Visual stress theory and its application to reading and reading tests

Arnold Wilkins
University of Essex, UK

Jie Huang
Michigan State University, USA

Yue Cao
University of Michigan, USA

This paper presents a theory of visual stress. The theory is applied to the assessment of symptoms of visual stress and its subsequent treatment with coloured filters. The theory has implications for standard reading assessments that relate both to the visual skills and the age of the children taking the tests. These implications are reviewed, with the conclusion that insufficient attention is paid to visual factors in reading, particularly is the design of reading tests.

Introduction

Very many tests are currently used in the assessment of reading. The Southwest Educational Laboratory (www.SEDL.org) has so large a database that it lists tests according to following categories: reading comprehension, language comprehension, decoding, linguistic knowledge, background knowledge, cipher knowledge, lexical knowledge, phonological awareness, semantics, syntax, phonology, phonemic awareness, knowledge of alphabetic principle and letter knowledge and concepts about print. It is striking that although reading patently requires visual function, none of the above tests seeks to assess the visual function necessary for reading.

It is widely assumed that the tests administered by optometrists are sufficient to pick up any refractive or binocular difficulties with near vision that are likely to impair reading, and that these are the only visual difficulties that can do so. An optometrist will report ‘perfect sight’ if the two eyes are well coordinated and his patient is able to read with either eye a row of small letters subtending a visual angle at the eyes of about five minutes of arc. There are other aspects of vision to consider, however. The ability to see...
small targets is not necessarily a good predictor of the ability to see words in which the letters are clustered and crowded (Chung, 2002; Tripathy and Cavanagh, 2002). The flanking letters reduce the visibility of the letters they surround, particularly in the case of dyslexic individuals (Spinelli, De Luca, Judica and Zoccolotti, 2002). Reading difficulties may also arise from other, more subtle visual deficits. Over the last few years increasing attention has been paid to the deficits that certain dyslexic individuals experience. These have been attributed to magnocellular dysfunction (Stein, 2001). The consequences of these deficits for reading ability are contentious (Ramus et al, 2003).

There is another aspect of visual dysfunction that has been, until recently, almost universally ignored. It is a dysfunction that is not typically measured in the optometric clinic or assessed by psychophysicists, and it occurs in dyslexic and non-dyslexic individuals alike. It is the inability to see comfortably without distortion and discomfort, and is often referred to as visual stress (Wilkins, 1995). Visual stress can interfere with the ability to read for long periods (Tyrrell et al, 1995), and it is surprisingly prevalent in the general population (Wilkins, Lewis, Smith and Rowland, 2001). The purpose of this paper is to draw attention to the mechanisms of visual stress and its association with migraine, outline some methods for assessment and treatment and assess the implications for the measurement of reading skill.

A theory of visual stress

The theory of visual stress will now be developed in some detail to show the implications it has for symptoms of visual stress when reading, the treatment of these symptoms, and the design of written material for children, including reading tests.

Links with migraine

About 10% of the population suffers from severe headaches known as migraine. These individuals usually experience an aversion to light (photophobia) during their headaches, and between the headaches they find certain visual stimuli aversive. These visual stimuli are very similar indeed to the visual stimuli that can provoke seizures in patients with photosensitive epilepsy. The stimuli include flicker and geometric patterns, particularly stripes with certain spatial characteristics. These stripes are generally aversive for the population at large, but are particularly so for people with migraine. Migraineurs are unusually susceptible to the illusions of colour, shape and motion that these patterns induce. Although the illusion susceptibility is present between headaches, it increases in the 24 hours before headache onset. The illusions are asymmetric in individuals who suffer head pain on one side of the head. It has been hypothesised that the illusions are the result of an inappropriate firing of cortical cells due to a spread of excitation within the cortex. The stripes that evoke the illusions are strong visual stimuli and may compromise the inhibitory resources of the cortex that are shared by pyramidal neurons. This would explain why more illusions are reported after sleep deprivation, which is known to increase cortical excitability (Wilkins, 1995).

There are three independent lines of evidence indicating that in migraine the cortex is hyperexcitable. First, migraineurs are more likely than others to report visual phenomena (phosphenes) when the occipital cortex of the brain is subjected to transcranial magnetic stimulation (Aurora and Welch, 1998). Second, three anticonvulsant drugs have now
been shown in double-masked trials to have an anti-migraine as well as anti-epileptic action (Diener and Limmroth, 2001). Third, when individuals with migraine observe stripes that have epileptogenic characteristics the blood oxygenation level-dependent (BOLD) response is abnormally large (Huang, Cooper, Satana, Kaufman and Cao, 2003). This last finding is quite consistent with the greater susceptibility in migraineurs to illusions in patterns of stripes, and with the attribution of these illusions to the effects of a cortical excitability. The susceptibility to illusions in patterns of stripes may have consequences for reading, as we will see.

