	-	-	-	-	-	-	0.33 (0.67)
5	35.3 ± 9.8 (17)	-5.8 ± 4.2	-	1	-	-	-	-	-	2	-	-	-	0.82 (0.18)
6/7	36.4 ± 6.9 (0)	+0.3 ± 4.8	1	-	-	-	-	-	1	1	-	-	-	0.86 (0.14)
Overall	40.3 ± 10.0 (16)	-6.1 ± 5.7												
<b>READING RULER GROUP</b>														
1/2	44.4 ± 8.6 (15)	-6.0 ± 5.4	2	1	1	-	-	-	3	1	-	1	-	0.67 (0.33)
3	43.6 ± 15.3 (13)	-5.2 ± 4.7	3	-	-	-	1	-	-	-	1	2	-	0.69 (0.31)
4	39.3 ± 4.0 (13)	-5.0 ± 6.5	-	-	-	-	-	-	2	-	1	-	-	0.80 (0.20)
5	39.2 ± 8.2 (15)	-6.0 ± 4.4	-	2	1	-	-	-	3	2	-	-	-	0.64 (0.36)
2A	28.7 ± 5.8 (16)	-5.0 ± 3.5	-	-	-	-	-	-	-	-	1	-	-	1.0 (0.0)
Overall	41.0 ± 10.5 (14)	-5.5 ± 5.0												

### 3.7. Order effects

For 78 children, the reading aids were administered in the same order, following the protocol described by Wilkins (2003) and Crossbow Education. It is possible that there may have been order effects, such as practice at the start of the test sessions, and boredom or tiredness towards the end. To assess this, 28 of the children were tested with the reading aids administered in random order.

For those children who were tested with each coloured reading aid in the same order, there was no clear pattern consistent with order effects for either type of reading aid. Decreases and increases in RT, compared to baseline, were distributed across the different colours for each child, and therefore occurred at different places in the test sequence. It is, therefore, unlikely, that an increased familiarity with the already familiar words, nor practice, accounted for the RT differences, as there then would have been a trend for a decrease in RT as each test passage was read, which was not observed (Tables 2 and 6). Finally, visual acuity was tested at the end of the test session and was found to improve on average with the most effective coloured aid (Table 4), which is also not consistent with boredom or fatigue. To pursue this further, however, the performance of the 78 pupils tested with a fixed order were compared to the 28 who were tested with the order of colour aid randomised.

A mixed two-factor ANOVA with log<sub>e</sub> transformed RTs was performed with colour of reading aid (none or the 10 coloured aids) as a within subjects factor and group (fixed or random order) as a between subjects factor produced only a significant main effect of reading aid colour [F(10,1040) = 11.0, p ≤ 0.001]: the RTs differed for the different coloured reading aids. The main effect of group, and the interaction between colour and group were not statistically significant (F's < 1.4). For both overlay and reading ruler groups, there were no significant differences in log<sub>e</sub> transformed RTs between the most effective or most comfortable colours [overlay—most effective: t(51) = -0.2, NS; most comfortable: t(51) = 0.01, NS; reading ruler—most effective: t(51) = -1.9, NS; most comfortable: t(51) = -1.7, NS].

## 4. Discussion

In this study, all but one child reported improvements in the clarity of the text and their RT decreased when tested with either type of coloured reading aid: A5 overlay or reading ruler (Tables 4 and 5). The size of the reading aid did not affect RT performance significantly. The number of skipped words and the number of errors/mis-read words also decreased when reading with the coloured aids (Tables 1, 4 and 5). There were not very many of either type of error, however, so only a few of the decreases were significant (Table 6). Both groups had a comparable range of reading abilities prior to being tested with the reading aids, so initial reading ability does not confound these comparisons (Table 1). Practice from multiple testing, or then boredom, may have been present during the test session but they are unlikely to account for these results as the performance of those tested with a fixed test sequence, and those tested with a randomized test sequence, did not differ significantly. From the results of the present study, coloured aids improved RTs on the RRT employed here and the size of the reading aid was not crucial to facilitate performance (see also Waldie & Wilkins, 2003). Both sizes of reading aid were also preferred to no reading aid (baseline) to a similar extent (Table 2). While a repeat of the baseline test was not completed at the end of the reading of the sets of text passages, acuity was tested at the end of the test session, with and without the most effective reading aid. The most effective reading aid was found to improve the reading of the letters on the acuity chart. Boredom or fatigue cannot account for this result (see also below).

The task and procedure were based on the guidelines described by Wilkins (2003), however, here they were extended by determining both the most effective as well as the most comfortable colours. The children were asked to read a set passage with each colour, as well as to judge

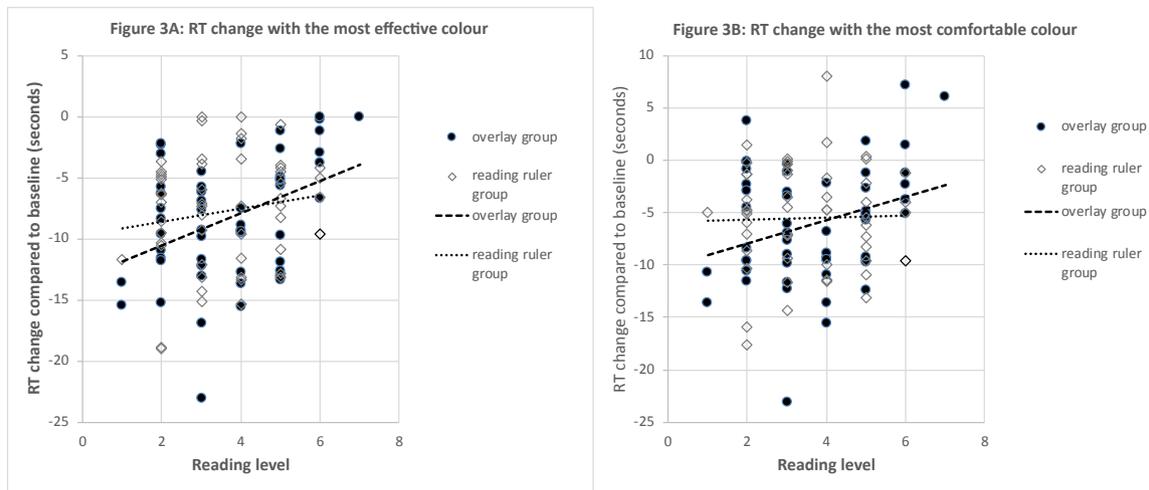


Fig. 3. The correlations between reading level and RT improvements with the most effective and most comfortable colours. Reading level was coded as 1 to 7. Negative RT changes indicate that the child read more quickly with the reading aid compared to the baseline, no reading aid, condition.

which colour gave text the greatest perceptual clarity and comfort. This method may decrease effects of placebo or demand characteristics (children tend to like to impress the experimenter), because they have no clear idea about which colour should be the most effective or the most comfortable. They appeared to do the task conscientiously and did not, for example, just choose their favourite colour. The most effective coloured aid also improved over a third of the children’s visual acuity (Table 4). Changes in acuity have not been assessed in previous studies. The children commented that their most effective coloured aid decreased the blurring or fuzziness of the letters on the acuity chart, which comprises closely spaced letters. The children reported the letters were more readable with the coloured aid overlaying the chart rather than with black letters on white a white background. If the aids help with identification of letters (the acuity chart) and the reading of random words (the text passages), this early stage assistance may help subsequently with reading normal text: the flow of semantics, grammar and comprehension at these later stages of reading would become a less taxing task.