**Characteristics of aversive stripes**

It is only certain very specific patterns that provoke illusions and are capable of inducing seizures in patients with photosensitive epilepsy. The spatial characteristics of these patterns are shown in Figure 1. The solid curves show the probability of epileptiform EEG activity in patients with photosensitive epilepsy, and the broken curves show the number of illusions of colour, shape and motion reported by normal observers. As can be seen from the figure, it is bright patterns of high-contrast stripes that are most likely to evoke both illusions and seizures, provided the bright and dark stripes have approximately equal width (i.e. the pattern has a duty cycle of 50%), and provided the

**Figure 1.** The characteristics of patterns of stripes that induce epileptiform abnormalities in patients with photosensitive epilepsy (solid curves) and perceptual distortions (illusions of colour, shape and motion) in normal observers (broken curves).

Notes: A diagram beneath each graph show how the pattern changes with the value on the x-axis, and the y-axis shows the probability of epileptiform activity (0–1.0) and mean number of illusions (0–2.0). The shaded areas show the values of the x-axis for adult text, considered as a striped pattern (after 32Wilkins and Nimmo-Smith, 1987). The arrows on the x-axis show the values for a sample of text designed for five-year olds (Excerpt A from 10Hughes and Wilkins, 2000).
spatial frequency of the pattern is about three cycles per degree (i.e. one pair of stripes subtends about 20 minutes of arc at the eye). These are the very patterns that elicit the abnormal BOLD response in migraineurs (Huang et al, 2003).

**Text as stripes**

The lines of a paragraph of text can be considered as a pattern of stripes, in which case the pattern has the characteristics shown by the shaded regions in Figure 1. A wide variety of perceptual distortions can be seen in text, involving illusions of colour, shape and motion. The illusions of shape seen in text can include the rhomboid lattice, typical of that seen in gratings, and the lattice is similarly dependent on spatial frequency (Wilkins and Nimmo-Smith, 1987). The spatial characteristics of the distortions seen in patterns of stripes resemble those seen in text. Reading can evoke epileptiform EEG abnormalities in patients with photosensitive epilepsy, and these abnormalities are reduced when the lines of text are covered above and below the lines being read, using a reading mask (typoscope), suggesting that the pattern from the lines of text is responsible for the EEG abnormalities. It is therefore possible that some of the perceptual distortions of text and some of the discomfort and headaches that occur when reading do so because of the way in which text approximates a striped pattern (Wilkins, 1995).

**Treatment of visual stress using coloured filters**

The perceptual distortions seen in text can often be removed by covering the text with a coloured filter or by illuminating the text with light of a particular chromaticity (Wilkins, 2003). The colour has to be selected with precision to suit the individual, and new techniques for providing the appropriate colour using an *Intuitive Colorimeter* have been developed (Wilkins, 2003). The Colorimeter allows the perceptual effects of coloured illumination to be assessed while the eyes are adapted to colour (Wilkins and Sihra, 2000). A system for precision ophthalmic tinting has been introduced into general optometric practice in the UK, supported by a variety of clinical trials (Evans et al, 1999; Maclachlan, Yale and Wilkins, 1993; Wilkins et al, 1994). The techniques have been shown to increase reading speed by more than 25% in at least 5% of children in mainstream education (Wilkins et al, 2001).

**Neural mechanisms underlying treatment with coloured filters**

The above theory of visual stress can provide an explanation for the beneficial effects of coloured filters in reducing distortions of text and headaches when reading. The argument runs as follows. In photosensitive epilepsy, patients can be sensitive to patterns of stripes within a limited range of orientation, indicating that, at least in these patients, the cortical excitability can be local and confined to a limited number of orientation columns (the groups of cells in the cortex that respond to lines of a particular orientation). It is hypothesised that in migraine the cortical excitability can be similarly local. The representation of colour in the cortex has recently been shown to follow topographic maps in which cells responding to a certain colour are located together and arranged with neighbours of similar colour in a colour sequence reminiscent of the CIE colour diagram, for example, red, orange, yellow, green, turquoise, etc. (Xiao, Wang and Felleman, 2003). It is therefore reasonable to infer that coloured glasses or coloured overlays will change the distribution of the firing pattern within the visual cortex. Although von Kries adaptation (Hunt, 1991) will counteract the effects of the filters such adaptation is
unlikely to normalise the effects completely, given the saturation of the filters currently used.