The most effective and the most comfortable colours are not always assessed in research on reading aids, and here they were found to be not always the same. The lightness of the aids also varied with colour (Fig. 2). Those that reflected the most light (yellow, celery, orange, grass) would reduce the contrast of the text less than the remainder, but these differences do not appear to be associated with the colours that were the most effective or most comfortable (Table 2). The differences between the most effective and most comfortable colours may reflect

the very different experiences involved in simply observing text through a particular colour rather than being tested while reading with it.

As suggested by the reviewers, further research could, first, address the effects of contrast reduction, in combination with the colour of the aid, on reading performance. In this study, achromatic reading aids were not included, yet they could be included to provide a more appropriate baseline than the ‘no overlay’ condition included here as baseline. The three reading aids that produced the fastest reading times, and those that made the text more comfortable to view, included not only the lightest (yellow and celery), but also magenta, aqua, sky, pink and purple, which were darker (Table 2). The lightest aids (yellow and celery) were also selected to take home frequently, but so was magenta, one of the darker. In future research it could be asked, for example, if an aid that reduces contrast to a larger extent is followed by one that reduces contrast to a lesser extent (or vice versa), affects performance and the reported clarity of the second text passage that is read, regardless of its colour? This is a separate issue to avoiding presenting complementary colours in succession. It was not possible to disentangle colour from lightness and contrast reduction with the reading aids tested in the present study. Administering coloured and achromatic aids together, in future research, would be one way to keep lightness, and contrast reduction, constant and allow an assessment of colour *per se*.

Second, this study used the reading aids available from Crossbow Education. It was not possible to examine the effects of coloured aids that lay on particular lines in colour spaces that represent the early

Table 6

Changes in the number of errors (skipped words and mis-read words) for the most effective colour and the most comfortable colour (Tables 4 and 5) relative to the baseline error rates (Table 1), for each reading aid condition. Positive changes indicate that fewer errors were made in the baseline condition than in the colour condition, negative changes indicate that fewer errors were made in the colour condition than in the baseline condition. Statistical significances were tested with Sign tests.

	OVERLAY GROUP				RULER GROUP			
	Total Changes	Positive changes	Negative changes	Significance (one-tailed)	Total changes	Positive changes	Negative changes	Significance (one-tailed)
<i>Most effective colour</i>								
Skipped words	19	8	11	NS	19	5	14	$p = 0.02$
Mis-read words	27	8	19	$p = 0.01$	23	11	12	NS
<i>Most comfortable colour</i>								
Skipped words	18	7	11	NS	18	9	9	NS
Mis-read words	25	9	16	$NS (p = 0.06)$	25	14	11	not tested*

stages of colour processing in the visual system, such as cone- opponency. For example, colours can be selected that lie on tritan lines: lines of constant long (L) and middle (M) cone excitation, only short (S) cone excitation varies. Such colours plot on vertical lines in the MacLeod and Boynton (1979) diagram, Fig. 2. Nor was it possible to examine colours that lay on lines of constant S-cone excitation (horizontal lines in the MacLeod and Boynton (1979) diagram): the activity of only the L- and M-cones varies. The present data do not allow a formal assessment of these issues, although it is noteworthy that celery, yellow, magenta, aqua, sky, pink and purple span the ordinate of the MacLeod and Boynton (1979) diagram, indicating that a change in tritan (S-cone) sensitivity in those young readers who benefitted from these coloured reading aids is unlikely to explain the effects reported in this study. These issues could be addressed if aids can be manufactured that vary stimulation more systematically in the early cone-opponent pathways, and if lightness can be controlled, with, for example, achromatic aids that are used in conjunction with the coloured aids. to keep lightness and contrast reduction constant.

The differences between the most effective and most comfortable colours could also be explored further in future research. For example, the most effective and most comfortable colours could be assessed initially, and then again after the child has been reading continuously for some time and is beginning to tire (after Tyrell et al., 1995). If tested at that point, the most effective and most comfortable colours may be more likely to coincide. It would also be interesting to see whether the colours changed from the first to second test times, after the child has begun to tire. This study has examined RTs and errors when reading a short passage containing words in a nonsensical order, it could be extended in future work to assess comprehension as well by using sets of text passages comprised of simple, sensible sentences. The text could be restricted to include short, common words, so that the earliest readers are not disadvantaged, and then questions on comprehension after reading each text passage could be included, together with reading times and errors. Such an extension may help clarify or dissociate any effects coloured reading aids have on cognitive aspects of reading, including comprehension, and the visual clarity of the text being read.