One possible mechanism for the benefits of coloured filters is therefore that they redistribute the activity within the visual cortex in such a way as to reduce the amount of excitation in locally hyperexcitable regions.

Predictions of visual stress theory

This hypothesis has received support in recent studies of cortical activation in patients with migraine. As already mentioned, the blood oxygenation level-dependent (BOLD) response to epileptogenic patterns is abnormally large in migraineurs (Huang et al, 2003). In preliminary studies we have recently investigated a patient with migraine and shown that glasses with a therapeutic colour reduce the size of the response, whereas control colours have no effect. Figure 2 shows preliminary data from the patient examined using the procedure described by Huang et al (2003). The BOLD response in visual areas V1, V2 and V3 is shown separately for spectacles with the prescribed colour selected by the patient as reducing distortions of print (a shade of green), a control colour (one that was similar but did not reduce the distortions, a shade of turquoise) and grey lenses of similar transmission. The glasses with the prescribed colour selectively reduce the response at epileptogenic spatial frequencies in V3. This is close to an area (V3a) where Hadjikhani et al (2001) have reported a focal onset of migraine aura. This is a surprising finding: the BOLD response has been normalised at the appropriate spatial frequencies, with the appropriate colour, and in the appropriate putative cortical area. It will be interesting to see if it replicates.

Figure 2. The percentage change in BOLD response when a pattern of stripes is viewed, shown as a function of spatial frequency (in cycles per degree) for three putative visual areas, V1, V2 and V3.

Notes: The data are for an individual who suffers from migraine. The triangles show the change activation when green precision tinted lenses were worn ($u' = 0.144$, $v' = 0.514$). The squares show the activation when grey lenses having the same photopic transmission were worn, and the crosses show the activation with lenses of a control colour (turquoise-blue, $u' = 0.145$, $v' = 0.469$). The photopic transmission of all lenses was between 27% and 30% and the mean luminance of the display was 5.4 cd.m $^{-2}$. The scanning methods were described by Huang et al, 2003.
As would be predicted by the theory of visual stress, the children who benefit most from coloured filters usually have migraines, or have migraine sufferers in the family (Maclachlan, Yates and Wilkins, 1993). As would also be predicted, the benefits of filters extend to a variety of disorders of the central nervous system in which visual symptoms are present and in which the cortex is hyperexcitable. These disorders include not only photosensitive epilepsy (Wilkins et al, 1999) and migraine (Evans, Patel and Wilkins, 2002; Wilkins, Patel, Adjamian and Evans, 2002), but perhaps also multiple sclerosis, autism and head injury, although the data are currently only anecdotal in the case of the last three disorders.

It would therefore appear that the idea of visual stress as a consequence of cortical hyperexcitability is a theoretical notion that has widespread explanatory utility. It explains the widespread prevalence and links with migraine and provides a plausible explanation for the effects of coloured filters in various neurological diseases. According to this conception the colour that reduces distortions is specific for each individual because the distribution of hyperexcitability in the visual cortex varies from one person to another. Unfortunately the theory is currently insufficient to predict which colours suit which people. It does, however, make important predictions concerning symptoms when reading and text design.

Measurement of visual stress

Visual stress can be measured in many different ways. There are trait measures that register the overall susceptibility of an individual, and measures that record the state of visual stress in the here and now. The trait can be assessed using symptom questionnaires including the one proposed by Irlen (1991) and the one evaluated by Conlon and Hine (2000). The questions include those concerning distortions of print, and typical behaviour during reading. There is generally a reasonable agreement between state and trait measures. Individuals with high levels of state or trait visual stress are generally slower readers, although the correlation with reading speed is rather weak, explaining typically about 10–20% of the variance in reading speed (Conway, unpublished data). To assess the state of visual stress we used a simple set of questions asked of the individual while he or she observed a page of text. Though much shorter, this test explained almost as much variance in reading speed as an extensive questionnaire (unpublished data). The questions we used were ‘Do the letters stay still or do they move’; ‘Are they clear or are they blurred’; ‘Is the page too bright, not bright enough, or just about right’; ‘Does it hurt your eyes to look at the page?’ (Wilkins et al, 2001).1

Consequences for visual aspects of reading

The consequences of visual stress for reading can most readily be assessed using the Rate of Reading Test, a test in which the reader is required to read aloud randomly ordered common words as quickly as possible. The use of common words ensures that the words are familiar to poor readers. The use of a random word order ensures that each word cannot be guessed from context, but has to be seen to be read. The fact that the text is

1A fifth question ‘Are the words too close together or far enough apart’ was not significantly associated with an increase in reading speed with a coloured overlay (Wilkins et al, 2001).
meaningless has the advantage that readers who make errors of omission and transposition are usually unaware of their errors, and they can perform the test without a sense of failure. The text is printed in small and closely spaced text to increase the visual stress, enabling its effects on reading speed to be measured in the short term (Wilkins, Jeanes, Pumfrey and Laskier, 1996).