Ninety-six per cent of the children read more quickly with at least one of the coloured aids, which is a considerably larger percentage than that reported in previous studies. Wilkins (2003), for example, reported twenty per cent of children are likely to experience some benefit from overlays. This discrepancy may reflect the different methods: here all colours were used while RT for a set passage was recorded, whereas previous studies have used just the one that was reported to make the text the most comfortable to view. As mentioned above, the most effective and most comfortable colours were not always the same. This suggestion is unlikely to account entirely for such a large difference in the number of children who had decreased RTs, however. Another reason might be the inclusion of very young children in this study. As mentioned in the Introduction, many of the children in the earlier studies were older, and sometimes the youngest previously tested have been older than even the oldest children tested here (e.g. Waldie & Wilkins, 2003).

Future research could recruit a larger number of the younger readers (4–6 year olds) to explore this aspect further, although a modified method may be preferable, perhaps using multiple test sessions over several days and using more than one reading test with each colour. Here, the study was designed to keep the test session short, so had one reading test for each colour, so as not to tire the young readers. A future study that included more than one reading test for each colour would enable the size of RT changes to be assessed and compared for different colours. This may yield a figure for the number of children who benefit from coloured reading aids that is more in line with previous estimates, or it may confirm the trend for greater effects in younger children.

This trend was also shown by the correlations between the children's reading level and the maximum improvement achieved by the

child with their most effective colour. These correlations revealed that the less experienced readers benefited more from the usage of overlays or rulers: the largest decreases in RTs were observed with children whose reading skill was low (Fig. 3). It is possible that, while a child is learning to read, the reading aids are most effective and, subsequently, the effect of the overlays on RTs may diminish with age or reading skill. The four children whose RTs did not decrease with either type of reading aid (two with overlays, two with rulers) were reading at higher levels to begin with (levels 4 to 7). The correlations between reading level and RTs with the overlays were significant for both the most effective and the most comfortable colours. For the group given reading rulers, the correlations were not statistically significant. This pattern of correlations could also be confirmed in future research using a larger sample of younger readers.

As mentioned, in this study, one focus was on the comparison of size of reading aid and little difference was found between the two sizes. The reading rulers, however, presented a coloured strip above a darker section where the manufacturer's name was printed, so they combine a coloured strip with a form of typoscope (see Fig. 1). Future research could include a condition where the children read with an achromatic reading ruler or typoscope as a second baseline condition to see whether any benefit from reading rulers is due to their typoscopic properties, or to their colour.

Most of the coloured aids were preferred by most children compared to none. This is not to say all of the colours were beneficial, however: the slowest RTs occurred with no reading aid for only 15 children given overlays and for only 8 given reading rulers (Table 2). Thus, all of the coloured aids may be preferred by the majority of children and be reported to make the text more comfortable to view, but not all of them improved performance. Improved performance occurred only for the most effective and most comfortable colours. For each of these, the colour selected seemed to be idiosyncratic (Table 2). This agrees with other reports that, to obtain benefit, the coloured aids need to be individually determined for each child (for reviews, see Wilkins, 2002; Wilkins, 2003). It is possible that the children said they preferred so many colours as they became aware they would be given one to take home. The most effective and the most comfortable colours were, however, chosen to be taken home in approximately equal numbers. The expectation of being given a reading aid may have led to them preferring so many of the colours, but this does not account for the frequent choice of the most effective colour to take home as they were not told which colour that was and different children read most quickly with different colours. This pattern is also not consistent with a general placebo effect (Wilkins et al., 1994), nor the Hawthorne effect (Mayo, 1933, chap. 3) whereby participants modify their behaviour because they know they are being observed while reading with the overlays or rulers.