Individuals who report benefits from coloured filters usually read more quickly on the Rate of Reading Test when the text is appropriately coloured with a filter. The increase in reading speed can predict the individuals who will subsequently use the coloured filter without prompting (Jeanes et al., 1997). In more naturalistic reading tasks the effects of visual stress are more difficult to measure and take longer to appear (Tyrrell et al., 1995). Nevertheless, increases in reading speed on the Rate of Reading Test predict improved performance of tasks requiring silent reading for comprehension (Wilkins, 2002).

Although visual stress affects reading speed, and these effects can be measured in the short term, the stress has longer-term effects on the willingness of children to expose themselves to the effort and discomfort of reading.

Controversy concerning colour

The use of colour to treat visual stress in reading is a recent development, and one with a highly controversial beginning. The controversy continues in the United States to this day where established medical opinion still has it that treatment with coloured filters offers no more than a placebo. Because the assessment is subjective and the outcome of treatment is also subjective (a change in symptoms), the client’s belief is likely to have a major influence on the results of treatment. Nevertheless, we have shown that individuals who complain of visual stress have an abnormality of accommodation, and this cannot be attributed to belief (Simmers, Gray and Wilkins, 2001). We have also used double-masked studies to eliminate contamination with placebo effects, as will now be described.

The double-masked studies were possible because of the techniques available with the Intuitive Colorimeter. During assessment for coloured filters using the Intuitive Colorimeter, the eyes are colour adapted, and it is possible to assess the effects of coloured light without the client being aware of the precise chromaticity under investigation. It is therefore possible to check the consistency of colour choice without interference from the client’s belief. As a result of colour adaptation in the colorimeter, the coloured light will appear less saturated than expected. Lenses that provide the colour when worn under white light will appear strongly coloured. Any knowledge of the effective colour can therefore be separated from the colour appearance of the lenses, and this allows double-masked controlled studies to be undertaken. The lenses with the appropriate colour can be compared with those having a slightly different colour under conditions in which all the participants (the patients and those caring for them) are unaware as to which pair of spectacles has the appropriate tint. Two studies with such a design have been undertaken, the first in children with reading difficulties (Wilkins et al., 1994), and the second in adults with migraine (Wilkins et al., 2002). Both have shown reduced symptoms of visual stress with the appropriately coloured lenses, but not with other lenses having a slightly different colour.

Precision required

Recent studies of individuals who use coloured filters have shown that the effects of colour on reading speed are very specific. Little benefit remains once the colour
has departed from optimum by a colour difference (CIE $L^* U^* V^*$) of about 100 (Wilkins, 2003). The implication of this finding is that tinting systems that do not offer a very wide range of tint options cannot provide tints that will cater optimally for all individuals.

**Impact on reading behaviour and development**

Some children begin to fail in reading at the age of seven when they are put off reading by the discomfort and distortions they experience when looking at print. This is the age at which the text children are required to read starts to decrease in size and to resemble the characteristics of an epileptogenic pattern. The arrows on the x-axis of Figure 1 show the characteristics of the lines of a typical sample of text designed for five-year-olds, considered as if it were a striped pattern. Notice that the spatial frequency and average Michelson contrast of a line are outside the epileptogenic/aversive range. These may be compared with equivalent values for adult text (shaded areas), which are within the epileptogenic range. The difference suggests that adult text is more likely to evoke distortions and discomfort, as indeed appears to be the case.

Hughes and Wilkins (2000) made versions of the *Rate of Reading Test* in typography typical of children’s reading schemes. They varied the typeface size and letter spacing so as to resemble the reduction in size and spacing that occur with increasing reading age. They showed that children with a reading age of 7–9 read the text designed for five-year-olds more quickly than material designed for seven-year-olds, even though they were presumably more familiar with the latter. Clearly the differences in reading speed could only be attributed to the differences in the characteristics of the type because the reading tasks were otherwise identical. Their findings would seem to suggest that the text in children’s reading schemes gets too small too quickly. These authors also showed that children who reported illusions in patterns of stripes were disproportionately affected by the change in typeface, which would be consistent with the explanation in terms of visual stress outlined above. More recent unpublished data have confirmed the finding that reading speed in children age 7–8 increases with typesize, and further that the increase in speed continues at least up to 18-point text (Rachel G. Evans, unpublished data).

**Extension of the theory**

Figure 1 provides a specification of the stressful properties of patterns, but only patterns that are actually stripes. Although we have shown that the illusions reported in text resemble those reported in stripes and are similarly dependent upon the spatial frequency of the successive lines, there is considerable work to be done before the theory of visual stress is sufficiently specified to be applicable to typographic design. It is unclear whether it is the energy in the fundamental Fourier component of the text image that is principally responsible for visual stress or whether, as seems likely, higher frequency components also play a role. The text can provide patterns of stripes not only from the horizontal lines but also from the successive vertical strokes that form letters. In preliminary unpublished work students were asked to report the distortions seen in isolated lines of text printed in various fonts. The reports of distortions appear to be related to the periodicity of the letter strokes, as assessed by peaks in the horizontal autocorrelation function. Further work is under way to explore this relationship further.
Implications for the assessment of reading

Although the theory of visual stress cannot provide a precise indication of optimal typeface design, it has a number of predictions that are worth exploring, particularly when it comes to the design of reading tests. For the most part, the typographic characteristics of text in reading tests resemble those in other text for children of equivalent reading age. For example, in the *Neale Analysis of Reading Ability* (Neale, 1997) the height of the central body of the letters (x-height) varies with reading age according to the large points in Figure 3, which also shows the x-height in a sample of reading schemes for children. As can be seen, the text in the *Neale Analysis* is typical of the reading material the child is likely to encounter. It therefore has the advantage of presenting realistic material, even though, as reported in a previous section, Hughes and Wilkins (2000) have found that such material is too small and provides an underestimate of potential ability in children in the age range 7–10. The *Wechsler Objective Reading Dimensions* (Wechsler, 1993) uses text with an x-height of 4 mm except for children over the age of 7 years 3 months, when the x-height reduces to 2.9 mm. The text is larger than conventionally experienced, and this may mean atypically high scores relative to the reading of conventional material. The same cannot be said for the *Primary Reading Test* (France, 1981), which is designed for an age range of 6–12 years and yet uses text with an

![Figure 3](image_url). The height of the central body of the letters (x-height) in samples of text from children’s reading schemes (Courtesy of P. Greenfield and A. Cove-Smith). The large points show the x-height of letters in the Neale Analysis of Reading (Neale, 1997).
x-height of only 2.4 mm. As can be seen from Figure 3, this size is customarily used only for age eight and above, and even then, as we have seen, it is too small. At 2.4 mm the x-height is smaller than in any of the text in the Wechsler Objective Reading Dimensions, which is designed to cater for age groups up to 17 years.

A different problem arises with the Word Reading Card of the British Ability Scales. Although the x-height is similar to that of material typically used in the age group, the test uses an unusually bold typeface in which the letter strokes have a width similar to that of the distance between them, forming stripes (i.e. a pattern with high autocorrelation). Some perceptual distortion of ‘striped’ words, like jump or running is likely in individuals predisposed to visual stress, according to the preliminary results concerning autocorrelation described above.

There is quite a different problem with the NFER-Nelson Group Reading Test (NFER Nelson, 1992): the various forms are printed on backgrounds of different colour and in texts of different colour, and as we have seen from the effects of coloured filters described above, this is likely to be a disadvantage for some children, and an advantage for others, depending on the colours from which they benefit.

Conclusion

Insufficient attention is paid to visual factors in reading, and specifically so as regards the design of reading tests. The text in children’s reading material gets small and closely spaced at a rate that disadvantages most children, and especially those subject to visual stress. The text in many reading tests in widespread use is poorly designed – too small and more closely spaced than is conventional even for children’s reading material, which is already too small. Given that so many children in mainstream education report distortions of text, and given that these distortions are associated with a reduction in reading speed, it seems sensible, indeed an obligation, to discover how these distortions can be reduced. Although the use of coloured filters is one method, there is not one colour that suits all children and it would be preferable to design text so that distortions were less apparent in the first place. In the author’s experience many children report distortions for small closely spaced text more than for large widely spaced text, but a formal study of ways of minimising distortions at a given letter density is overdue. Typographic practice needs to change to accommodate the fact that children’s vision for text is subject to distortion, and the fact that children under the age of ten need far larger text than they are currently required to read.

References


**Address for correspondence:** Arnold Wilkins, Department of Psychology, University of Essex, Colchester CO4 3SQ, UK