Most of the children enjoyed the participation in this study and continued to use their reading aids in the classroom and to ask for them when they did not have them in silent reading time. Unfortunately, it was not possible to conduct a follow up study to determine for how long they continued to use their reading aids. Nevertheless, whether days, weeks, or months, their attention had been drawn to their reading and clarity of text. During the assessments children frequently reported that, with certain colours, they see the text “properly” or they said “wow, the letters got bigger!”, “the coloured side goes bigger” or “it lets me read”. Some of them said about the overlay that “I just read, read, read” or “it allows my eyes to run”. Thirteen of the 33 children who chose the reading rulers rather than the overlays to take home with them had compulsive, repetitive behaviour patterns while reading accompanied with finger pointing, head tilts, eccentric viewing and swinging upper-body movements, which are overt signs of reading difficulties and visual stress (Wilkins, 2003). These children expressed that they wanted to have the reading ruler, because it “helped me to follow the line properly”, “it keeps me on the line, I am not getting lost”. Both types of reading aid that produced the fastest reading time were equally

effective, on average, but one or other size may be better for a particular child depending on their degree of visual stress, that child's abilities and their reading style. These suggestions could be explored in further research.

### Acknowledgements

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors. This research was completed with the support of Crossbow Education Ltd who donated all of the overlays and reading rulers. They also generously offered the participating pupils their reading aids to keep. The study could not have been completed without the support of Sue Head, the Head of Earlsmead Primary School in Tottenham, for providing facilities, and Fiona George for providing support for the experimental sessions. Thanks are also due to Professor Arnold Wilkins, for initial advice and for measuring the reflectances of each of the coloured reading aids (depicted in Fig. 2), and to Dr Deacon Harle for advice on the manuscript.

### Disclosure

The authors report no conflicts of interest and have no proprietary interest in any of the materials mentioned in this article and no commercial associations. Crossbow Education provided the test materials, but we have no further association with them and they were not involved in the production of this paper.

### References

- Albon, E., Adi, Y., & Hyde, C. (2008). *The effectiveness and cost-effectiveness of colored filters for reading disability: A systematic review*. Birmingham, United Kingdom: University of Birmingham Department of Public Health and Epidemiology.
- Conlon, E. G., Lovegrove, W. J., Chekaluk, E., & Pattison, P. E. (1999). Measuring visual discomfort. *Visual Cognition*, 6, 637–666.
- Ellis, A. W. (1993). *Reading, writing and dyslexia: A cognitive analysis*. Hove, UK: Lawrence Erlbaum Associates.
- Evans, B. J. W., Allen, P., & Wilkins, A. J. (2009). Vision and reading difficulties Part 2: Optometric correlates of reading difficulties. *Optometry Today*, 27, 30–38.
- Evans, B. J., Patel, R., & Wilkins, A. J. (2002). Optometric function in visually sensitive migraine before and after treatment with tinted spectacles. *Ophthalmic and Physiological Optics*, 22, 130–142.
- Harle, D. E., Shepherd, A. J., & Evans, B. J. W. (2006). Visual stimuli are common triggers of migraine and are associated with pattern glare. *Headache*, 46, 1431–1440.
- Henderson, L. M., Tsogka, N., & Snowling, M. J. (2012). Questioning the benefits that coloured overlays can have for reading in students with and without dyslexia. *Journal of Research in Special Educational Needs*, 13(1), 57–65.
- Hughes, L. E., & Wilkins, A. J. (2000). Typography in children's reading schemes may be suboptimal: Evidence from measuring reading rate. *Journal of Research in Reading*, 23(3), 314–324.
- Hyatt, K. J., Stephenson, J., & Carter, M. (2009). A review of three controversial educational practices: Perceptual motor programs, sensory integration, and tinted lenses. *Education & Treatment Children*, 32(2), 313–342.
- MacLeod, D. I. A., & Boynton, R. M. (1979). Chromaticity diagram showing cone excitation by stimuli of equal luminance. *Journal of the Optical Society of America*, 69(8), 1183–1186.
- Mayo, E. (1933). *The human problems of an industrial civilization*. New York: MacMillan.
- National Curriculum Assessments in England (2011/12, revised): <http://media.education.gov.uk/assets/files/pdf/m/sfr33-2012v2.pdf> (page 4.) <http://www.education.gov.uk/researchandstatistics/statistics/a00217264>.
- Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, 124(3), 372–422.
- Rayner, K. (2009). Eye movements and landing positions in reading: A retrospective. *Perception*, 38(6), 895–899.
- Ritchie, S. J., Della Sala, S., & McIntosh, R. D. (2011). Irlen colored overlays do not alleviate reading difficulties. *Pediatrics*, 128(4), e932–e938.
- Royal College of Ophthalmologists. (2009). Reading and the visual system. Available at: [www.rcophth.ac.uk/core/core\\_picker/download.asp?id=361](http://www.rcophth.ac.uk/core/core_picker/download.asp?id=361).
- Shepherd, A. J. (1997). A vector model of colour contrast in a cone excitation colour space. *Perception*, 26, 455–470.
- Shepherd, A. J. (1999). Remodelling colour contrast: Implications for visual processing and colour representation. *Vision Research*, 39, 1329–1345.
- Shepherd, A. J., Beaumont, H. M., & Hine, T. J. (2013). Color and spatial frequency are related to visual pattern sensitivity. *Headache*, 53, 1087–1103.
- Smith, L., & Wilkins, A. J. (2007). How many overlay colours are necessary to increase reading speed? A comparison of two systems. *Journal of Research in Reading*, 30, 332–334.
- Solan, H. A., & Richman, J. (1990). Irlen lenses: A critical appraisal. *Journal of the American Optometry Association*, 61(10), 789–796.
- Tyrell, R., Holland, K., Dennis, D., & Wilkins, A. (1995). Coloured overlays, visual discomfort, visual search and classroom reading. *Journal of Research in Reading*, 18(1), 10–23.
- Waldie, M., & Wilkins, A. J. (2003). How large does an overlay have to be? *Ophthalmic and Physiological Optics*, 24(1), 57–60.
- Wilkins, A. J. (1994). Overlays for classroom and optometric use. *Ophthalmic and Physiological Optics*, 14, 97–99.
- Wilkins, A. J. (1995). *Visual stress*. Oxford, UK: Oxford University Press.
- Wilkins, A. J. (1996). A system for precision ophthalmic tinting and its role in the treatment of visual stress. In C. M. Dickinson, I. J. Murray, & D. Garden (Eds.). *John Dalton's Colour Vision Legacy* (pp. 709–715). London: Taylor and Francis.
- Wilkins, A. J. (2002). Coloured overlays and their effects on reading speed: A review. *Ophthalmic and Psychological Optics*, 22, 448–454.
- Wilkins, A. J. (2003). *Reading through colour*. Chichester, UK: Wiley.
- Wilkins, A. J., Evans, B. J. W., Brown, J. A., Busby, A. E., Wingfield, A. E., Jeanes, R. J., & Bald, J. (1994). Double-masked placebo-controlled trial of precision spectral filters in children who use coloured overlays. *Ophthalmic and Physiological Optics*, 14, 365–370.
- Wilkins, A. J., Jeanes, R. J., Pumfrey, P. D., & 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, 491–497.
- Wilkins, A. J., Lewis, E., Smith, F., & Rowland, E. (2001). Coloured overlays and their benefit for reading. *Journal of Research in Reading*, 24(1), 41–64.
- Wilkins, A. J., Nimmo-Smith, M. I., Tait, A., McManus, C., Della Sala, S., Tilley, A., ... Scott, S. (1984). A neurological basis for visual discomfort. *Brain*, 107, 989–1017.